DEEP LEARNING BASED MODELS FOR GRAPEVINE PHENOLOGY

Authors: Federico OLDANI[1*], Dario SALZA[1], Giacomo BLANCO[1], Claudio ROSSI[1], Boris BASILE[2*], Fabrizio CARTENI[2], Núria PÉREZ-ZANÓN[3], Antonio DENTE[4], Fernando ALVES[5], Joana VALENTE[5], Montse TORRES[6], Carlos EZQUERRA[6], Rosa ARAUJO[7]

[1] *LINKS Foundation, Turin, Italy*
[2] *Department of Agricultural Sciences, University of Naples Federico II, Portici (Napoli), Italy*
[3] *Barcelona Supercomputing Center, Barcelona, Spain*
[4] *Mastroberardino, Atripalda (Avellino), Italy*
[5] *Symington Family Estates, V. N. Gaia, Portugal*
[6] *Familia Torres, Vilafranca del Penedès, Spain*
[7] *Eurecat, Centre Tecnològic de Catalunya, Barcelona, Spain*

*Corresponding author: federico.oldani@linksfoundation.com; boris.basile@unina.it

Abstract:

Context and purpose of the study – the phenological evolution is a crucial aspect of grapevine growth and development. Accurate detection of phenological stages can improve vineyard management, leading to better crop yield and quality traits. However, traditional methods of phenological tracking such as on-site observations are time-consuming and labour-intensive. This work proposes a scalable data-driven method to automatically detect key phenological stages of grapevines using satellite data. Our approach applies to vast areas because it solely relies on open and satellite data having global coverage without requiring any in-field data from weather stations or other sensors making the approach extensible to other areas.

Material and methods - we leveraged historical phenological observations and developed a supervised deep-learning model that uses the land surface temperature estimated by the Copernicus Sentinel-3 satellite to estimate the current phenological stage at the parcel level. We compared the performances of our model with traditional approach based on Growing Degrees Days (GDD).

Results – we train our algorithm on manually collected phenological observations of four winegrape cultivars in three European vineyards (Italy, Spain, and Portugal) from 2017 to 2022. Preliminary results indicated that our deep learning phenology model outperforms the traditional methods based on GDD, decreasing the Mean Absolute Error from 33.8 to 7.8 days (-76.5%).

Keywords: satellite imagery, earth observation, machine learning, Sentinel-3, Copernicus, climate change

1. Introduction

Phenology monitoring plays a critical role in grapevine cultivation, because it enables growers and vineyard managers to track the progress of key events in the grapevine life cycle, such as budburst, flowering, fruit set, veraison and leaf fall. These events are important because they determine the timing of critical operations, such as pruning, irrigation, fertilization, and pest and disease management that can impact the quality and quantity of grape yield. With the effects of climate change increasingly impacting the grapevine growth cycle, monitoring phenology is becoming even more critical. Changes in temperature, precipitation, and other environmental factors can cause variations in the timing of key phenological stages, which can impact grape quality and yield. Remote sensing technologies, such as satellite imagery, can provide a cost-effective and scalable way to monitor phenology across large areas, providing valuable information for growers and vineyard managers to make informed decisions and optimize their grapevine cultivation practices. Traditionally, phenological analysis has been conducted through on-site observations, which can be time-consuming and labor-intensive. The recent advances in satellite remote sensing have opened new possibilities for monitoring phenological changes in grapevines on a large scale. This study aims to develop a phenology model for grapevines using satellite data. The most used tools to predict phenology are temperature-driven models. Such models differ in the articulation between the thermal effects of the dormancy phase and the grapevine vegetative phase, which can be modeled as sequential, parallel, deepening rest, four phases, alternating etc. (Kramer et al., 1994; Chuine et al., 2013). Thermal models of the vegetative cycle are based on the occurrence of a given phenological stage through a sum of temperatures, starting on a pre-defined date and progressing until a given threshold is reached for each phase. Most models rely on the concept of growing degree-days (GDD) or heat units which are accumulated from a certain value called the base temperature (lower thermal limit). This value is conventionally set at 10 °C for grapevine (Winkler, 1995), other authors (Parker et al. 2011) recommended a value of 0 °C. Despite being very simple models, the degree-day approach also comes with important constraints (Garcia de Cortazar et al., 2009). More advanced models have been proposed, considering other vineyard information such as soil moisture and irrigation (Carteni et al., 2019).

2. Material and methods

Plant material and data sources

Plant materials - In this study, we focus on three grapevine (*Vitis vinifera* L.) cultivar across southern Europe: (i) 'Aglianico' cultivated in the province of Avellino (Mastroberardino); (ii) 'Syrah' in Aranyó (Familia Torres), which is located in Catalonia, Spain; (iii) the 'Touriga Nacional' collected at the Quinta do Ataíde estate (Symington Family Estates) , which is located in the Douro Valley region of Portugal.

Plant measurements - Phenology data relative to five stages (bud break, flowering, fruit set, veraison, and harvest) were collected in the vineyards for the three wine grape cultivars from 2017 to 2022. Budbreak, flowering, fruit set and veraison were collected using the Baggiolini scale as a reference (Baggiolini et al. 1952). For Symington, only three phenological stages were collected: budbreak, flowering, and veraison.

Data sources – Copernicus is the European Union's Earth observation programme coordinated and managed by the European Commission in partnership with the European Space Agency (ESA), the EU Member States and EU agencies. ESA is currently developing seven missions under the Sentinel programme (Sentinel 1, 2, 3, 4, 5P, 5, 6). Sentinel 3 provides high-accuracy optical, radar and altimetry data for marine and land services. Since it has been proven that the phenology has a strict correlation with the accumulated temperature, for this specific use case, the study focused on the output of SLSTR (Sea and Land Surface Temperature) sensor. The SLSTR thermal bands used to retrieve LST (Land Surface Temperature) are three infra-red channels S7, S8 and S9 at 3.74 µm, 10.85 µm and 12 µm). The twin satellites Sentine-3A and Sentinel-3B acquire data twice a day in free cloud conditions with a spatial resolution is 1 km^2.

Deep learning phenology model

Data processing – By gathering the satellite data measurements, it is possible to retrieve the historical time series of the temperature of the land surface for each Sentinel-3 channel (Figure 1 and Figure 2) in the considered period 2017-2022. In case of cloudy or misty weather, for example, the values are not reliable since the Sentinel-3 SLSTR acquires the temperature of the first layer that the wavelength encounters, in such cases the clouds. In these cases, we face temperatures below a normal range, and they are filtered out. The previous filtering might generate missing data, which is finally filled by injecting the moving average of the last 7 days. Moreover, since the first acquisition of the day is often missing (Figure 1), we chose to consider the second acquisition only (Figure 2) and given the strict correlation between the signal of the three bands (S7, S8, S9), we selected the band S9 because slightly more reliable in case of high temperature. In fact, S7 and S8 reach the maximum value in some cases, making their signals unreliable (Figure 1).

Model design - Recurrent Neural Networks (RNN) is a specific class of neural network able to handle time series keeping memories of the previous inputs (Rumelhart et al., 1985). They adapt well to time series because they are not limited to the actual input, but they also consider the temporal dependencies with the previous input values. Among the different RNN architectures, the Gated Recurrent Unit (GRU) (Kyunghyun, et al. 2014) has proven to perform well in the case of long-time series. We adopt a GRU neural network specific for each variety and location, which takes the temperature timeseries from January 1^{st} and detects the different phenological stages. The analysis has been implemented as a classification problem in which every class represents a different phenological phase. The output is expressed as a binary array containing zeros once the phase has been reached for each day of the year (Eibe Frank et al., 2001) (Figure 3).

3. Results and discussion

A GRU model with two hidden layers (30 neurons each) for every variety was trained using the AdamW (Loshchilov, 2017) optimizer and evaluated with 5-fold cross validation, using all data from four years (i.e. 2017, 2018, 2019, 2020) as training samples and one year (i.e., 2021) as test sample, and repeating for all the five possible combinations, averaging the results (Table 1). The results are compared with the Growing Degree Days (GDD) (Table 2) which has been calibrated with the same methodology and on the same temperature measurements from Sentinel-3. The Mean absolute error (MAE) obtained by the GDD is 33.2 days, the MAE obtained by the deep learning model is 7.8 days, with a reduction of -76.5% of days. The deep learning model has demonstrated to correlate the available information with the historical phenology data and provided a result that outperforms the traditional GDD model computed on the same data source, demonstrating the feasibility and effectiveness of using Sentinel-3 temperature data to monitor the phenological stages of grapevines. The GDD model assumes a linear relationship between temperature and plant development, however, this relationship is often not accurate, as plant growth and development are influenced by various factors, including soil moisture, solar radiation, and atmospheric CO_2 levels. The deep learning phenology model can learn complex and nuanced relationships between temperature and plant development stages and it is able to compensate the noise and the errors of the data source remaining even after the preprocessing operations. The approach suffered some limitations, such as the spatial resolution of the measurements (1 km^2) and the sampling frequency (once a day) that cannot give information about the average temperature of the entire day.

4. Conclusions

In this study, we developed a deep-learning phenology model for grapevines using satellite data. Remote sensing technologies, provide a cost-effective and scalable way to monitor phenology across large areas, providing valuable information for grape growers and vineyard managers to make informed decisions and optimize their grapevine cultivation practices. With the effects of climate change increasingly impacting the grapevine growth cycle, monitoring phenology is becoming even more critical. By using remote sensing technologies to monitor phenology, growers can monitor vast areas without installing in-field weather stations and take timely action to mitigate the impact on grape quality and yield. The presented approach could help to justify the investment in these technologies for smaller vineyards and growers who may not have the resources to invest in more traditional, labor-intensive phenology monitoring methods. The results of the study indicate that the deep learning phenology model using the satellite data is able to monitor the phenological stages of the grapevine and it outperforms the traditional method based on growing degree days (GDD) computed on the same dataset decreasing the Mean Absolute Error from 33.8 to 7.8 days (-76.5%). Overall, the use of satellite remote sensing for phenology monitoring shows great potential to revolutionize the grape-growing industry and help growers to adapt to the challenges of climate change.

5. Acknowledgments

This work has been partially funded by the EC through the VITIGEOSS project (GA 869565).

6. Literature cited

Kramer K (1994) Selecting a model to predict the onset of growth of Fagus sylvatica. J Appl Ecol 31:172–181
Chuine, I., de Cortazar-Atauri, I. G., Kramer, K., & Hänninen, H. (2013). Plant development models. Phenology: an integrative environmental science, 275-293.

Amerine, M., & Winkler, A. (1944). Composition and quality of musts and wines of California grapes. Hilgardia, 15(6), 493-675.

Parker, A. K., DE CORTÁZAR-ATAURI, I. G., Van Leeuwen, C., & Chuine, I. (2011). General phenological model to characterise the timing of flowering and veraison of Vitis vinifera L. Australian Journal of Grape and Wine Research, 17(2), 206-216.

Carteni, F.; Rossi, C.; Marcos, R.; Porras, I.; Basile, B.; Scognamiglio, P.; Teobaldelli, M.; Mataffo, A.; Mazzoleni, S.; Garcia-Tejera, O.; et al (2019). LEAF: A process-based model of berry ripening in vineyards. In Proceedings of the 2019 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor), Portici, Italy, 24–26 October 2019; pp. 152–157

Garcia De Cortazar-Atauri I, Brisson N, Gaudillere J (2009) Performance of several models for predicting budburst date of grapevine (Vitis vinifera L.). International Journal of Biometeorology, 53, 317–326.

Baggiolini, M. (1952). Les stades repères dans le développement de la vigne et leur utilisation pratique, Station Féd. Essais Agric., Lausanne.

Rumelhart, David E; Hinton, Geoffrey E, and Williams, Ronald J (Sept. 1985). Learning internal representations by error propagation. Tech. rep. ICS 8504. San Diego, California: Institute for Cognitive Science, University of California.

Cho, Kyunghyun; van Merrienboer, Bart; Bahdanau, Dzmitry; Bougares, Fethi; Schwenk, Holger; Bengio, Yoshua (2014). "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation". arXiv:1406.1078

Frank, E., & Hall, M. (2001). A simple approach to ordinal classification. In Machine Learning: ECML 2001: 12th European Conference on Machine Learning Freiburg, Germany, September 5–7, 2001 Proceedings 12 (pp. 145-156). Springer Berlin Heidelberg.

Loshchilov, I., Hutter, F. (2017) Decoupled weight decay regularization. In: 7th International Conference on Learning Representations, ICLR 2019.

Table 1 Mean absolute error of the deep learning model. It is computed as the absolute value of temporal distance in days between the outcome of the models and the phenology date observed on field.

MAE (Days, average of 5 folds)		Bud break	Flowering	Fruit-set	Veraison	Harvest	*Average*
Aranyó	**Syrah**	12.2	6.4	6.2	11.2	7.6	8.7
Ataíde	**Touriga Nacional**	15.0	7.0	-	8.6	-	10.2
Mirabella Eclano	**Aglianico**	8.2	4.6	2.8	1.6	5.8	4.6
Average		*11.8*	*6.0*	*4.5*	*7.1*	*13.4*	*7.8*

Table 2: Mean absolute error of Growing Degree Days used as a baseline. It is computed as the absolute value of temporal distance in days between the outcome of the models and the phenology date observed on field.

MAE (Days, average of 5 folds)		Bud break	Flowering	Fruit-set	Veraison	Harvest	*Average*
Aranyó	**Syrah**	14.0	16.8	33.0	65.0	54.0	36.6
Ataíde	**Touriga Nacional**	16.0	13.2	-	47.8	-	25.6
Mirabella Eclano	**Aglianico**	11.0	11.0	26.0	41.8	61.8	30.3
Average		*13.6*	*13.6*	*29.5*	*51.5*	*57.9*	*33.2*

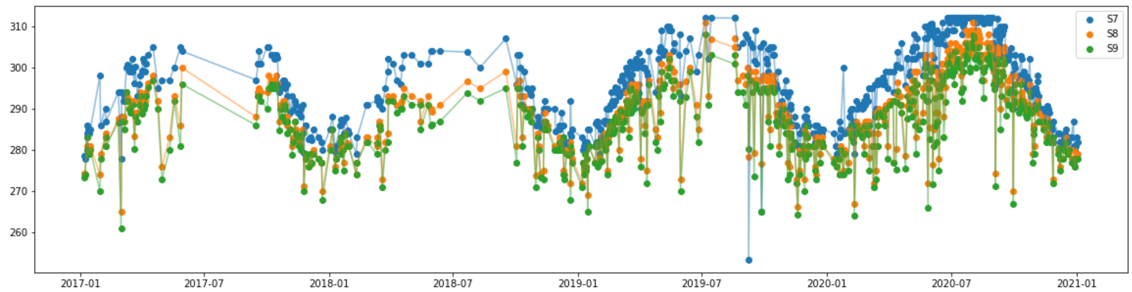


Figure 1: Temperature acquired by Sentinel-3 during the first acquisition of the day in the Torres vineyard. The Sentinel-3 thermal bands are three (S7, S8 and S9 at 3.74 µm, 10.85 µm and 12 µm wavelength respectively)

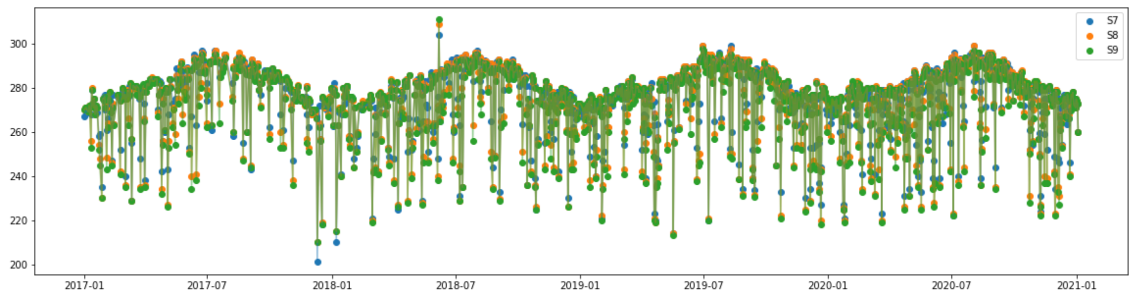
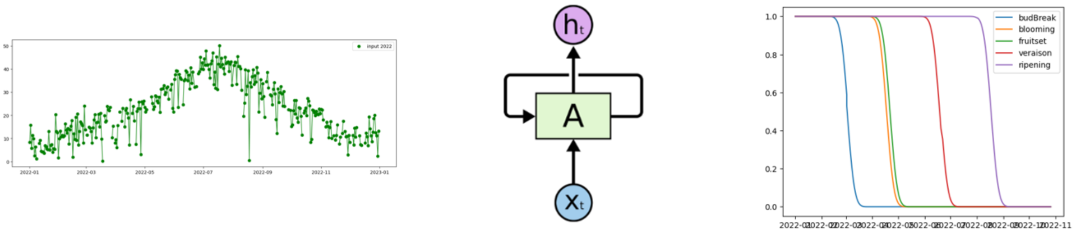


Figure 2: Temperature acquired by Sentinel-3 during the second acquisition of the day in the Torres vineyard. The Sentinel-3 thermal bands are three (S7, S8 and S9 at 3.74 µm, 10.85 µm and 12 µm wavelength respectively)



input: estimated temperature time series of the considered vineyard

processing: deep learning model (GRU) trained on the considered variety

output: sequence of phenological phases through time

Figure 3 Deep learning inference pipeline