Evaluation of Time Series Distance Functions in the Task of Detecting Remote Phenology Patterns

José C. Conti*, Fabio A. Faria†, Jurandy Almeida§†, Bruna Alberton‡,
Leonor P. C. Morellato‡, Luiz Camolesi Jr.*, and Ricardo da S. Torres†
* Faculty of Technology, University of Campinas – UNICAMP
13484-332, Limeira, SP – Brazil
Email: conti30@gmail.com,camolesi@ft.unicamp.br
† RECOD Lab, Institute of Computing, University of Campinas – UNICAMP
13083-852, Campinas, SP – Brazil
Email: ffaria@ic.unicamp.br, rtorres@ic.unicamp.br
‡ Phenology Lab, Dept. of Botany, São Paulo State University – UNESP
13506-900, Rio Claro, SP – Brazil
Email: bru.alberton@gmail.com, pmorella@rc.unesp.br
§ Institute of Science and Technology, Federal University of Sao Paulo – UNIFESP
12231-280, São José dos Campos, SP – Brazil
Email: jurandy.almeida@unifesp.br

***Abstract*—Phenology is the study of periodic natural phenomena and their relationship to climate. Usually, phenology studies consider the identification of patterns on temporal data. In those studies, several phenological change patterns are often encoded in time series for analysis and knowledge extraction. In this paper, we evaluate the effectiveness of several time series similarity functions in the task of classifying time series related to phenological phenomena characterized by near-surface vegetation indices extracted from images. In addition, we performed a correlation analysis to identify potential candidates for combination.**

I. Introduction

Plant phenology, the study of recurrent life cycles events and its relationship to climate, is a key discipline in climate change research [1]. To increase the range of study sites and species and the accuracy of phenological observations, digital cameras have been successfully applied as multi-channel imaging sensors, providing measures to estimate changes on phenological events, such as leaf flushing and senescence [2], [3], [4].

We have monitored leaf changing patterns of a tropical cerrado-savanna vegetation by taken daily digital images [5]. For that, we extracted leaf color information from the RGB (red, green, and blue) channels and correlated the changes in pixel levels over time with leaf phenology patterns [5]. The image analysis was conducted by defining regions of interest (ROI) based on the random selection of plant species identified in the digital image [4], [5].

Time series associated with each ROI have been obtained, raising the need of using appropriate tools for mining patterns of interest [6], [7], [8]. In this context, it is necessary to define a quantitative measure of the dissimilarity between time series, aiming at approximating the perceptual dissimilarity as well as possible. This has fundamental implications for the understanding of texture pattern and color perception of plant species.

In this paper, we aim to identify a consistent measure of distance, or dissimilarity, between time series obtained from phenological observations. Our main goals are: i) to determine which color channel is better for characterizing plant species; ii) to analyze the impact of the sunshine on the discriminability of time series; and iii) to evaluate the individual responsiveness of plant species with respect to distance functions.

The remainder of this paper is organized as follows. Section II discusses the methodology adopted for acquiring time series. Section III presents the evaluated distance functions and shows how to apply them to compare time series. Section IV reports the results of our experiments and compares the performance of distance functions for different scenarios. Finally, we offer our conclusions and directions for future work in Section V.

II. Time Series Acquisition

The near-remote phenological system was set up in a 18m tower in a Cerrado *sensu stricto*, a savanna-like vegetation located at Itirapina, São Paulo State, Brazil. A digital hemispherical lens camera (Mobotix Q24) was setup at the top of the phenology tower, attached in an iron arm facing northeast.

The first data collection from the digital camera started on 18th August 2011. We set up the camera to automatically take a daily sequence of five JPEG images (at 1280 $\times$ 960 pixels of resolution) per hour, from 6:00 to 18:00 h (UTC-3). The present study was based on the analysis of over 2,700 images (Figure 1), recorded at the end of the dry season, between August 29th and October 3rd 2011, day of year 241 to 278, during the main leaf flushing season [5].

1051-4651/14 $31.00 © 2014 IEEE
DOI 10.1109/ICPR.2014.539

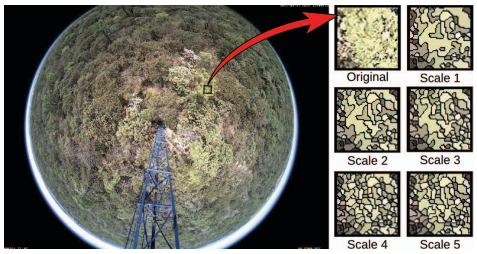

Fig. 1: Sample image of the cerrado savanna recorded by the digital camera on September 17th, 2011; and the segmentation results for the selected scales in a subimage sample.

The image analysis was conducted by defining different regions of interest (ROI), as described by Richardson *et al.* [2], Richardson *et al.* [4], and Ahrends *et al.* [3]. We defined eleven ROIs (Figure 2) based on the random selection of six plant species identified by phenology experts in the hemispheric image and on the ground: (1) three regions associated with *Aspidosperma tomentosum* (blue areas), (2) one region for *Caryocar brasiliensis* (green area), (3) one region for *Myrcia guianesis* (cyan area), (4) three regions for *Miconia rubiginosa* (red areas), (5) two regions for *Pouteria ramiflora* (yellow areas), and (6) one region for *Pouteria torta* (magenta areas).

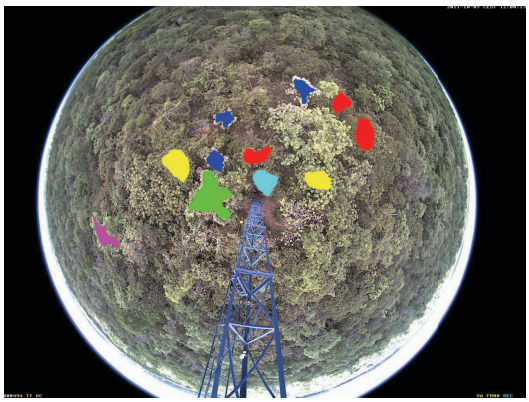

Fig. 2: Regions of interest (ROIs) defined for the analysis of cerrado-savanna digital images.

We analyze each region in terms of the contribution of the primary colors (Red, Green, and Blue), as proposed by Richardson *et al.* [2]. Initially, we analyze each color channel and compute the average value of the pixel intensity. After that, we compute the relative (or normalized) brightness of each color channel, as:

$$\begin{aligned} Total_{avg.} &= Red_{avg.} + Green_{avg.} + Blue_{avg.} \quad (1) \\ Red\ \% &= \frac{Red_{avg.}}{Total_{avg.}} \\ Green\ \% &= \frac{Green_{avg.}}{Total_{avg.}} \\ Blue\ \% &= \frac{Blue_{avg.}}{Total_{avg.}} \end{aligned}$$

where $Red_{avg.}$, $Green_{avg.}$, and $Blue_{avg.}$ are the average pixel intensity of the R, G, and B bands, respectively. The normalization of those values reduces the influence of the incident light, decreasing the color variability due to changes on illumination conditions [9].

Figure 3 hows the behavior of those values for each ROI along the whole period, considering only the digital images taken at different hours of a day. Each line corresponds to a time series for the variation of the normalized brightness of the G channel. Notice the differences between the behavior of each species, reflecting the leaf color changes over the leaf life cycle or aging process.

III. Dissimilarity Measures

Between all categories of dissimilarity measures existing in the literature, we have used three shape-based distances (Manhatan [10], Euclidean [11], and Dynamic Time Warping [12]), three edit-based distances (Edit Distance on Real Sequence [13], Edit Distance with Real Penalty [14], and Longest Common Subsequence [15]), and one feature-based distances (Zero-mean Normalized Cross Correlation [16]). In the following subsections, each of those distance functions is explained in more detail.

A. Manhatan Distance (L_1)

The Manhattan distance, also known as L_1, is one of the widely used distance function employed to compare time series [10]. Let T_a and T_b be two time series. The Manhatan distance between T_a and T_b is defined as follows:

$$Manhattan(T_a, T_b) = \sum_{i=1}^{K} |T_a[i] - T_b[i]| \quad (2)$$

B. Euclidean Distance (L_2)

The Euclidean distance, also known as L_2, between T_a and T_b is defined as follows [11]:

$$Euclidean(T_a, T_b) = (\sum_{i=1}^{K} (T_a[i] - T_b[i])^2)^{1/2} \quad (3)$$

C. Dynamic Time Warping (DTW)

The Dynamic Time Warping distance between two series T_a and T_b with m and n events, respectively, is defined as follows [12]:

$$DistDTW(T_a, T_b) = DTW(m, n) \quad (4)$$

where

$$DTW(i,j) = \begin{cases} 0, if \quad i = 0, j = 0 \\ \infty, if \quad i = 0 \\ \infty, if \quad \text{J} = 0 \\ dist(T_a[i], T_b[j]]) + \\ min\{DTW(i-1, j-1), \\ DTW(i-1, j), DTW(i, j-1)\}, otherwise \end{cases} \quad (5)$$

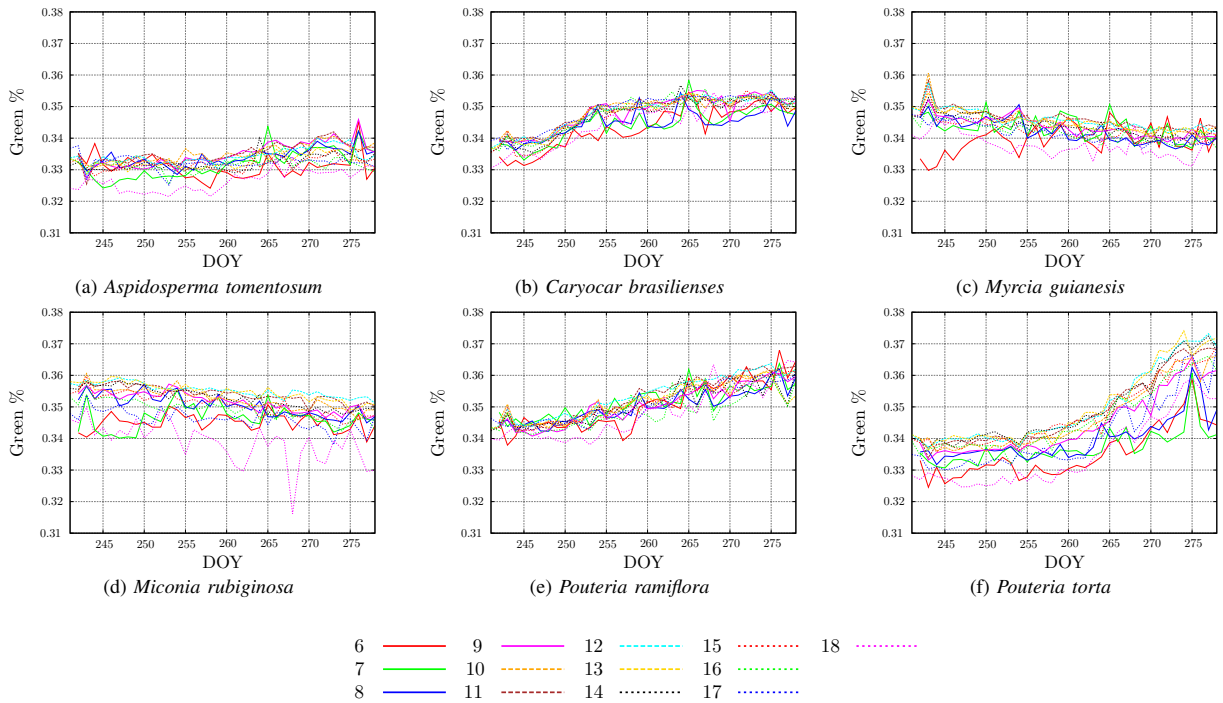

Fig. 3: The variance of the normalized brightness of each color channel from the digital images taken at different hours of a day, each Julian day (August 28th to October 3rd, 2011), in the cerrado savanna using different regions of interest (ROIs).

D. Edit Distance on Real Sequence (EDR)

$$EDR(T_a, T_b) = \begin{cases} n, if \quad m = 0 \\ m, if \quad n = 0 \\ min\{EDR(Rest(T_a), Rest(T_b)) + subcost, \\ EDR(Rest(T_a), T_b) + 1, \\ EDR(T_a, Rest(T_b)) + 1\}, otherwise \end{cases} \tag{6}$$

where $subcost = 0$, if $match(T_a[1], T_b[1]) = true$ and $subcost = 1$, otherwise. $Rest(T)$ stands for the time series obtained from T by eliminating the first element [13].

E. Edit Distance with Real Penalty (ERP)

$$ERP(T_a, T_b) = \begin{cases} \sum_{i=1}^{n} dist(T_b[i], g), if \quad m = 0 \\ \sum_{i=1}^{m} dist(T_a[i], g), if \quad n = 0 \\ min\{ERP(Rest(T_a), Rest(T_b)) + \\ dist(T_a[1], T_b[1]), \\ ERP(Rest(T_a), T_b) + dist(T_a[1], g) \\ ERP(T_a, Rest(T_b)) + dist(T_b[1], g)\}, \\ otherwise \end{cases} \tag{7}$$

where $dist(a, b)$ is the distance between two elements and g is a gap of edit distance. ERP also uses the real distance between elements ($dist(T_a[1], T_b[1])$) as the penalty to handle local time shifting [14].

F. Longest Common Subsequence (LCSS)

$$LCSS(i, j) = \begin{cases} 0, if \quad i = 0 \\ 0, if \quad j = 0 \\ 1 + LCSS(i-1, j-1), \\ if \quad |T_a[i] - T_b[i]| < \epsilon \\ \max\{LCSS(i-1, j), LCSS(i, j-1)\}, otherwise \end{cases}$$

$$dist_{LCSS}(T_a, T_b) = 1 - \frac{LCSS(i, j)}{min(n, m)} \tag{8}$$

G. Zero-mean Normalized Cross Correlation (ZNCC)

$$ZNCC(T_a, T_b) = \max_x \frac{\sum_{i=0}^{n-1} (T_a[i] - \overline{T_a})(T_b[x+i] - \overline{T_b})}{\sqrt{\sum_{i=0}^{n-1} (T_a[i] - \overline{T_a})^2 \times \sum_{i=0}^{m-1} (T_b[x+i] - \overline{T_b})^2}} \tag{9}$$

where $\overline{T_a}$ and $\overline{T_b}$ are the mean values of T_a and T_b, respectively [16].

IV. Experiments and Results

In this work, we adopted the evaluation method used in [17], [18]. It relies on the classification of time series extracted from pixels associated with individuals of a same species. For that, regions are defined by using a hierarchical segmentation based on the Guigues algorithm [19]. The image used to obtain the hierarchy of segmented regions was taken

on September 17th, 2011 (Figure 1). Five segmentation scales were selected from the hierarchy to extract the time series, as described in Section II. The finest scale is composed of $27,380$ regions and the coarsest scale contains $8,849$ regions. A given region is considered as belonging to a ROI if at least 80% of its size is overlapped by such a ROI. In our experiments, we have used only regions from the coarsest scale, as they have been shown the most effective ones [17], [18].

The distance functions described in Section III are used in the classification of time series using the k-Nearest Neighbor algorithm (kNN) classifier. The main motivations for employing kNN rely on the use of an explicit distance function in the classification process and its low complexity. A time series can be represented as a point in a high-dimensional space. A K-Nearest Neighbor query asks for the K records, in a dataset, more similar to a query q (a time series in our case), according to a distance metric. More precisely, the objective is to find the set $S = \{T_i \mid T_i \in DB\}$ such that $|S| = K$ and $\forall T_j \notin S, dist(q, T_i) \leq dist(q, T_j)$, where DB is the set containing all available time series and $dist$ is a distance function described in Section III. In our implementation, we assign q to the most frequent class found in S.

We use the weighted accuracy as evaluation measure to assess the quality of each kNN classifier. Let $|C|$ be the number of time series associated with class C and N_C the number of instances correctly classified as belonging to class C. The average accuracy AC of a classifier, considering instances of class C as input, is defined as follows:

$$AC = \frac{N_C}{|C|} \tag{10}$$

Conducted experiments aim to assess the accuracy of different distance functions in classification tasks involving different phenology study scenarios: (1) we consider time series obtained for different hours of a day (from 6 am to 6 pm) and, in this case, we are interested in determining which distance function yields the best classification results for each hour; (2) we consider time series obtained from three different color channels (R, G, and B) and, in this case, we are interested in determining which distance function yields the best results for each color channel; and (3) we consider different masks that are associated with individuals of different species and, in this case, we are interested in determining which distance function yields the best results for each species. In the following subsections, we discuss obtained results considering each one of those scenarios.

A. Evaluation according to the Hour of a Day

Figure 4 presents the average accuracy performance for different 1-NN classifiers implemented using the different distance functions, considering different time stamps. As it can be observed, for all classifiers, the best results are usually found in extreme hours (early in the morning and late in the afternoon). Note also that there is no clear winner classifier, except for the ZNCC-based kNN classifier whose accuracy performance is worse than all others.

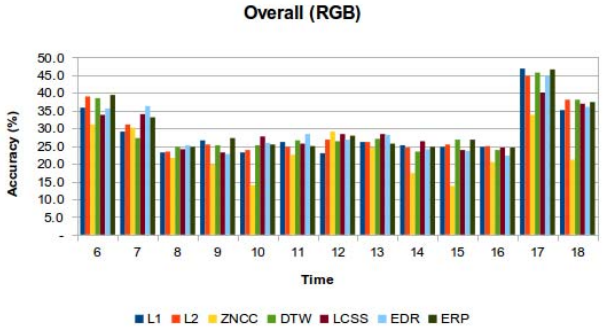

Fig. 4: Accuracy of all kNN classifiers with different distance functions using RGB Channels.

B. Evaluation according to Color Channels

Figures 5, 6, and 7 present the average accuracy performance for different 1-NN classifiers implemented using the different distance functions, considering the R, G, and B color channels, respectively. As it can be observed, for all color channels, there is no clear winner distance function. For all classifiers, again the best results are usually found in extreme hours (early in the morning and late in the afternoon). For the G color channel, the observed behavior of the accuracy performance of all classifiers over time is more steady. Note also that again that the ZNCC-based KNN classifiers usually yield the worse accuracy performances.

C. Evaluation according to Species

Figures 8, 9, and 10 present the average accuracy performance for different 1-NN classifiers implemented using the different distance functions, considering time series of the species *Aspidosperma tomentosum*, *Miconia rubiginosa*, and *Pouteria ramiflora*, respectively.

For the *Aspidosperma tomentosum*, it was observed high performance accuracy scores for all classifiers associated with the time series of 5 pm (more than 80% in terms of accuracy). The accuracy performance of the evaluated classifiers is worse for the species *Miconia rubiginosa*. The best results (around 40%) are observed at 6 pm. Note also that the performance of the ZNCC-based classifiers are not good again. For the species *Pouteria ramiflora*, observed accuracy performances are poor for all classifiers, except for the ZNCC-based KNN classifiers at 7 am.

D. Correlation Analysis of Classifiers

In our experiments, we have performed a correlation analysis among all classifiers implemented using the different distance functions. We used the well-known *Correlation Coefficient* ρ (COR) and this measure is defined as follows [20]:

$$COR(c_i, c_j) = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}. \tag{11}$$

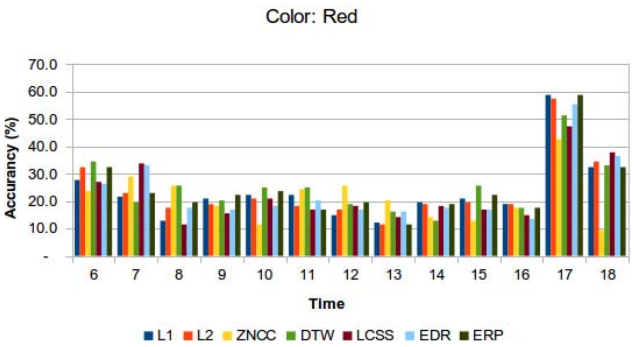

Fig. 5: Accuracy of all kNN classifiers with the different distance functions using **R color channel**.

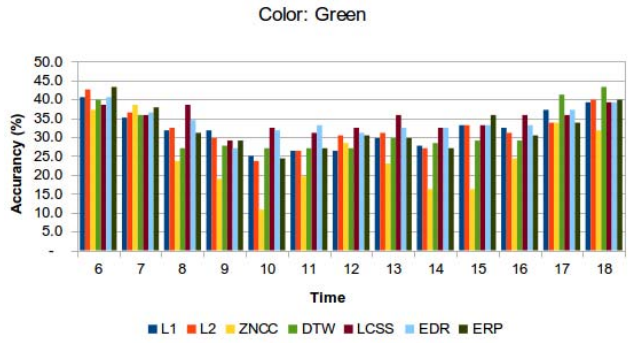

Fig. 6: Accuracy of all kNN classifiers with the different distance functions using **G color channel**.

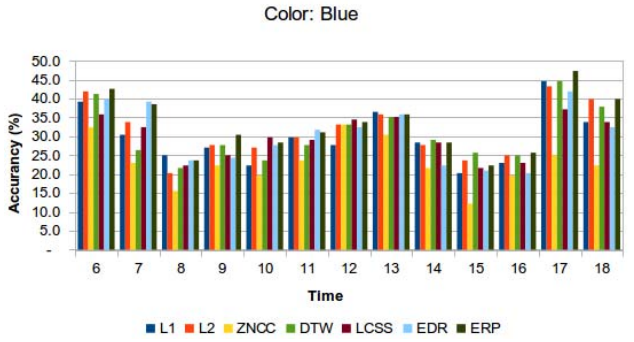

Fig. 7: Accuracy of all kNN classifiers with the different distance functions using **B color channel**.

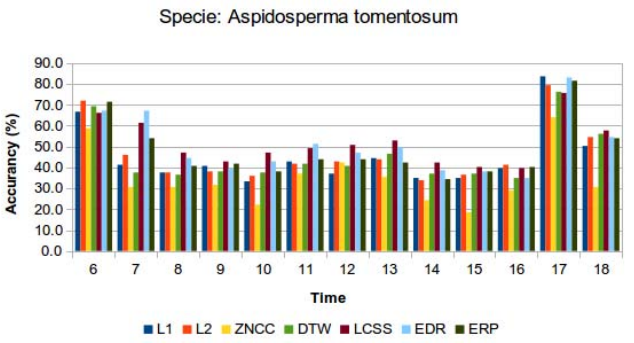

Fig. 8: Accuracy of all kNN classifiers with the different distance functions for the specie *Aspidosperma tomentosum*.

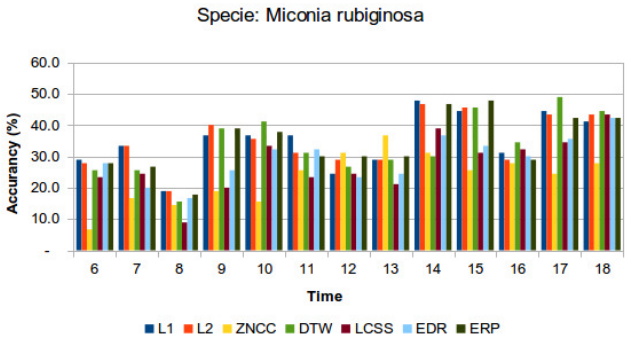

Fig. 9: Accuracy of all kNN classifiers with the different distance functions for the specie *Miconia rubiginosa*.

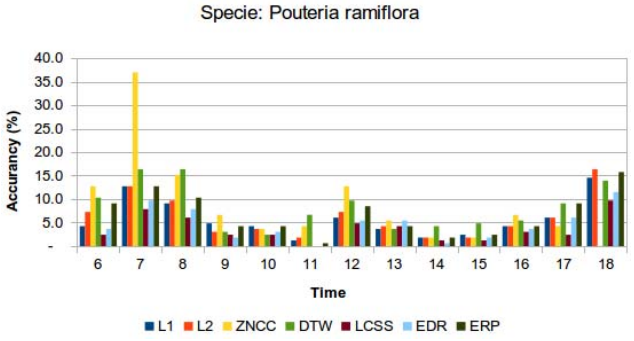

Fig. 10: Accuracy of all kNN classifiers with the different distance functions for the specie *Pouteria ramiflora*.

To calculate $COR(c_i, c_j)$, we have used a 2×2 matrix containing the relationship between a pair of classifiers with the percentage of agreement. This relationship matrix has the percentage of hit and miss for each classifier c_i and c_j. The value a is the percentage of instances that both classifiers c_i and c_j classified correctly in the testing set. Values b and c are the percentage of instances that c_j classified correctly but c_i missed and vice-versa. The value d is the percentage of instances that both classifiers missed.

Table I shows the correlation analysis scores between all seven classifiers considered in this work. As it can be observed, the classifiers that consider the distance functions L_1, L_2, and ERP are highly correlated. The $ZNCC$-based classifier shows to be the less correlated to other. There are no distance functions exactly equal between others, thus it means a good indicative to be used classifier fusion approaches [21], [22].

Classifier	kNN-1							
	Distances	L_1	L_2	$ZNCC$	DTW	$LCSS$	EDR	ERP
kNN-1	L_1	1.00	0.86	0.31	0.76	0.71	0.74	0.85
	L_2	-	1.00	0.33	0.73	0.71	0.74	0.94
	$ZNCC$	-	-	1.00	**0.28**	**0.28**	**0.28**	0.33
	DTW	-	-	-	1.00	0.66	0.67	0.72
	$LCSS$	-	-	-	-	1.00	0.83	0.71
	EDR	-	-	-	-	-	1.00	0.75
	ERP	-	-	-	-	-	-	1.00

TABLE I: Correlation analysis between all seven classifiers. The less correlated classifiers are highlighted in bold.

V. Conclusions

In this paper, we performed a comparative study of different time series distance function for identifying phenology patterns associated with leaf color variation over time. Similar study has never been done before in the literature. Performed experiments demonstrate that the kNN classifiers implemented with different distance functions yield better results for extreme hours (early in the morning and late in the afternoon). The best accuracy results have been achieved by distance functions L_1 and ERP using Red Channel at 5pm, 58.8%. It is possible to observe a behavior a bit more stable over time in the green and blue channels. Although the classifier based on $ZNCC$ distance function yields the worst results in almost all experiments, it achieved good results for the most difficult species (*Pouteria ramiflora*) of the dataset. Furthermore, a correlation analysis has been performed between all classifiers. This analysis showed that the distance functions L_1, L_2, and ERP are more correlated each other. However, there is no function that is totally equal other. Thus, in future work is possible consider the combination those different kNN classifiers using fusion approaches, and the investigation of novel distance functions.

Acknowledgments

This research was supported by the FAPESP-Microsoft Research Virtual Institute (grant 2010/52113-5). FAF is grateful to FAPESP (grant 2010/14910-0) and CAPES (grant 1260-12-0). BA a master scholarship from CAPES; LPCM and RST receive a Productivity Research Fellowship from CNPq (grants 306243/2010-5 and 306587/2009-2). Also, we have been benefited from funds of CNPq, CAPES, and FAPESP (grants 2007/52015-0 and 2009/18438-7)

References

[1] M. D. Schwartz, *Phenology: An Integrative Environmental Science*. Academic Publishers, 2003.

[2] A. D. Richardson, J. P. Jenkins, B. H. Braswell, D. Y. Hollinger, S. V. Ollinger, and M. L. Smith, "Use of digital webcam images to track spring greep-up in a deciduous broadleaf forest," *Oecologia*, vol. 152, pp. 323–334, 2007.

[3] H. Ahrends, S. Etzold, W. Kutsch, R. Stoeckli, R. Bruegger, F. Jeanneret, H. Wanner, N. Buchmann, and W. Eugster, "Tree phenology and carbon dioxide fluxes: Use of digital photography for process-based interpretation at the ecosystem scale," *Climate Research*, vol. 39, pp. 261–274, 2009.

[4] A. D. Richardson, B. H. Braswell, D. Y. Hollinger, J. P. Jenkins, and S. V. Ollinger, "Near-surface remote sensing of spatial and temporal variation in canopy phenology," *Ecological Applications*, vol. 19, pp. 1417–1428, 2009.

[5] B. Alberton, J. Almeida, R. Henneken, R. da S. Torres, A. Menzel, and L. P. C. Morellato, "Using phenological cameras to track the green up in a cerrado savanna and its on-the-ground validation," *Ecological Informatics*, vol. 19, pp. 62–70, 2014.

[6] R. S. Torres, M. Hasegawa, S. Tabbone, J. Almeida, J. A. Santos, B. Alberton, and L. P. C. Morellato, "Shape-based time series analysis for remote phenology studies," in *IEEE International Geoscience and Remote Sensing Symposium*, 2013, pp. 3598–3601.

[7] J. Almeida, J. A. Santos, B. Alberton, L. P. C. Morellato, and R. S. Torres, "Visual rhythm-based time series analysis for phenology studies," in *IEEE International Conference on Image Processing*, 2013, pp. 4412–4416.

[8] J. Almeida, J. A. Santos, B. Alberton, L. P. C. Morellato, and R. S. Torres, "Plant species identification with phenological visual rhythms," in *IEEE International Conference on eScience*, 2013, pp. 148–154.

[9] H.-D. Cheng, X. Jiang, Y. Sun, and J. Wang, "Color image segmentation: Advances and prospects," *Pattern Recognition*, vol. 34, no. 12, pp. 2259–2281, 2001.

[10] B.-K. Yi and C. Faloutsos, "Fast time sequence indexing for arbitrary lp norms," in *International Conference on Very Large Data Bases*, 2000, pp. 385–394.

[11] C. Faloutsos, M. Ranganathan, and Y. Manolopoulos, "Fast subsequence matching in time-series databases," *SIGMOD Record*, vol. 23, no. 2, pp. 419–429, May 1994.

[12] D. J. Berndt and J. Clifford, "Using Dynamic Time Warping to Find Patterns in Time Series," in *KDD Workshop*, 1994, pp. 359–370.

[13] L. Chen and M. T. zsu, "Robust and fast similarity search for moving object trajectories," in *ACM SIGMOD International Conference on Management of Data*, 2005, pp. 491–502.

[14] L. Chen and R. Ng, "On the marriage of lp-norms and edit distance," pp. 792–803, 2004.

[15] M. Vlachos, "Discovering similar multidimensional trajectories," in *International Conference on Data Engineering*, 2002, pp. 673–684.

[16] J. Martin and J. L. Crowley, "Experimental comparison of correlation techniques," in *International Conference on Intelligent Autonomous Systems*, 1995.

[17] J. Almeida, J. A. dos Santos, B. Alberton, R. S. Torres, and L. P. C. Morellato, "Remote phenology: Applying machine learning to detect phenological patterns in a cerrado savanna," in *IEEE International Conference on eScience*, 2012, pp. 1–8.

[18] J. Almeida, J. A. Santos, B. Alberton, R. S. Torres, and L. P. C. Morellato, "Applying machine learning based on multiscale classifiers to detect remote phenology patterns in cerrado savanna trees," *Ecological Informatics*, 2013. DOI: 10.1016/j.ecoinf.2013.06.011

[19] L. Guigues, J. Cocquerez, and H. Le Men, "Scale-sets image analysis," *International Journal of Computer Vision*, vol. 68, pp. 289–317, 2006.

[20] L. I. Kuncheva and C. J. Whitaker, "Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy," *Machine Learning*, 2003.

[21] F. A. Faria, J. A. dos Santos, A. Rocha, and R. da S. Torres, "A framework for selection and fusion of pattern classifiers in multimedia recognition," *Pattern Recognition Letters*, vol. 39, no. 0, pp. 52 – 64, 2014.

[22] F. S. P. Andrade, J. Almeida, H. Pedrini, and R. S. Torres, "Fusion of local and global descriptors for content-based image and video retrieval," in *Iberoamerican Congress on Pattern Recognition*, 2012, pp. 845–853.