A machine learning-based system for early and accurate vineyard yield forecasting

By Jose Carlos Caris-Maldonado, Jose Cuevas-Valenzuela, Ricardo Luna-Hernandez and Álvaro González-Rojas, Centre for Research and Innovation (CRI) Viña Concha y Toro, Ruta k-650 km 10, Pencahue, Maule, Chile

Vineyard yield forecasting is a key issue for vintage scheduling and optimising winemaking operations. Researchers from the Centre for Research and Innovation at Chilean-based wine company Viña Concha y Toro report on a trial they conducted in a high yielding Cabernet Sauvignon vineyard that used data from various sources such as satellite imagery, agro climatic conditions and historical yields. Machine learning models were then applied to generate yield forecasts at veraison.

IN BRIEF

- **Precise and timely information about yields is crucial to improve the efficiency of vineyard and winery operations, determine the need for investment in new winery equipment and make grape and wine purchase decisions.**
- **Currently, there is not a standardised methodology for achieving an appropriate vineyard yield prediction before veraison with an accuracy greater than or equal to 90%.**
- **Traditional forecasting methods show an average accuracy of around 70%-80% per vineyard block and are destructive, laborious, time consuming and/or expensive.**
- **Researchers from the Centre for Research and Innovation at Viña Concha y Toro proposed, developed and validated a new machine-learning-based system for yield forecasting before the completion of veraison and with an accuracy greater than 90% per block.**

INTRODUCTION

The digital transformation of agricultural processes implies a profound cultural change focussed on the acquisition and integration of various data sources from all critical points in the value chain. Application of new digital technologies can add value to information and allow better decisions based on data. At Viña Concha y Toro, the concept of digital transformation translates into connecting each of the stages and processes of the value chain to the digital world and generating value for the business through the capture, storage and processing of data integrated in an IIOT (Industrial Internet of Things) architecture.

Since its foundation in 2014, the Centre for Research and Innovation (CRI) has promoted the concept of digital transformation through the strategic program Smart Wine Industry and the application of engineering principles. The CRI seeks to promote applied research, technology transfer and innovation in order to ensure and maintain the productive excellence of Viña Concha y Toro, the sustainability of its processes and promote viticulture in the country.

The application of the concept of digital transformation in agricultural processes has been intensively addressed in academic and industrial fields, such as precision agriculture, which can be defined as a "management strategy that collects, processes and analyses temporal, spatial and individual data and combines them with other information to support management decisions according to the estimated variability, and thus improve the efficiency in the use of resources, productivity, quality, yield and sustainability of agricultural production "(ISPA 2019). In the case of the wine industry, this translates to precision viticulture, whose fundamentals have been used to optimise grape production processes, particularly those related to harvest forecasting.

Yield forecasting is one of the key elements in the viticultural value chain. It has the objective of predicting as accurately as possible the quantity of grapes for winemaking that will be harvested in a specific season (Dami & Sabbatini 2011, Sabbatini *et al.* 2012). Precise and on-time information about yield is crucial to improve the efficiency of vineyard and winery operations, to decide

SUPERIOR PROTECTION.

Nufarm

AUSTRALIAN THROUGH & THROUGH

The Tri-Base Blue Difference

Tri-Base Blue's unique liquid formulation makes measuring and pouring easy and dust-free. Providing reliable and sustained protection against various fungal diseases, it is availble in convenient 10L packs and 640L shuttles.

- Ultra-fine copper particle size enables superior coverage and enhanced disease control.
- 100% of the copper applied can release copper ions - 75% released immediately for fast protection, remaining 25% delayed release for extended duration of protection.
- Neutral pH formulation - ideal tank-mix partner for alkaline hydrolysis sensitive insecticides.
- Excellent crop safety.
- Ease of mixing and application.
- Only 1 day WHP to harvest provides flexibility with crop management.

nufarm.com.au

For more information, contact your local Nufarm Business Development Manager

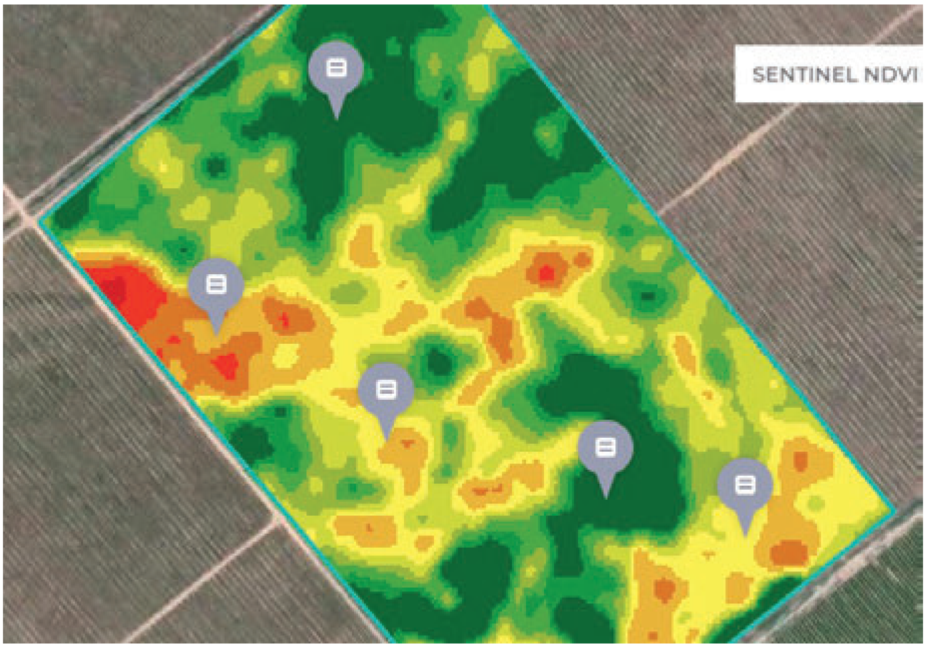

Figure 1. Satellite image corresponding to veraison in one of the blocks of Lourdes estate (Pencahue, Maule Valley, Chile), during season 2020.

on new investments in winery equipment, and to make grape and wine purchase decisions. It is also essential to successfully regulate vineyard yield, preventing under and over cropping and maintaining healthy and balanced vines each season (Bocco *et al*., 2015, Carrillo *et al*. 2016, Cunha *et al*. 2010, Dami & Sabbatini 2011, Diago *et al*. 2012, Dunn 2010, De la Fuente *et al*. 2015, Sabbatini *et al*. 2012).

Achieving a good result from yield forecasting is not easy. The high spatial variability of vineyards and increasingly unstable seasons due to climate change are among the factors that affect yield and make this task increasingly complex. Currently, there is not a standardised methodology for achieving an appropriate vineyard yield prediction before veraison and with an accuracy greater than or equal to 90% (Carrillo *et al*. 2016, Diago *et al*. 2012, Dunn and Martin 2004, Sabbatini *et al*. 2012). Today, forecasting methods (that we call 'traditional') show an average accuracy of around 70%-80% per vineyard block and they are destructive, laborious, time consuming and/or expensive (Blom and Tarara 2009, Diago *et al*. 2012, Dunn and Martin 2004). These traditional methods (TM) are based on manual grape cluster counting and weighing using a random sampling method that does not objectively consider spatial variability and is mainly based on the experience of managers.

Table 1. Relative accuracy(%) in grape yield forecasting reported for each method, traditional (TM) and ML-based method, after comparing against actual yield (in ton/ha). Standard deviation (SD) of the accuracy are also included. Forecasts were generated before veraison was completed (late January 2020) and actual yields were obtained during harvest (March-April 2020). In the case of the ML-based method, errors correspond to the application of the algorithm that presents the best performance between the three previously selected.

Block	Trellis system	Area(ha)	Actual yield (ton/ha)	TM (% accuracy)	ML-based method (% accuracy)
6	Free-cordon	21.8	13.6	89.7	92.6
10	Free-cordon	11.1	14.6	97.3	95.2
13	Free-cordon	9.77	12.4	79	91.1
29	Pergola	19.7	20.7	89.9	90.8
			MAPE(%)	89.9	92.4
			SD(%)	7.9	2.0

Therefore, improved methodologies are necessary for early and accurate vineyard yield forecasting. New methods, mainly based on digital technologies, have emerged in recent years as interesting alternatives for industrial applications. Most of them are supported by remote sensing technologies, such as satellite imagery, multispectral cameras coupled to a UAV (unmanned aerial vehicles) or mobile robots (multi-rotor and fixed wing drones). These technologies have been optimised and can be effectively implemented in vineyard operations to capture key data and characterise spatial variability (Matese *et al*. 2015, Santesteban 2019). In addition, computer vision and machine learning technologies are growing rapidly and they have demonstrated to be useful for analysing data and determining key vine and yield parameters (Di Gennaro *et al*. 2019, Kurtser *et al*. 2020, Liu *et al*. 2017, 2018, 2020, Nuske *et al*. 2014, Pérez-Zavala *et al.* 2018, Silver & Monga 2019).

As an alternative to traditional methods, we proposed, developed and validated a new machine-learning-based (ML-based) system for early and accurate vineyard yield forecasting, i.e. before the completion of veraison with a forecasting accuracy greater than 90% per block. We focus mainly on the characterisation of the spatial variability of the vineyard. Using this information we constructed a data acquisition protocol that takes this variability into account. Finally, we processed this information and developed the ML-based models to give a yield forecast.

METHODS

The development of the new harvest volume forecast system can be divided into five critical areas: (i) construction of a historical database; (ii) characterisation of vineyards spatial variability using satellite data; (iii) systematic cluster counting, sampling and weight measurements, based on the spatial variability (for the trial season); (iv) acquisition and integration of key agroclimatic data; and (v) the construction and application of yield forecasting models based on machine learning (ML) algorithms and collected data.

We tested and validated the new yield forecasting system in commercial vineyards belonging to Viña Concha y Toro (VCyT) in Chile. In particular, we carried out trials in high-yielding Cabernet Sauvignon (CS) vineyards located in Lourdes estate (Pencahue, Maule

Valley, Chile), during season 2020. We covered four blocks (66ha) and two trellis systems (pergola and free-cordon), which were selected because they present major challenges for forecasting in VCyT.

The main data sources were the historical harvest yield data from the VCyT databases for the blocks under study. These data corresponded to the type of trellis system, planting density, grape quality and its classification, among other data. Data from public agro-climatic stations were collected from the national meteorological network (Agromet). Here, different sources of data can be collected such as wind speed and direction, relative humidity, rainfall and air temperature. Finally, satellite image data were obtained from the Copernicus Open Access Hub, from the Sentinel-2 satellite (Figure 1). The images were downloaded for each season between the periods of November and January of the following year (before veraison). These data were structured, curated and added to the database.

Once the database was built, data preprocessing steps were carried out (normalisations and encodings), followed by feature selection and finally various machine learning algorithms were trained. The mean absolute percentage error (MAPE) was used as a performance measure. Subsequently, the best machine learning algorithm was selected to generate the harvest forecasts (tonnes/ha) for veraison (late January 2020). At the end of the season, the results of the yield forecast obtained with the machine learning-based system were compared with those obtained through the use of traditional methods, using the actual yield for each block.

RESULTS

The results showed that the ML-based system achieved a 92.4% average accuracy, whereas the traditional method obtained 89.9%. Moreover, the standard deviation of the forecasts was calculated as a measure of precision. In this case, the results showed that the method based on machine learning obtained 2%. Conversely, the standard deviation obtained by the traditional method was 7.9% (Table 1), indicating a lower precision in the harvest volume forecast.

CONCLUSIONS

The use of remote sensing technologies such as satellite imagery and weather stations are key to be able to perform sampling in a systematic way that represents the variability of different blocks and trellis systems. This, combined with the use of machine learning methods, has proven to be an effective tool to generate early and accurate yield forecasts.

We have obtained a cost-efficient, early and accurate new system for vineyard yield forecasting based on machine learning models.

The new method significantly reduces the forecasting yield accuracy above the threshold of 90% per block, giving an overall accuracy of 92.4% for the Cabernet Sauvignon blocks considered in the trial (before veraison is completed).

REFERENCES

Blom, P.E. and Tarara, J.M. (2009) Trellis tension monitoring improves yield estimation in vineyards. HortScience 44(3):678-685. https://doi.org/10.21273/hortsci.44.3.678

Bocco, M.; Sayago, S. and Violini, S. (2015)

Modelos simples para estimar rendimiento de cultivos agrícolas a partir de imágenes satelitales : una herramienta para la planificación. 2o Simposio Argentino Sobre Tecnología y Sociedad. Modelos 26-35.

Carrillo, E.; Matese, A.; Rousseau, J. and Tisseyre, B. (2016) Use of multi-spectral airborne imagery to improve yield sampling in viticulture. Precision Agriculture 17(1):74-92. https://doi.org/10.1007/s11119-015-9407-8

Cunha, M.; Marçal, A.;R.;S. and Silva, L. (2010) Very early prediction of wine yield based on satellite data from vegetation. International Journal of Remote Sensing 31(12):3125-3142. https://doi.org/10.1080/01431160903154382

Dami, I., & Sabbatini, P. (2011). Crop Estimation of Grapes. Fact Sheet - Agriculture and Natural Resources, Stage I, 1–2.

Diago, M.-P.; Correa, C.; Millán, B.; Barreiro, P.; Valero, C. and Tardaguila, J. (2012) Grapevine yield and leaf area estimation using supervised classification methodology on RGB images taken under field conditions. Sensors (Basel, Switzerland) 12(12):16988-17006. https://doi.org/10.3390/s121216988

Dunn, G. (2010) Yield Forescasting. GWRDC Innovators Network, June 2010, 6.

Dunn, G.M. and Martin, S.R. (2004) Yield prediction from digital image analysis: A technique with potential for vineyard assessments prior to harvest. Australian Journal of Grape and Wine Research 10(3):196-198 https://doi.org/https://doi.org/10.1111/j.1755-0238.2004.tb00022.x

ISPA (2019) Precision AG Definition, International Society of Precision Agriculture. Available in: https://www.ispag.org/about/definition.

Kurtser, P.; Ringdahl, O.; Rotstein, N.; Berenstein, R. and Edan, Y. (2020) In-field grape cluster size assessment for vine yield estimation using a mobile robot and a consumer level RGB-D camera. IEEE Robotics and Automation Letters 5(2):2031-2038. https://doi.org/10.1109/LRA.2020.2970654

Liu, S.; Cossell, S.; Tang, J.; Dunn, G. and Whitty, M. (2017) A computer vision system for early stage grape yield estimation based on shoot detection. Computers and Electronics in Agriculture 137:88-101. https://doi.org/10.1016/j.compag.2017.03.013

Liu, S.; Li, X.; Wu, H.; Xin, B.; Tang, J.; Petrie, P.R. and Whitty, M. (2018) A robust automated flower estimation system for grape vines. Biosystems Engineering 172:110-123. https://doi.org/10.1016/j.biosystemseng.2018.05.009

Liu, S.; Zeng, X. and Whitty, M. (2020) A vision-based robust grape berry counting algorithm for fast calibration-free bunch weight estimation in the field. Computers and Electronics in Agriculture 173:105360. https://doi.org/10.1016/j.compag.2020.105360

de la Fuente, M.; Linares, R.; Baeza, P.; Miranda, C. and Ramón Lissarrague, J. (2015) Comparison of different methods of grapevine yield prediction in the time window. Journal International Science de La Vigne et Du Vin 49 (September 2014):27–35. https://doi.org/10.20870/oeno-one.2015.49.1.96

Nuske, S.; Wilshusen, K.; Achar, S.; Yoder, L.; Narasimhan, S. and Singh, S. (2014) Automated visual yield estimation in vineyards. Journal of Field Robotics 31(5):837-860. https://doi.org/https://doi.org/10.1002/rob.21541

Pérez-Zavala, R.; Torres-Torriti, M.; Cheein, F.A. and Troni, G. (2018) A pattern recognition strategy for visual grape bunch detection in vineyards. Computers and Electronics in Agriculture 151:136-149. https://doi.org/10.1016/j.compag.2018.05.019

Sabbatini, P.; Dami, I. and Howell, G.S. (2012) Predicting Harvest Yield in Juice and Wine Grape Vineyards. Michigan State University Extension, 1-12 November. https://www.canr.msu.edu/uploads/resources/pdfs/Predicting_Harvest_Yield_in_Juice_and_Wine_Grape_Vineyards_(E3186).pdf

Silver, D.L. and Monga, T. (2019) In Vino Veritas: Estimating Vineyard Grape Yield from Images Using Deep Learning BT - Advances in Artificial Intelligence (M.-J. Meurs & F. Rudzicz (eds.); pp. 212–224). Springer International Publishing. WVJ