MULTI-SCALE TIME SERIES ANALYSIS OF EVAPOTRANSPIRATION FOR HIGH-THROUGHPUT PHENOTYPING FREQUENCY OPTIMIZATION

Soumyashree Kar[1*], Ryokei Tanaka[2], Hiroyoshi Iwata[2], Jana Kholova[3], Surya S. Durbha[1], J. Adinarayana[1], Vincent Vadez[3,4*]

[1]Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay, Mumbai, India 400076;
[2]Laboratory of Biometrics and Bioinformatics, University of Tokyo, Japan
[3]International Crop Research Institute for Semi-Arid Tropics, Hyderabad, India, 502319;
[4]Institut de Recherche pour le Developpement (IRD) – Université de Montpellier – UMR DIADE, 911 Avenue Agropolis, BP 64501, 34394 Montpellier cedex 5, France

KEY WORDS: High-Throughput Plant Phenotyping (HTPP), Evapotranspiration, Sampling Frequency, ARIMA modeling, Time Series Classification, Entropy, Conditional Entropy, Optimization

ABSTRACT:

This work is undertaken considering the significance of functional phenotyping (primarily measured from continuous profiles of plant-water relations) for crop selection purposes. High-Throughput Plant Phenotyping (HTPP) platforms which largely employ state-of-the-art sensor technologies for acquisition of vast amount of field data, often fail to efficiently translate sensor information into knowledge due to the major challenges of data handling and processing. Hence, it is imperative to concurrently find a way for dissociating noise from useful data. Additionally, another important aspect is understanding how frequent should be the data collection, so that information is maximized. This paper presents a novel approach for identifying the optimal frequency for phenotyping evapotranspiration (ET) by assimilating results from both time series forecast as well as classification models. Thus, at the optimal frequency, plant-water relations can not only be desirably predicted but genotypes can also be classified based on the characteristics of their ET profiles. Consequently, this will aid better crop selection, besides minimizing noise, redundancy, cost and effort in HTPP data collection. High frequency (15 min) ET time series data of 48 chickpea varieties (with considerable genotypic diversity) collected at the LeasyScan HTPP platform, ICRISAT is used for this study. Time series forecast and classification is performed by varying frequency up to 180 min. Multiple performance measures of time series forecast and classification are combined, followed by implementation of entropy theory for sampling frequency optimization. The results demonstrate that ET time series with a frequency of 60 min per day potentially yield the optimum information.

1. INTRODUCTION

A run-through of the High-Throughput Plant Phenotyping (HTPP) methods implemented particularly in the last decade reveals the extensive use of Remote Sensing (RS) in characterizing plant processes. Availability of a plethora of imaging systems has facilitated anatomization of observable plant features (largely morphological), while weather sensors have effectuated detailed crop mapping and modeling under varying environmental conditions. The imaging systems can vary from very close-range fixed digital cameras to space-borne sensors, thereby covering leaf-level to farm-level details. Similarly, weather data can be collected from plant-level (i.e. microclimatic) to regional scales. The opportunity to generate insightful inferences from these data further broadens with increased frequency of data collection. However, despite the technological developments thereof, an aspect of HTPP which still remains relatively less explored is that of plant physiology i.e. functional phenotyping (Halperin et al, 2017; Gosa et al, 2019) which relates to canopy-conductance traits. Furbank, Tester, 2011 have even quoted phenomics as 'high-throughput plant physiology'.

There are essentially two primary limitations which impede the efficient utilization of HTPP for evaluating the differences in plant water use capacities: 1. the difficulty to measure plant water use frequently, and 2. identifying the optimal frequency for simultaneous measurement of canopy-conductance traits. Vadez et al, 2015 have addressed the first limitation by enabling simultaneous measurement of evapotranspiration (ET) under

non-controlled environmental conditions, while the second is addressed in this paper. Time series observations of ET are used to explore genotypic differences in the canopy-conductance traits, among 48 chickpea genotypes. The implications of ET in understanding canopy-conductance is elaborated in Isard, Belding, 1989 which asserts that "plant-water relations are majorly studied from temporal characterization of ET, since ET is the only one of the three principal transport phenomena between the earth's surface and lower atmosphere (water vapour, heat, and momentum flows) that can be measured directly". Halperin et al, 2017 have also reinforced the need for 'simultaneous and continuous monitoring of water relations in the soil–plant–atmosphere continuum' (i.e. the ET profile of each plant over time).

It has been studied that pot experiments equipped with gravimetric non-contact sensors (i.e. load cells) can be potentially used for estimating field ET (Lu et al, 2018). Former functional HTPP studies have utilized high frequency load cells data for ranking plants with respect to their drought tolerance and other physiological characteristics (Granier et al, 2006, Halperin et al, 2017). Furthermore, although these sensors enable evaluation of the impact of soil hydraulic characteristics on irrigation management strategies (Tolk et al, 2005) and HTPP of canopy-conductance traits (Lazarovitch et al, 2006; Denich et al, 2010), there has rarely been any mention about optimizing the data collection frequency at such platforms. Identification of an optimum sampling frequency would not only prevent noise and redundancy in raw data, but can also promulgate systematic crop-specific irrigation scheduling. Hence, this work is focused on

* ksoumya2301@gmail.com; karsoumya@iitb.ac.in; V.VADEZ@cgiar.org

optimizing the sampling frequency of load cells observations for efficient categorization of chickpea genotypes in terms of ET.

High frequency ET time series exploited for this work was obtained from the LeasyScan HTPP platform at ICRISAT, Hyderabad, India which uses load cells to measure pot weight at a frequency of 15 minutes. Due to several extraneous influences, data generated in such platforms in non-controlled conditions tend to be noisy. Hence, these data sets require adequate de-noising (or pre-processing) and relevant feature extraction, in order to facilitate precise distinction in the characteristics of the desired trait.

In this work, a thorough set of steps was adopted for the conversion of raw load cells time series into interpretable ET profiles, followed by multi-scale Autoregressive Integrated Moving Average (ARIMA) modeling. The ARIMA based models were built for forecasting each ET time series at frequencies (interchangeably used with scales), ranging from 15 min to 180 min. The genotypes were clustered at the base frequency i.e. 15 min and classification accuracy was estimated for each successive frequency. Subsequently, a combined set of metrics obtained from forecasting as well as classification, was used for entropy-based optimization of sampling frequency. The details of each step is presented in Section 2 and the corresponding results and description are elaborated in section 3.

2. MATERIALS AND METHODOLOGY

2.1 Materials

For this study, 4 replications of 48 chickpea genotypes were phenotyped at the LeasyScan HTPP platform, ICRISAT-Patancheru (17.5111° N, 78.2752° E). The platform is equipped with Phenospex® 3D laser scanners which traverse on top of the plants, and provide 3-D images (used for canopy-growth trait analysis, not covered in this paper) every 2 hours. The total capacity of LeasyScan is ~ 5000 plants, grown in sectors (unit of analysis) across the trenches of platform. For simultaneous phenotyping of canopy-conductance traits, each sector is associated with gravimetric sensors (load cells) that measure weight of respective sectors every 15 minutes. The whole system operates in non-controlled or open-environment conditions and provides HTPP data close to real field conditions (see Vadez et al, 2015 for details). The time series data set examined in this work comprised load-cell weights collected at every 15 minutes' interval, from 20th February to 6th March 2017 i.e. 15 days. The following analysis was conducted by considering the mean (of 4 replications) time series of each variety.

2.2 Methodology

2.2.1 Pre-processing:

The complete pre-processing was done in two stages. The first stage of pre-processing included de-noising load-cell values to obtain ET time-series. In this step, pot weight values lying beyond the range of 60,000-100,000 grams were first discarded, as per the limits of the experimental set-up (Vadez et al, 2015). It was observed that there were both phenomenon-based as well as latent anomalies, which were removed independently. The phenomenon-based anomalies were primarily caused due to intermittent irrigation events which resulted in sudden rise and drop in pot weight, after the discharge of excess water. This resulted in sudden spikes in each of the time series. Contrarily, abrupt changes in ambient environmental conditions, besides system-generated noise comprised the latent anomalies. While the former anomalies needed complete removal from data, the latter required meticulous selection of the observations to be

retained (i.e. due to ambient effects) and the ones that should be removed (i.e. system-generated and external noise). Discrete wavelet transform (DWT) was implemented on each time series up to the first three levels, and outliers in the coefficients of each level were identified using boxplot procedure. The locations of those outliers in the DWT coefficients were retrieved, and the corresponding observations were removed from the dataset. Thus, anomalous peaks from the raw data were eliminated. The consequent missing values were linearly interpolated, and their difference at lag 1 was calculated to get the ET time series at 15 minutes' frequency, for each plant.

The next stage in pre-processing included computation of reference ET i.e. Penman Monteith ET in mm/mm2/15 minutes [6] using the weather variables, collected from the Campbell sensors, installed at the LeasyScan platform. Reference ET values were used as a benchmark to analyse possible abnormalities in the observed ET values of all the genotypes. The ratios of observed |ET| to reference ET were calculated for day-time (solar radiation greater than 0) and night-time (solar radiation equal 0). Every observation which resulted in a ratio beyond 1.0 and 1.5, for day- and night-time respectively, was considered abnormal and consequently filtered out before smoothing each ET time series using the cubic spline method.

2.2.2 Multi-scale ARIMA Modeling:

Auto Regressive Integrated Moving Average (ARIMA) models are used to explain a given time series in terms of its own past values, i.e. its own lags, and the lagged forecast errors are used to forecast future values (Ho, Xie, 1998). Here, ARIMA modeling was performed to assess forecast ability at different scales using different forecast metrics. However, ARIMA models require the series to be stationary. Hence, the Augmented Dickey-Fuller test was executed prior to ARIMA modeling for inspecting the stationarity of each time series (Cheung, Lai, 1995). A linear trend in the time series was identified due to the growth of plants during the later days and increased water loss (through higher transpiration). The presence of trend resulted in a p-value greater than 0.05 for each time series. Hence, the first differenced time series was used for modeling. Through this test, it was also identified that majority of the time series were auto-correlated only till lag 12 i.e. 180 minutes. Therefore, multi-scaling was done by reducing the frequency of each time series by 15 minutes i.e. the interval of data points was increased from 15 minutes (originally) to 30, 45, 60, …, 180 minutes, through aggregation as suggested in Wu et al, 2013.

The entire data set was converted into 12 consecutive scales, with an interval of 15 min between each. Subsequently, in order to assess the forecast ability at each scale the ARIMA model was used, and a unique model was built for ET time series of each plant. The auto.arima function in R was implemented for selecting the best fit model for each time series, and prediction was performed on the last 5 days (the initial 10 days were used for training). The corresponding Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Root Mean Square Error (RMSE) for each model at each sampling frequency were then retrieved (Tsay, 1984). While AIC denotes the out-of-sample prediction error, BIC measures the trade-off between model fit and complexity of the model, and RMSE represents the absolute fit of the model (i.e. how close is predicted value to the observed value). The mean AIC, BIC and RMSE across all the models (one per plant) were retained as the representative prediction measures per frequency.

2.2.3 Multi-scale Classification:

In this study, there was insufficient prior information about the class a particular genotype belongs to, as well as the number of classes inherently present in the data. Therefore, an optimal number of clusters was first identified from the 15 min interval ET time series using Silhouette method (Rousseeuw, 1987). It was found that the data set could be optimally segregated into three groups. This was followed by feature-based time series clustering, since time series features are known to be more informative and offer better representation of the data than the raw time series (Laurinec, 2018). Features also help in significantly reducing the time series dimensionality, besides emphasising the fundamental characteristics that are most instrumental in delineating the differences among samples (here, the genotypes).

Time series features can be extracted using different methods e.g. nondata adaptive, data adaptive, model based and data dictated (Laurinec, 2018). While the nondata adaptive methods transform every time series using the same parameters, irrespective of the characteristic differences between them, the parameters of data adaptive transformations vary according to each data series. In data dictated methods, a compression ratio (Aghabozorgi et al, 2015) is defined per time series. However, model based methods assume that each observed time series is based on a basic model, and every two time series which could be explained by the same model parameters are similar. For this study, model based feature extraction was implemented; the parameters obtained from ARMA models were used as features for clustering the time series (Xiong, Yeung, 2002; Otranto, 2008). Unsupervised Random Forest (Shi, Horvath, 2005) method was employed for obtaining the cluster indices, subsequently used as class labels for estimating the supervised Random Forest based classification performance measures; accuracy (in percentage), F1-score (harmonic mean of precision and recall), sensitivity (the true-positive rate) and normalized ZeroOneLoss (each correct classification has loss=0, and incorrect classification has loss=1), at each scale.

Thus, the cluster-then-label approach to multi-class classification (de Sa, V.R., 1994; Nguyen, Smeulders, 2004; Nilashi et al, 2017) was included as a part of the procedure for optimizing the sampling frequency. The four classification metrics, calculated on test data (30% of the total samples, selected randomly) were coalesced with the AIC, BIC and RMSE estimates obtained from the previous step. Thus, the data set used for entropy-based optimization (see next section) was a table with seven attributes per scale, representative of the respective forecast ability as well as classification performance.

2.2.4 Entropy-based Frequency Optimization:

Entropy (E) per frequency (Equation 1) was calculated from the vector of performance attributes (as detailed in the previous section), followed by computation of conditional entropy (CE) between the base frequency and every other successive frequency (Equation 2). E and CE were calculated as proposed in Kullback, 1997 and Karamouz et al, 2009. While E enabled estimation of the average amount of information (pertaining to forecasting and classification performance) at each frequency, CE was incorporated as the measure of redundancy in the information (between each frequency and the base frequency). Thus, each frequency level was associated with its respective information measures, leveraged to define the cost (or objective) function for optimization. The objective of the study was to minimize redundancy between two frequency levels, and ensure low entropy (i.e. less randomness) at individual scale. Hence, the objective function was defined as shown in Equation 3.

$$H(y) = -\sum_{i=1}^{N} p(y_i) \log p(y_i) \qquad (1)$$

$$T(x,y) = H(x) + H(y) - H(x,y) \qquad (2)$$

$$Minimize\ Z = \frac{CE_s - E_s}{CE_s} \qquad (3)$$

where N represents the number of attributes y_i with probabilities $p(y_i)$ and $i = 1, ..., N$. $H(x,y)$ is the joint Entropy between the base scale, x and every other scale, y. CE_s and E_s are the CE and E at a given scale, s respectively ($s = 1, ..., n_s$; n_s = total number of sampling frequencies or scales).

The constraints for optimization are given as:

$CE - E \leq 0$
$E_{min} \leq E \leq E_{max}$
$CE_{min} \leq CE \leq CE_{max}$

It may be noted that, here, the objective was not to identify the optimum y and x values, rather the frequency at which the cost was minimum. This procedure was therefore, similar to the implementation of internal cluster validation optimization methods (Bezdek, Pal, 1995; Gajawada et al, 2012) used in machine learning. The reportedly robust Nelder-Mead optimization procedure (Singer, Nelder, 2009) was implemented using the optim function in R (R, core team, 2013).

3. RESULTS AND DISCUSSION

3.1 Pre-processing

The raw time series data from each load cell contained both missing as well as erroneous observations (Fig. 1a). The extremely anomalous values were removed and the remaining observations (Fig. 1b) were used to generate the ET time series. Fig. 1b also shows the efficacy of using DWT in identifying the spikes in the data set. In this step, the objective was to devise a basic set of methods to systematically transform raw data into temporal profiles (i.e. pre-processing) of plant responses. Hence, after complete pre-processing of the raw data, ET time series of few genotypes were compared for visual inspection (Fig. 2).

The chickpea genotypes used for this analysis either belonged to wild, cultivated or check varieties. There were also some genotypes which were expected to have high transpiration efficiency (TE), while another set of genotypes had low TE. Fig. 2 compares the ET profiles of three different varieties from the high (G22, blue), low TE (G32, red) and the wild types (G1, purple), with reference ET (green) as a theoretical baseline. Reference ET represented the maximum possible water loss under the given environmental conditions, and the genotypic ET profiles were quite lower since the plants were at the very initial growing stages with smaller canopy sizes (hence, lesser stomatal water loss). The overlaid plot of reference ET also shows the consistency of pre-processing results, under the same ambient conditions. The practically more important result was however, the visually discernible differences among the ET observations. In the following steps, the differences in ET are segregated through clustering, and classification performance is assessed at different sampling frequencies.
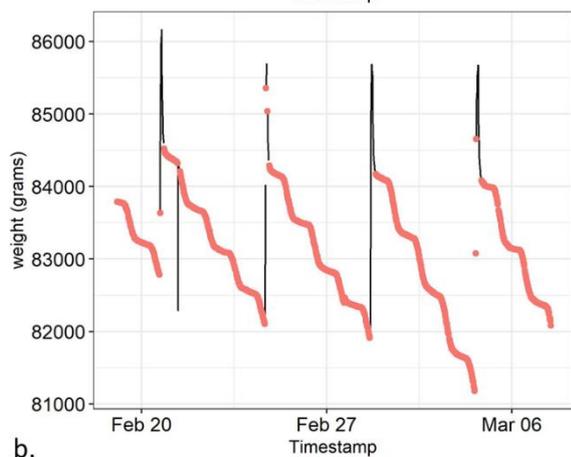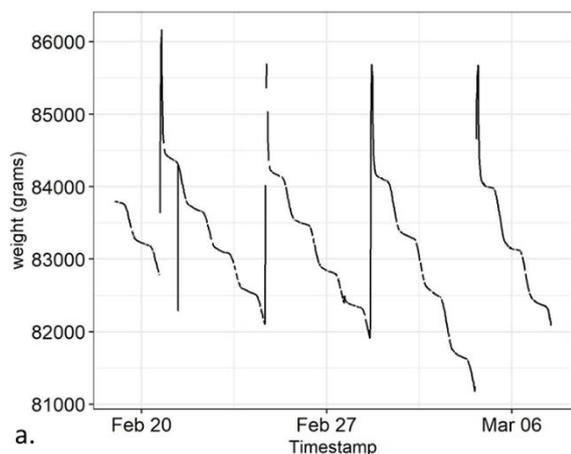
Figure 1. An example of raw load cell time series (a) and stage-1 pre-processed time series (b; data points retained after outlier removal is shown in red)
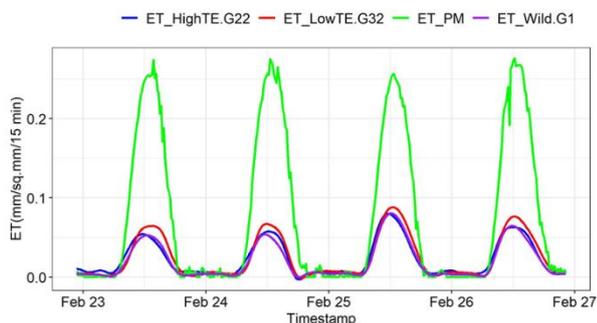
Figure 2. Comparison of smooth ET time series (stage-2 pre-processing) of three different genotypes and reference ET (ET_PM)

3.2 Multi-scale ARIMA Modeling

The ET trend and seasonality plots of one of the genotypes is shown in Fig. 3. It was observed that there was an increasing trend due to increased water loss with the eventual growth of plants, while seasonality denoted the daily cyclic patterns. Fig. 4 represents the forecast for the last five days of the series at 15 min interval and a confidence interval of 95%. The predicted values were found to lie quite close to the original values, thereby ensuring acceptable performance of the selected models. Furthermore, normalized values of the forecast metrics (Table 1), AIC, BIC and RMSE revealed the best predictability at 15 min scale. However, prediction RMSE was found to be lowest at 60 min scale. While a remarkable increase in RMSE was noticed only after the 105 min scale, AIC and BIC were found to have

considerably large estimates only after 60 min scale. Thus, the results (as per time series forecast performance) implied that, the optimal sampling frequency could lie between 15 min to 105 min. The relevance of considering these three metrics for time series model selection have been affirmed in several other studies as well (Atkinson et al, 2012; Boroojeni et al, 2017). This study however, includes these performance measures to compare model prediction at different scales.
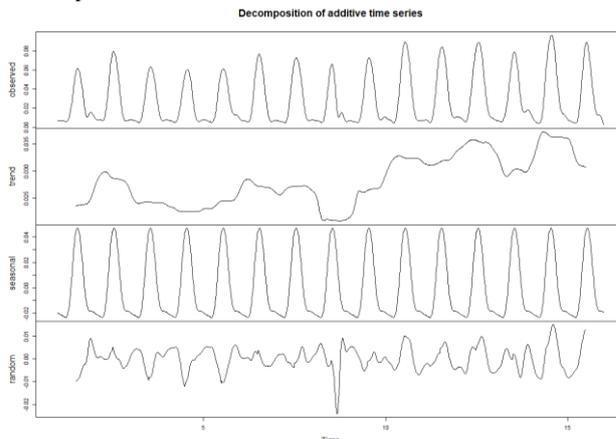
Figure 3. ET time series decomposition plot of a single genotype (as an example) representing the observed, trend, seasonal and random series (in the order of top to bottom).
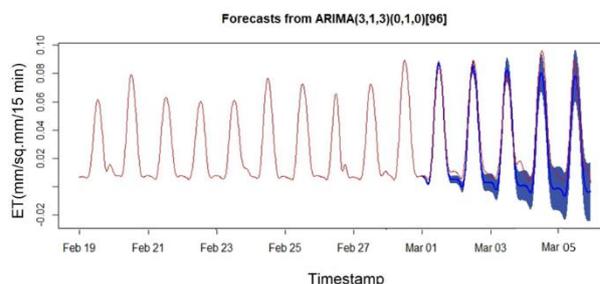
Figure 4. ET time series forecast plot of a single genotype as per the best fit ARIMA model
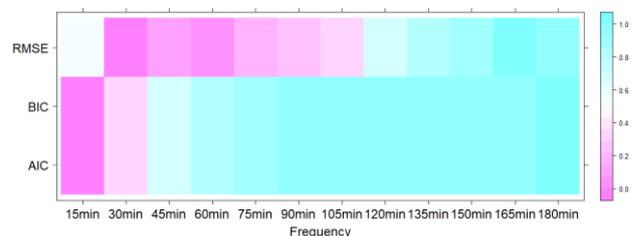
Figure 5. Levelplot of the time series forecast metrics (AIC, BIC, RMSE) across the entire range of frequencies

3.3 Multi-scale Classification

One important objective of this work was also to enable delineation of differences among the varieties with respect to the differences in their ET profiles. However, discrete segregation of the ET time series feature space required prior label information. Hence, cluster indices at 15 min scale were used for classification. The separation between clusters is shown on the first and second dimensions of the Multidimensional Scaling (MDS) plot (Fig. 6a). Additionally, the differences between the average ET profiles of each cluster is plotted in Figure 6b, which clearly identifies that water loss due to ET is the least in Cluster 1 and highest in Cluster 3.

In Table 2 the normalized values of classification performance measures are given which indicate that maximum classification accuracy was achieved at 15, 60 and 105 min scales while the F1-score is highest at 60 and 165 min. The true-positive rate was found to be consistently high only until 60 min scale beyond which it started fluctuating. Finally, ZeroOneLoss was also found to be the least at 60 min and 105 min scales. Thus, the intuitive inference about the optimal sampling frequency obtained from an overall examination of the classification measures coincided with that of the time series forecast measures. These results also state the need to examine several classification performance metrics, since conclusions differ with different metrics i.e. performance cannot be completely ascertained with a single metric, particularly in case of imbalanced datasets (Ferri et al, 2009). The classification measures leveraged in this work are local measures i.e. these are estimated considering the frequency of each class known to have low sensitivity to class noise (Ferri et al, 2009).

a.

b.
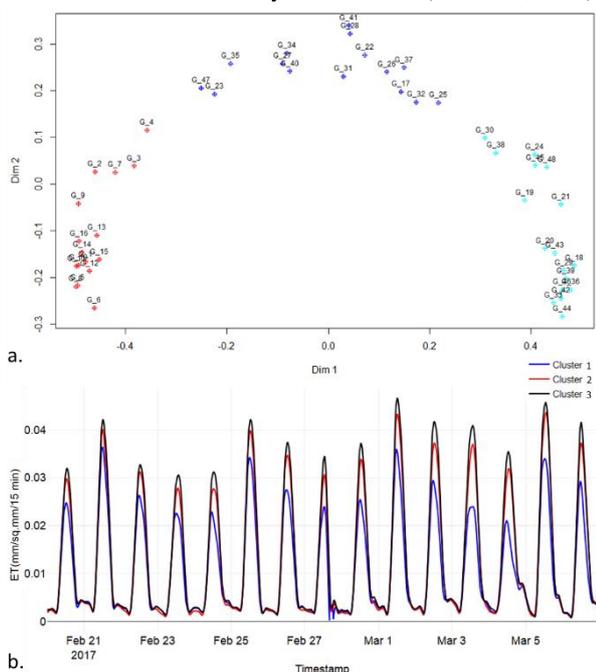
Figure 6. MDS plot genotypic clusters (a) and average ET time series of each cluster (b)
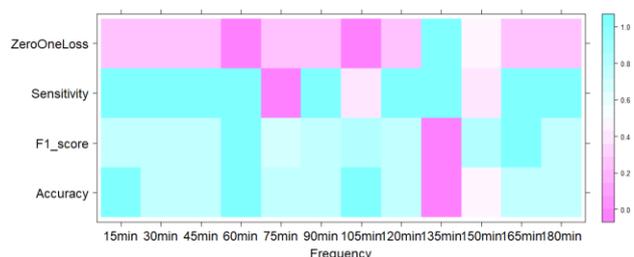
Figure 7. Levelplot of the time series classification metrics (Accuracy, F1_score, Sensitivity, ZeroOneLoss) across the entire range of frequencies

3.4 Entropy-based Frequency Optimization

The proposed optimization procedure aims to identify a suitable frequency for HTPP data collection at which both time series predictability as well genotype classification performance is high. Hence, the combined set of metrics from both the methods were used to compute Entropy (E) and Conditional Entropy (CE) per frequency level (or scale). Subsequently, Function Value (FV) at each scale was calculated as per Equation 3. The variation in E, CE and FV is represented in Fig. 8, for the entire range of frequency examined in this work. The optimal frequency (at

which FV is closest to the optimal value, obtained from the Nelder-Mead algorithm) was found to be 60 min, at which both E and CE were quite low. Furthermore, an overall inspection of E and CE implied that beyond the 150 min scale, randomness was higher and redundancy was lower. Thus, it was inferred that observed data would tend to be more random (insufficient) and the cyclic ET patterns will be less predictable if data collection frequency is less than 150 min.
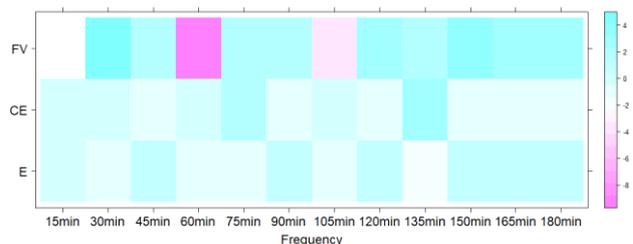
Figure 8. Levelplot of Entropy (E), Conditional Entropy (CE) and Function Value (FV) across the entire range of frequencies

4. CONCLUSIONS

This study demonstrated an extensive method for optimizing frequency of data collection at HTPP platforms, particularly operated in non-controlled conditions. Dynamic changes in the ambient environment often result in noisy and redundant data, thereby impeding efficiency of information extraction. Hence, an intuitive set of steps was thoroughly devised for denoising and converting sensor data to ET time series. ARIMA modeling was performed by changing the daily frequency of the entire time series data set from 15 min to 180 min. It was found that there was no single frequency at which all the forecast measures (AIC, BIC and RMSE) yielded the best value. Rather, it was noticed (based on only ARIMA modeling) that the optimal frequency lied in the range of 15 min to 105 min. Similar inferences were deduced based on classification performance measures as well. The metrics from both the methods were used for estimating E and CE at each frequency, and optimal sampling frequency was identified such that redundancy and randomness was the minimum at that frequency. It was finally concluded that a sampling interval of 1 hour or 60 min might be sufficient for desired information extraction from ET time series with reduced noise in the observed data.

REFERENCES

Aghabozorgi, S., Shirkhorshidi, A.S. and Wah, T.Y., 2015. Time-series clustering–A decade review. Information Systems, 53, pp.16-38.

Atkinson, P.M., Jeganathan, C., Dash, J. and Atzberger, C., 2012. Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. Remote sensing of environment, 123, pp.400-417.

Bezdek, J.C. and Pal, N.R., 1995, November. Cluster validation with generalized Dunn's indices. In Proceedings 1995 Second New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems (pp. 190-193). IEEE.

Boroojeni, K.G., Amini, M.H., Bahrami, S., Iyengar, S.S., Sarwat, A.I. and Karabasoglu, O., 2017. A novel multi-time-scale

modeling for electric power demand forecasting: From short-term to medium-term horizon. Electric Power Systems Research, 142, pp.58-73.

Cheung, Y.W. and Lai, K.S., 1995. Lag order and critical values of the augmented Dickey–Fuller test. Journal of Business & Economic Statistics, 13(3), pp.277-280.

de Sa, V.R., 1994. Learning classification with unlabeled data. In Advances in neural information processing systems (pp. 112-119).

Denich, C. and Bradford, A., 2010. Estimation of evapotranspiration from bioretention areas using weighing lysimeters. Journal of Hydrologic Engineering, 15(6), pp.522-530.

Ferri, C., Hernández-Orallo, J. and Modroiu, R., 2009. An experimental comparison of performance measures for classification. Pattern Recognition Letters, 30(1), pp.27-38.

Furbank, R.T. and Tester, M., 2011. Phenomics–technologies to relieve the phenotyping bottleneck. Trends in plant science, 16(12), pp.635-644.

Gajawada, S., Toshniwal, D., Patil, N. and Garg, K., 2012. Optimal clustering method based on genetic algorithm. In Proceedings of the International Conference on Soft Computing for Problem Solving (SocProS 2011) December 20-22, 2011 (pp. 295-303). Springer, New Delhi.

Gosa, S.C., Lupo, Y. and Moshelion, M., 2019. Quantitative and comparative analysis of whole-plant performance for functional physiological traits phenotyping: new tools to support pre-breeding and plant stress physiology studies. Plant science, 282, pp.49-59.

Granier, C., Aguirrezabal, L., Chenu, K., Cookson, S.J., Dauzat, M., Hamard, P., Thioux, J.J., Rolland, G., Bouchier-Combaud, S., Lebaudy, A. and Muller, B., 2006. PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in Arabidopsis thaliana permitted the identification of an accession with low sensitivity to soil water deficit. New phytologist, 169(3), pp.623-635.

Halperin, O., Gebremedhin, A., Wallach, R. and Moshelion, M., 2017. High-throughput physiological phenotyping and screening system for the characterization of plant–environment interactions. The Plant Journal, 89(4), pp.839-850.

Ho, S.L. and Xie, M., 1998. The use of ARIMA models for reliability forecasting and analysis. Computers & industrial engineering, 35(1-2), pp.213-216.

Isard, S.A. and Belding, M.J., 1989. Evapotranspiration from the alpine tundra of Colorado, USA. Arctic and Alpine Research, 21(1), pp.71-82.

Karamouz, M., Kerachian, R., Akhbari, M. and Hafez, B., 2009. Design of river water quality monitoring networks: a case study. Environmental Modeling & Assessment, 14(6), p.705.

Kullback, S., 1997. Information theory and statistics. Courier Corporation.

Laurinec, P., 2018. TSrepr R package: Time Series Representations. J. Open Source Software, 3(23), p.577.

Lazarovitch, N., Ben-Gal, A. and Shani, U., 2006. An automated rotating lysimeter system for greenhouse evapotranspiration studies. Vadose Zone Journal, 5(2), pp.801-804.

Lu, Y., Ma, D., Chen, X. and Zhang, J., 2018. A simple method for estimating field crop evapotranspiration from pot experiments. Water, 10(12), p.1823.

Nguyen, H.T. and Smeulders, A., 2004, July. Active learning using pre-clustering. In Proceedings of the twenty-first international conference on Machine learning (p. 79). ACM.

Nilashi, M., Ibrahim, O., Ahmadi, H. and Shahmoradi, L., 2017. A knowledge-based system for breast cancer classification using fuzzy logic method. Telematics and Informatics, 34(4), pp.133-144.

Otranto, E., 2008. Clustering heteroskedastic time series by model-based procedures. Computational Statistics & Data Analysis, 52(10), pp.4685-4698.

R Core Team, 2013. R: A language and environement for statistical computing. Vienna, Austria. URL http://www.R-project.org/.

Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics, 20, pp.53-65.

Shi, T. and Horvath, S., 2006. Unsupervised learning with random forest predictors. Journal of Computational and Graphical Statistics, 15(1), pp.118-138.

Singer, S. and Nelder, J., 2009. Nelder-mead algorithm. Scholarpedia, 4(7), p.2928.

Tolk, J.A., Howell, T.A. and Evett, S.R., 2005. An evapotranspiration research facility for soil-plant-environment interactions. Applied engineering in agriculture, 21(6), pp.993-998.

Tsay, R.S., 1984. Order selection in nonstationary autoregressive models. The Annals of Statistics, 12(4), pp.1425-1433.

Vadez, V., Kholová, J., Hummel, G., Zhokhavets, U., Gupta, S.K. and Hash, C.T., 2015. LeasyScan: a novel concept combining 3D imaging and lysimetry for high-throughput phenotyping of traits controlling plant water budget. Journal of Experimental Botany, 66(18), pp.5581-5593.

Xiong, Y. and Yeung, D.Y., 2002, December. Mixtures of ARMA models for model-based time series clustering. In 2002 IEEE International Conference on Data Mining, 2002. Proceedings. (pp. 717-720). IEEE.