

SYNERGISTIC USE OF SENTINEL-1 AND SENTINEL-2 TIME SERIES FOR POPLAR PLANTATIONS MONITORING AT LARGE SCALE

Y. Hamrouni^{1,2,*}, É. Paillassa³, V. Chéret¹, C. Monteil¹, D. Sheeren¹

¹ Université de Toulouse, INRAE, UMR DYNAFOR, Castanet-Tolosan, France

² Conseil National du Peuplier, Paris, France

³ Centre National de la Propriété Forestière, Institut pour le Développement Forestier, Bordeaux, France

Commission III, ICWG III/IVc

KEY WORDS: Poplar plantations, Active learning, Large scale, Mapping, SAR, Stand age

ABSTRACT:

The current context of availability of Earth Observation satellite data at high spatial and temporal resolutions makes it possible to map large areas. Although supervised classification is the most widely adopted approach, its performance is highly dependent on the availability and the quality of training data. However, gathering samples from field surveys or through photo interpretation is often expensive and time-consuming especially when the area to be classified is large. In this paper we propose the use of an active learning-based technique to address this issue by reducing the labelling effort required for supervised classification while increasing the generalisation capabilities of the classifier across space. Experiments were conducted to identify poplar plantations in three different sites in France using Sentinel-2 time series. In order to characterise the age of the identified poplar stands, temporal means of Sentinel-1 backscatter coefficients were computed. The results are promising and show the good capacities of the active learning-based approach to achieve similar performance (Poplar F-score $\geq 90\%$) to traditional passive learning (i.e. with random selection of samples) with up to 50% fewer training samples. Sentinel-1 annual means have demonstrated their potential to differentiate two stand ages with an overall accuracy of 83% regardless of the cultivar considered.

1. INTRODUCTION

Poplar (*Populus* spp.) is a fast-growing and wood producing tree which is considered as an important economic resource due to the increasing demand for its by-products such as lightweight packaging and plywood. In France, poplar cultivation is a key local industry. However, over the last two decades this sector has faced several economic, social and environmental upheavals that have led to a continuous decrease in planted surfaces. The future of poplar depends mainly on areas replanted after harvesting. It is therefore crucial to have spatially explicit information on newly planted and lost areas, which provides essential baseline data for industrial and socio-economic dynamics. Accurate and updated maps of poplar plantations are not yet available at the national scale. The update rate of the French forest database (10 to 20 years) is unsuitable for this species because of its short rotation cycle (15 years on average). The availability of Sentinel data with high spatial and temporal resolutions has provided new opportunities for identifying and characterising poplar plantations over large areas.

Several works have already demonstrated the potential of remotely sensed data for mapping plantations and few have particularly focused on poplars like (Chardenon, Flouzat, 1981), (Borry et al., 1993) and (Heyman et al., 2003). Nevertheless, in most cases, the studies were conducted at a local scale and the reported performances were highly dependent on the data used to train and validate the classification models. Consequently, these models generally exhibit limited generalisation capabilities and their application over large areas remains challenging.

In this paper, Sentinel-2 optical time series are used to differentiate poplars from other deciduous species in three study sites.

* Corresponding author

In order to minimise the number of samples required for training and to build a generic model tailored to the different study sites, we propose the use of a transfer learning-based approach, namely Active Learning. Secondly, we were interested in characterising the age of poplar plantations. Due to the sensitivity of SAR to vegetation structure, we explored the potential of Sentinel-1 data to distinguish between two main stand ages.

2. POPLAR IDENTIFICATION WITH SENTINEL-2 TIME SERIES

In this section, we investigate the potential of Sentinel-2 data to identify poplar plantations. We first assess its ability to recognise the plantations at a local scale (Sentinel-2 tile) (Section 2.3). We secondly focus on the adaptation of the resulting local classifiers to distinct areas with active learning (Section 2.4).

2.1 Study area

Three poplar sites with contrasting silvicultural practices and climatic conditions were chosen with the forest partners. They are located in northeastern, central and southwestern France and are covered by three Sentinel-2 tiles with a surface area of 100 km² each (tile codes respectively: 31UEQ, 30TYT and 31TCJ) (Figure 1).

2.2 Reference data

Samples for deciduous forest classes were retrieved from the national forest database (BD Forêt®IGN). A photo interpretation was conducted to collect poplar references in order to ensure updated samples. It should be noted that the poplar samples created correspond to plantations older than two years

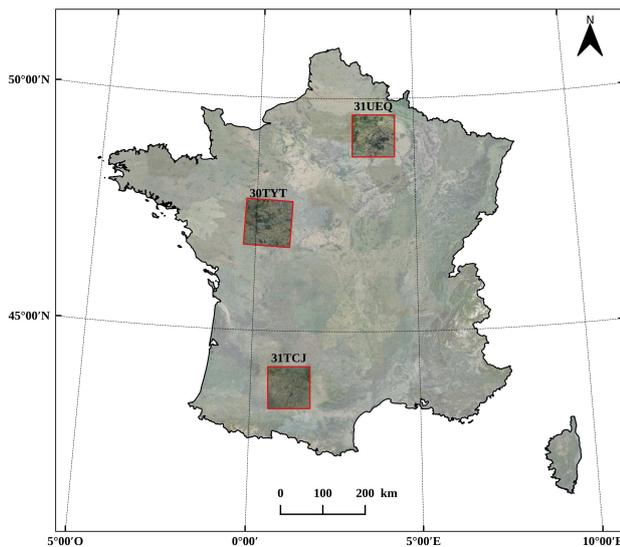


Figure 1. Overview of the three study sites represented by three Sentinel-2 tiles: 31UEQ, 30TYT and 31TCJ.

old because below this limit, the plantations cannot be differentiated. Regardless of the photo interpretation's bias, this phase is considerably time-consuming because it needs to be performed repeatedly for each study site until achieving a statistically representative data set.

2.3 Tile-scale classification with Sentinel-2

Sentinel-2 optical time series from 2017 were downloaded from Theia platform over the three study tiles (Table 1). They are level 2A products (surface reflectance) provided with a cloud mask and after atmospheric correction.

Tile code	Relative orbit number	No. dates in 2017
31UEQ	51	26
30TYT	94	34
31TCJ	51	36

Table 1. Properties of the Sentinel-2 tiles used in the study

A temporal gap-filling was applied to the Sentinel-2 images in order to replace cloudy pixels with an interpolated value based on the nearest cloud-free pixels of the temporal time series (Inglada, 2016). The time series were then resampled at 10-m spatial resolution with a 10-day time step common to all tiles. In order to assess the potential of Sentinel-2 to recognise poplar plantations at local scale in the same tile, a random forest (RF) supervised classification was performed in each tile independently.

Reference polygons were randomly split into 50% for training and 50% for testing with a stand-based stratified random sampling. Sampling was repeated 30 times in order to account for variability related to random selection. However, with this approach we assume that samples are available everywhere which is not practical at large scale. It was necessary to define a more automatic process able to detect poplars with good classification performances but with a minimum of training samples. It is in this context that we propose the use of Active Learning (AL).

2.4 Towards a large-scale generalisation with AL

Over the past decade, AL has received the attention of the remote sensing community (Tuia et al., 2011). It has mostly been applied to select a reduced set of training samples required for classification tasks (Tuia et al., 2009), (Ma et al., 2018). Only a few works have addressed transfer learning between two distant regions (Matasci et al., 2012), (Alajlan et al., 2014).

AL is based on the assumption that an algorithm is able to reach better classification results not by focusing on the number of samples (randomly selected), but rather on their quality while choosing the most relevant ones (Settles, 2012). The AL process uses a ranking criterion to select in an intelligent way the most informative samples from a pool of candidates. The selection is guided by the algorithm needs which is iteratively enriched with new and carefully chosen training samples until certain predefined stopping criteria are met (e.g. maximum score or maximum number of iterations).

In this study, AL was performed between the three Sentinel-2 tiles (in pairs) along the six possible directions of learning (north-east to south-west, south-west to center, etc.). In each case, the process started with a trained classifier on a first tile (source) and is used to predict a second one (target). At each iteration 10 samples were queried until reaching a maximum of 1000 samples (100 iterations). Uncertainty was used as an informativeness criterion to select the best samples (i.e. the most uncertain instances are considered to be the most informative) and two measures were tested: entropy (H) and margin sampling (MS). As detailed in Equation 1, while entropy uncertainty measure takes the probabilities of belonging to all the model classes into account, the MS metric considers only the first two most probable labels (Equation 2).

$$x_H^* = \operatorname{argmax}_x - \sum_y P_\theta(y | x) \log P_\theta(y | x) \quad (1)$$

$$x_{MS}^* = \operatorname{argmin}_x [P_\theta(\hat{y}_1 | x) - P_\theta(\hat{y}_2 | x)] \quad (2)$$

where x^* = the best instance selected
 y = all possible labels of x
 P_θ = the probability value under the model θ
 \hat{y}_1, \hat{y}_2 = first and second most probable labels

In this paper we present only the results obtained with MS uncertainty metric. For comparison purposes, the AL process was run against a classifier using the same number of samples but randomly selected.

3. STAND AGE RETRIEVAL WITH SENTINEL-1

Several studies have demonstrated the sensitivity of SAR information and in particular C-band data to monitor vegetation dynamics due to their sensitivity to structure (Bouman, Hoekman, 1993), (Satalino et al., 2014), (Veloso et al., 2017), (Vreugdenhil et al., 2018). Specifically, the VH/VV ratio has proved to be of great interest for monitoring the vegetation growth cycle and showed a strong correlation with NDVI (Vreugdenhil et al., 2018), (Frison et al., 2018).

In this section, we investigate the potential of Sentinel-1 for stand age assessment.

Field data for 94 poplar plots located in the south of France were provided by the forest partners (Table 2). For each plot, the plantation year as well as the cultivar type were reported. The poplar stands are between two and seven years old, but not all ages are available for each cultivar.

Cultivar	No. plots	No. pixels
Koster	34	3083
I45/50	28	3708
I214	22	934
Soligo	7	589
Raspalje	3	760

Table 2. Field data summary

Sentinel-1 Ground Range Detected (GRD) data from 2017 were downloaded from the French distribution and processing platform (PEPS) over the study zone. They were acquired in Interferometric Wide swath mode (IW) in an ascending orbit direction (Relative orbit 30). The images were calibrated, orthorectified and filtered from speckle noise with a spatio-temporal filter using the S1-Tiling tool (Koleck, 2017). Annual means of radar backscatter coefficients were then calculated for VV and VH polarisations as well as their ratio VH/VV. A color composite image of the three derived means is shown in Figure 2 with examples of poplar plots at different ages.

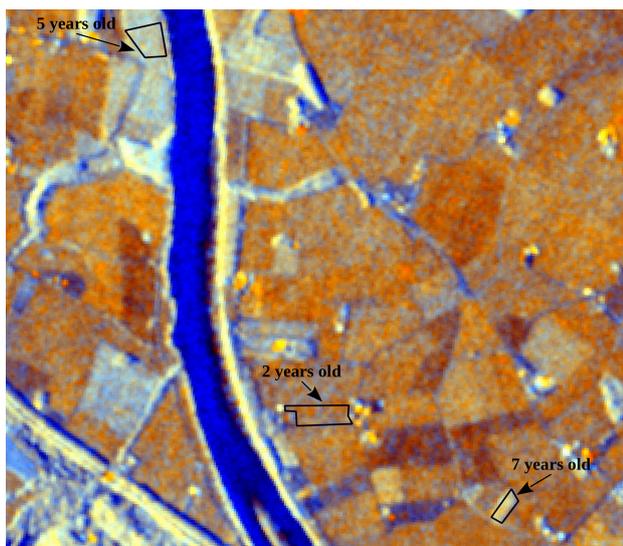


Figure 2. Multi-polarisation color composite of annual means of Sentinel-1 backscatter coefficients (Red: σ_{VV}^0 , Green: σ_{VH}^0 , Blue: $\sigma_{VH}^0/\sigma_{VV}^0$).

Considering the lack of representativeness of age classes and the scarcity of plots per cultivar, a supervised random forest classification was carried out on the entire data set in order to identify two main age groups: young plantations (two to four years old) and mature plantations (four to seven years old). The classification was performed at two scales: plot scale (the plot is defined by the average value of all its pixels) and pixel scale (all pixels are considered regardless of the plot).

As in the previous section, samples were randomly split into 50% for training and 50% for testing with a stand-based stratified random sampling. Sampling was repeated 30 times in order to account for variability related to random selection.

4. RESULTS AND DISCUSSION

4.1 Sentinel-2-based classification: high capacity to identify poplar plantations at local and global scales

The random forest classification results are reported in Table ?? for each of the three study tiles. As mentioned earlier, the aim is to assess the potential of Sentinel-2 data to discriminate poplar plantations from the other deciduous species at a local scale, in this case the tile extent.

Tile code	No. samples ¹ per class	No. classes ²	OA (*30)	Poplar F-score (*30)
31UEQ	1250	6	73.7%	89.5%
30TYT	2000	6	74.9%	99.3%
31TCJ	3850	6	80.0%	97.9%

¹ Training samples represent 50% of the available reference data. The number is given in pixels of 10m² area.

² The classes are the same for both central (30TYT) and southwestern (31TCJ) tiles: *poplar*, *locust*, *chestnut*, *oak*, *open mixed forest* and *closed mixed forest*. The *chestnut* is absent in the northeastern (31UEQ) tile but there is instead the *beech* class.

Table 3. Local classification results for each Sentinel-2 tile averaged over 30 independent repetitions.

The local classification results of Table ?? showed high capacity of Sentinel-2 data to identify poplars with F-score values of 90%, 99% and 98% respectively for the north-eastern, central and south-western tiles. However, when we tested the predict-

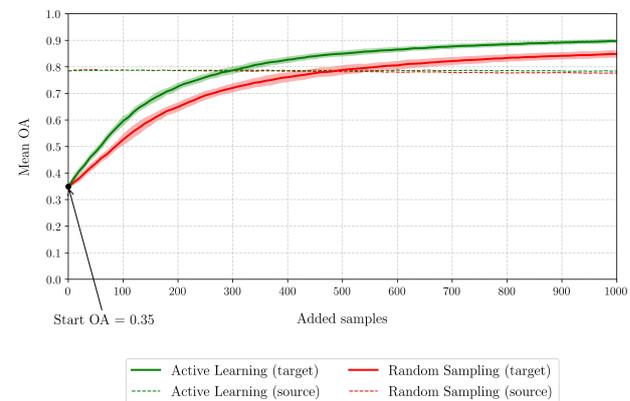


Figure 3. Changes in the average OA scores on the south-western target tile according to the number of added samples with the active learning (green) and random (red) models. The initial classifier is trained on the north-eastern source tile.

ive capabilities of these different local models, low accuracies were obtained. This is due to the non-stationarity of class distributions between the different study tiles. When we adapted the initial models with active learning based on the margin sampling uncertainty metric (AL_{MS}), the performance increased rapidly as samples were added but considerably faster than with a random selection of samples. The OA varied according to the adopted direction of the transfer (i.e. according to the tile on which the initial model is trained) but in all cases its values were up to 5.5% higher with AL_{MS} . An example is given in Figure 3 for a transfer from the north-eastern to the south-western tile. As it can also be observed, the OA values computed on the initial tile (source) remain fairly constant while adding new samples from the target for both AL_{MS}

and random selection. After querying new samples, the initial model improved its generalisation capabilities to better classify the target tile while still performing as well on the source tile.

When we focused on the poplar class and for the same F-score, the number of samples randomly selected was about eight times higher than those queried by AL_{MS} . Furthermore, the number of AL_{MS} queries was related to the starting F-score. As shown in the example of Figure 4, when this value was high enough (i.e. the initial model was able to accurately recognise the poplar plantations of the second tile), very few poplar samples were selected by AL_{MS} (green bars) unlike the random strategy which selected many samples (red bars) regardless of the initial performance (poplar F-score = 95%). AL_{MS} minim-

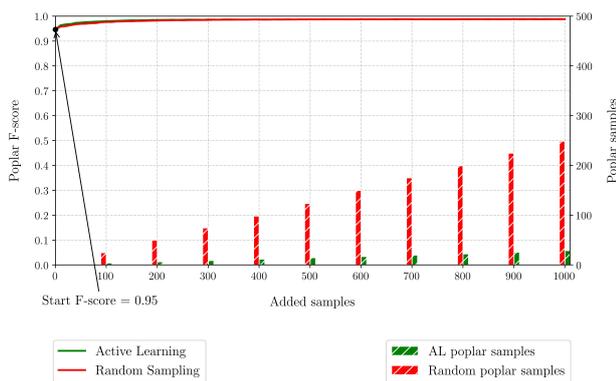


Figure 4. Changes in the average poplar F-score on the south-western target tile according to the number of added samples. The initial classifier is trained on the north-eastern source tile.

ised the need to label target poplar samples without sacrificing the classification performance. As reported in (Di, Crawford, 2011), when the class accuracy is high, the active learner avoids querying irrelevant samples.

Similarly to the poplar class, we have assessed the contribution of AL for the remaining deciduous classes. We particularly noted its potential regarding the hardest classes to discriminate (i.e. missing classes from the initial model or highly overlapping classes) as already demonstrated in (Joshi et al., 2009). The AL results based on entropy uncertainty metric were worse than with margin sampling. Indeed, as entropy takes the probabilities of belonging to all the classes into account, the AL selection is influenced by low probabilities of unimportant classes and is consequently less robust to noise.

4.2 SAR sensitivity to stand age classes

The annual means of the backscatter coefficients (VV, VH and VH/VV) were analysed and their relation to field data (stand ages) was evaluated at both the plot and pixel and scales. Determination coefficients (r^2) describing this relationship are reported in Table 4.

Whether the analysis was performed for plots or pixels, a weak correlation was observed with the VV polarisation. The VH/VV ratio was however capable of reproducing 67% of observed variability in the plots and 42% when considering all the pixels. These results are inline with the classification results reported in Table 5.

r^2	VV	VH	VH/VV
Plots	0.05	0.59	0.67
Pixels	0.06	0.43	0.42

Table 4. Determination coefficients (r^2) between annual means of Sentinel-1 backscatter coefficients (dB) and the field-derived stand ages.

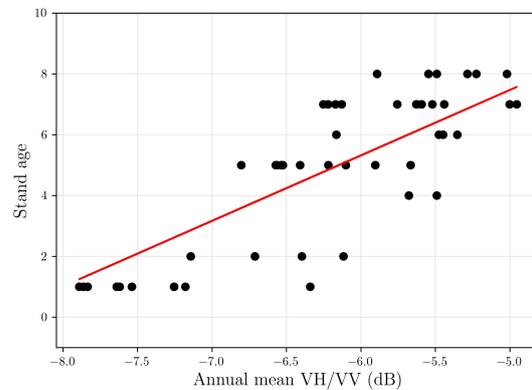


Figure 5. Plots-based correlation between field stand ages and VH/VV annual backscatter coefficient ($r^2=0.67$).

Classification results with the three bands VV, VH and their ratio VH/VV showed good performance at both plot and pixel scale with overall accuracy (OA) values of 82.5% and 83.1% respectively. When only the VH/VV ratio was considered, the results were slightly lower: -2.8% and -5% of OA for respectively plot and pixel scales. Although the best performances were recorded with the combination of the three bands, this highlighted the interest of the VH/VV ratio to reproduce very close results.

Features	Classification	OA _(*30)	F-score _(*30)
VV+VH +VH/VV	Plot-based	82.5%	81.7%
	Pixel-based	83.1%	83.0%
VH/VV	Plot-based	79.7%	79.6%
	Pixel-based	78.1%	78.0%

Table 5. Pixel and plot-based classification results with annual SAR backscatter coefficients.

The observed misclassifications could be related to the differences between the development stages according to the considered cultivar. Indeed, for a given cultivar the growing can be fast since the first year of plantation whereas for others it can take three to five years to get a detectable canopy. A classification was therefore carried out for each cultivar. Overall accuracies (pixel-based classification) ranged from 65% to 99% but as reported before, the number of samples available per cultivar was limited and not sufficiently representative of the two stand age classes.

5. CONCLUSION

In this letter, we proposed a combined use of optical and SAR imagery to monitor poplar plantations. The active learning approach showed promising results on Sentinel-2 data for identifying poplar plantations in two contrasting study sites by taking

advantage of the initial knowledge gained during local classification tasks and querying only if necessary a reduced set of relevant training samples. It is important to note that the AL may be penalised by noisy references that are most likely to be selected by the algorithm. Compared to entropy, the margin sampling uncertainty measure is more robust to noise. These results are very promising and have opened up interesting leads for national transfer. Temporal means of Sentinel-1 backscatter coefficients demonstrated their sensitivity to the plantation structure and their potential to differentiate two main stand ages particularly with the VH/VV ratio. Sentinel-1 temporal information could be further leveraged by filling the gaps in optical series during cloudy periods, creating forest mask based on seasonal backscattering means or by calculating phenological features to discriminate cultivar groups.

REFERENCES

- Alajlan, N., Pasolli, E., Melgani, F., Franzoso, A., 2014. Large-Scale Image Classification Using Active Learning. *IEEE Geoscience and Remote Sensing Letters*, 11(1), 259–263.
- Borry, F. C., de Roover, B. P., Leysen, M. M., de Wulf, R. R., Goossens, R. E., 1993. Evaluation of SPOT and TM Data for Forest Stratification: A Case Study for Small-Size Poplar Stands. 31(2), 483–490.
- Bouman, B. A. M., Hoekman, D. H., 1993. Multi-Temporal, Multi-Frequency Radar Measurements of Agricultural Crops during the Agriscatt-88 Campaign in The Netherlands. *International Journal of Remote Sensing*, 14(8), 1595–1614.
- Chardenon, J., Flouzat, G., 1981. The application of remote sensing to poplar growing: identification and inventory of poplar groves, prediction of timber production; France, Italy. *Revue Forestiere Francaise*, 33(6), 478–493.
- Di, W., Crawford, M. M., 2011. Critical class oriented active learning for hyperspectral image classification. *2011 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, Vancouver, BC, Canada, 3899–3902.
- Frison, P.-L., Fruneau, B., Kmiha, S., Soudani, K., Dufrêne, E., Le Toan, T., Koleck, T., Villard, L., Mougin, E., Rudant, J.-P., 2018. Potential of Sentinel-1 Data for Monitoring Temperate Mixed Forest Phenology. *Remote Sensing*, 10(12), 2049.
- Heyman, O., Gaston, G., Kimerling, A., Campbell, J., 2003. A Per-Segment Approach to Improving Aspen Mapping from High-Resolution Remote Sensing Imagery. *Journal of Forestry*, 101(4), 29–33.
- Inglada, J., 2016. Otb Gapfilling, A Temporal Gapfilling For Image Time Series Library.
- Joshi, A. J., Porikli, F., Papanikolopoulos, N., 2009. Multi-class active learning for image classification. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2372–2379.
- Koleck, T., 2017. S1TemporalSeries. <http://tully.ups-tlse.fr/koleckt/s1tiling>.
- Ma, L., Fu, T., Li, M., 2018. Active Learning for Object-Based Image Classification Using Predefined Training Objects. *International Journal of Remote Sensing*, 39(9), 2746–2765.
- Matasci, G., Tuia, D., Kanevski, M., 2012. SVM-Based Boosting of Active Learning Strategies for Efficient Domain Adaptation. 5(5), 1335–1343.
- Satalino, G., Balenzano, A., Mattia, F., Davidson, M. W. J., 2014. C-Band SAR Data for Mapping Crops Dominated by Surface or Volume Scattering. 11(2), 384–388.
- Settles, B., 2012. Active Learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 6(1), 1–114.
- Tuia, D., Ratle, F., Pacifici, F., Kanevski, M., Emery, W., 2009. Active Learning Methods for Remote Sensing Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 47(7), 2218–2232.
- Tuia, D., Volpi, M., Copa, L., Kanevski, M., Munoz-Mari, J., 2011. A Survey of Active Learning Algorithms for Supervised Remote Sensing Image Classification. 5(3), 606–617.
- Veloso, A., Mermoz, S., Bouvet, A., Le Toan, T., Planells, M., Dejoux, J.-F., Ceschia, E., 2017. Understanding the Temporal Behavior of Crops Using Sentinel-1 and Sentinel-2-like Data for Agricultural Applications. *Remote Sensing of Environment*, 199, 415–426.
- Vreugdenhil, M., Wagner, W., Bauer-Marschallinger, B., Pfeil, I., Teubner, I., Rüdiger, C., Strauss, P., 2018. Sensitivity of Sentinel-1 Backscatter to Vegetation Dynamics: An Austrian Case Study. *Remote Sensing*, 10(9), 1396.