



# Innovative approaches to orchard management: assessing the variability in yield and maturity in a ‘Gala’ apple orchard using a simple management unit modeling approach

L. Manfrini<sup>1</sup>, L. Corelli Grappadelli<sup>1</sup>, B. Morandi<sup>1</sup>, P. Losciale<sup>2</sup> and J.A. Taylor<sup>3</sup>

<sup>1</sup> Department of Agricultural and Food Sciences, University of Bologna, Bologna, Italy

<sup>2</sup> Department of Soil, Plant and Food Sciences – DiSSPA, University of Bari “Aldo Moro”, Bari, Italy

<sup>3</sup> ITAP, INRAE, University of Montpellier, Montpellier, France

## Summary

Stratification of spatial data into management classes is a common way of interpreting and managing spatial agricultural data. High-resolution environmental and crop production information was collected within a 2.2-ha apple orchard (*Malus domestica* cv. Gala) near Sydney, NSW, Australia. Classifying the block into management units using the environmental data did not help to interpret the observed apple spatial variation in apple production. A backwards modelling approach was subsequently undertaken, effectively analyzing whether the crop production explains environmental variation. The backwards management unit model identified a different spatial patterning that revealed both an environmental and managerial effect on fruit production in different sections of the orchards. The unknown management effect had masked any spatial environmental-crop interactions in the initial (forward) management unit analysis. The combined forward and backward modeling approach significantly assisted the interpretation of the spatial variation in production in the orchard and identified a previously unknown management effect. When using management units to assess spatial relationships, the inclusion of a backwards modeling approach may help industry and growers to better understand and interpret spatial relationships and facilitate orchard management.

## Keywords

*Malus domestica*, precision horticulture, precision orchard management, spatial relationships

## Introduction

Apple yield is a function of fruit number and size, and both may vary significantly within an orchard (Manfrini et al., 2012, 2015; Aggelopoulou et al., 2010). However, it is difficult for fruit growers to quantify the amount of production variability within a season and to determine the factors that drive this variability spatially within an orchard. Better knowledge of within-orchard variability would facilitate better orchard management (Zude-Sasse et al., 2016).

Precision fruit growing techniques and spatial data sets are becoming increasingly available for orchard management (Manfrini et al., 2012; Käthner et al., 2017). With this comes a need to recognize that spatial data, particularly

## Significance of this study

*What is already known on this subject?*

- Precision fruit growing techniques and spatial data sets are becoming increasingly available for horticultural management. Spatial and temporal data have been used in fruit production with the aim at recognizing its variability and the implementation of sensor technologies. Management units are a common way of combining spatial environmental information in broadacre systems to explain spatial crop production patterns.

*What are the new findings?*

- Modeling using management zones to explain spatial crop variation using spatial environmental data does not always capture spatial managerial effects. Inverting this approach, a counter-intuitive approach to explain spatial environmental variation using spatial crop data, may help to uncover relationships between environmental and crop parameters that are impacted by differential management within orchards

*What is the expected impact on horticulture?*

- The implementation of management units to assess spatial relationships, including a backwards modelling approach, may help industry and growers to better understand and interpret spatial relationships and facilitate orchard management.

high-resolution sensor-derived data, has inherently different properties to agricultural data derived from ‘traditional’ plot experimentation. Spatial data needs a spatial analysis. This means that spatial data often needs to be pre-processed prior to analysis and the density of the data may often hide relationships. Increasingly, modeling approaches are being used to tease out relationships in these high density (high information) data sets. These are often sophisticated regression-based techniques (Usha and Singh, 2013; Peeters et al., 2015; Käthner et al., 2017) that are not easy to perform or interpret for orchard management.

Management units (MUs) have been used in site-specific horticultural management (Miranda et al., 2018; Aggelopoulou et al., 2013; Taylor and Whelan, 2011; Paoli et al., 2007). Recent advances in sensor technologies, such as soil sensors, canopy sensors and fruit/yield monitors, can now provide

many different layers to help derive these management units (Bresilla et al., 2019; Boini et al., 2019; Losciale et al., 2015). In many cases horticulture producers already have some of these data layers but do not have the tools to apply them spatially to crop management (Praat et al., 2001; Taylor et al., 2007a; Manfrini et al., 2009; Zude-Sasse et al., 2016). Spatial relationships between crop and environmental parameters (layers) can be modeled to assist in understanding correlations and determining potential causal effects in the data. However, a simpler, more effective approach to help in the initial interpretation of spatial data may be to use a management unit model. Management unit models integrate multiple data layers into a single layer that is often easier for practitioners not familiar with spatial data analysis to analyze and interpret. Management units are generally considered as a means of implementing differential (zonal) management, however, they can also be used as an analysis tool to help interpret spatial relationships. Experiences in dryland production systems have shown that within-field production variation is strongly influenced by within-field environmental variation, particularly edaphic variation (e.g., Kitchen et al., 2003). Similar experiences have been reported in viticulture systems, although these systems tend to be placed under water-stress during ripening, thus emphasizing the effect of edaphic variation on production (e.g., Tisseyre et al., 2007). In horticulture crops, where irrigation is non-limiting, the influence of soil type may be less obvious.

Information of orchard variability in apple is scant with only a few studies published to date (Manfrini et al., 2009; Aggelopoulou et al., 2010; Gemtos et al., 2013; Liakos et al., 2015; Käthner et al., 2017). There is an increasing demand

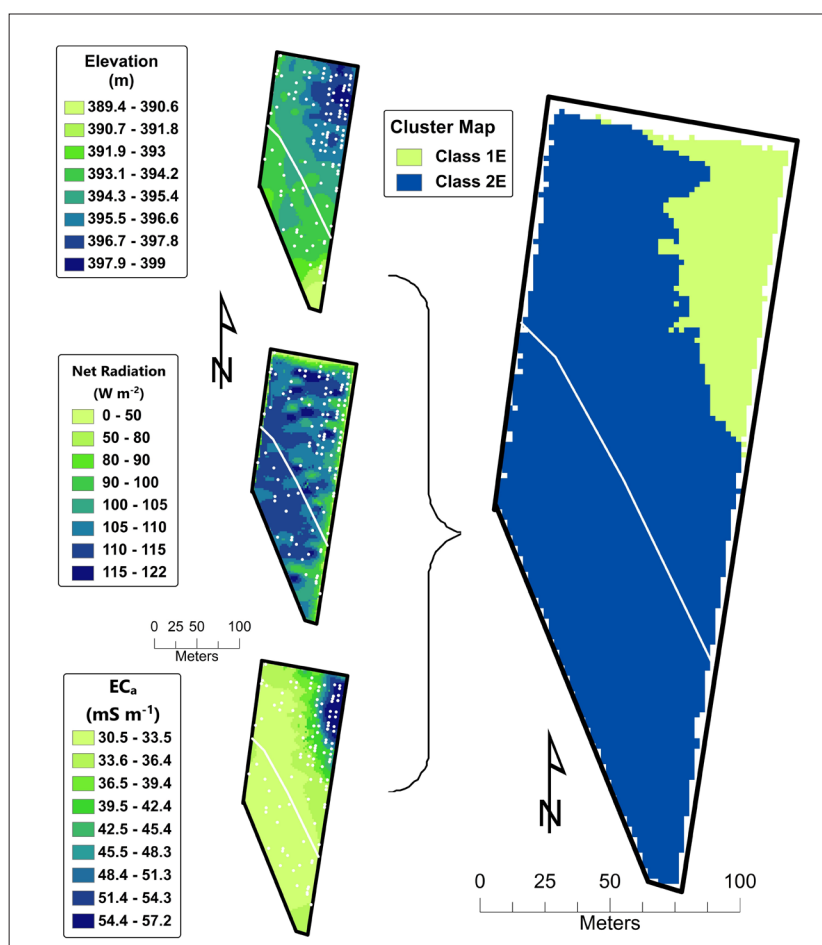
from the apple industry for information about orchard variability and protocols that may facilitate differential management (Manfrini et al., 2009, 2015), but this needs to be supplemented with effective ways for end-users to interpret the data.

In this study, a 2.2-ha apple 'Gala' orchard in Australia was used to test the utility of using management units to interpret spatial apple orchard data. The basic hypothesis is that a management unit analysis is able to identify if and how production variation is related to environmental variation, particularly soil type. However, there is a recognition that spatial variation is often confounded in agricultural data sets by managerial as well as environment effects and their interactions. For this reason, this study proposes using the concept of MUs as a means of spatially organizing and interpreting the data using both a forward and backward management unit modeling approach. Mixing both forwards and backwards modelling is a novel approach to deriving and interpreting MUs in precision agriculture. Backward or inverse modeling, as the name suggests, is simply a process of switching the dependent and independent variables in a model. This is typically applied in precision agriculture to crop simulation models (Florin et al., 2008).

## Materials and methods

### Site description and environmental data layers

The study was conducted during the season 2007–2008 in a commercial 'Gala' apple (*Malus domestica* Borkh.) orchard located at Darkes Forest (near Sydney), NSW, Australia (Lat. -34.227327°, Long. 150.907844°, elevation 373 m). The



**FIGURE 1.** Maps of the *a priori* environmental data for the study site on the left-hand side (from top to bottom) Elevation, soil apparent electrical conductivity ( $EC_a$ ), and net radiation compressed into a two-cluster management class map ( $Cluster_E$ ) using *k*-means clustering (right-hand side). The two-cluster map identifies also the breakdown into two different regions (North-East and North-West/South).

trees were 12 years old and grafted onto M26 rootstock. They were planted 3.0 m apart in rows 5.0 m apart and trained to a vertical axis in a north-south orientation. 'Granny Smith' and 'Pink Lady' pollinators were planted every three rows. The pollinator rows were omitted from any observations and analysis. The orchard was managed according to commercial practices including drip irrigation, fertilization and fruit thinning. 'Gala' is considered a suitable crop for this 1-year study as it is a non-alternate bearing cultivar and the data coming from a single growing season is valid for management in the following season.

The orchard was chosen for its ease of access and the existence and availability of *a priori* environmental information. This *a priori* information consisted of a soil apparent electrical conductivity ( $EC_a$ ) map (performed with a Geonics EM38 (Geonics Ltd., Mississauga, ON, Canada)) and a high-quality elevation map (from a carrier-phase global navigation satellite system (GNSS) receiver) collected commercially (Terrabyte, Wagga Wagga, NSW, Australia). This data was used to identify a 2.2-ha block of apple trees within the orchard that exhibited a large variation in  $EC_a$ . The block was dissected north-west/south-east by an internal track for machinery. This split the block into two sections, with an approximate ratio of 1:2 between the south and north sections (Figure 1).

The following derived environmental information layers were available for the orchard block:

- A Digital Elevation Model (DEM) of the orchard on a 2-m grid interpolated using local block-kriging with a linear variogram model.
- An  $EC_a$  map, interpolated onto the same 2-m grid using local block kriging with an exponential variogram model.
- A surface of modelled Net Radiation Interception. The incident net solar radiation (NR) modeled using the SRAD program (McKenney et al., 1999) using a modified Digital Surface Model (DSM) based on the interpolated DEM. The northern and eastern boundaries of the 2.2-ha block are bordered by native Eucalyptus (*Eucalyptus* sp.) forest with an approximate height of 10 m. The DSM was adjusted to reflect this and the potential effect that shading, particularly early morning shading, may have on NR interception.

#### Forward MU model and crop sampling scheme

The three data layers ( $EC_a$ , Elevation and NR) were clustered, using a *k*-means clustering algorithm after the protocol of Taylor et al. (2007b), and the resulting map was used to determine samples within the orchard. A range of *k* values were used (2–5) and evaluated based on the spatial arrangement of the resulting units (i.e., is it sensible to manage?) and the difference in mean values of the environmental variables between classes. The 2-means result was selected as the most appropriate. The 2-means map highlighted distinct classes in the orchard; the higher  $EC_a$  and elevation 'hill' in the north-east corner (Class 1<sub>E</sub>) and the remainder of the block (Class 2<sub>E</sub>). Class 1<sub>E</sub> and Class 2<sub>E</sub> occupied 0.4 and 1.8 ha, respectively. The subscript 'E' has been used in this notation to identify that these classes were derived from the environmental layers.

The sampling scheme was defined by randomly selecting 89 trees across the orchard with a slight weighting to increase the sample size in the smaller Class 1<sub>E</sub>. From these 89 trees, 15 were chosen randomly and the adjacent tree to the south of these 15 trees was monitored. By pairing trees, the sampling ensures that the short range stochastic variation in production per tree is measured and can be accurately mod-

eled in any spatial analysis. Therefore, in total there were 104 sample trees (shown as white dots in Figures 1 and 2) with 41 in Class 1<sub>E</sub> and 63 in Class 2<sub>E</sub>.

#### Crop production data layers

The selected trees were monitored for yield components at harvest. Harvest was conducted over 5 days (12<sup>th</sup>, 14<sup>th</sup>, 15<sup>th</sup>, 19<sup>th</sup> and 21<sup>st</sup> February 2008). The initial trees were stripped picked. This corresponded to 19 trees in the most easterly row. Crop load ((FN) fruit tree<sup>-1</sup>) and yield (FW (kg tree<sup>-1</sup>)) were measured and then the average fruit weight (MFW (g)) was calculated. After this first row, the harvest strategy was altered and only saleable (mature) fruit was picked during the first pass. Unripened fruit was left on the tree. Subsequent passes through the orchard were made over the next 10 days until all the fruit were removed. The same parameters (FN, FW and MFW) were measured. This approach also allowed a maturity index calculation described by the fruit percentage picked at the first date (FP%). This was applied for the remaining 85 trees providing an indication of ripening within the orchard. All the harvest data is reported on a single tree basis.

The trunk circumference (TC) 20 cm above the graft union was measured during the season as an indication of tree size. The trees were georeferenced with a standalone Garmin GPS76 Global Positioning System (GPS) receiver.

To complete the spreadsheet, the environmental data ( $EC_a$ , elevation and NR) and corresponding 'E' class (Class 1<sub>E</sub> or Class 2<sub>E</sub>) was extracted at the sampling points (tree locations) and appended to the manually collected orchard data.

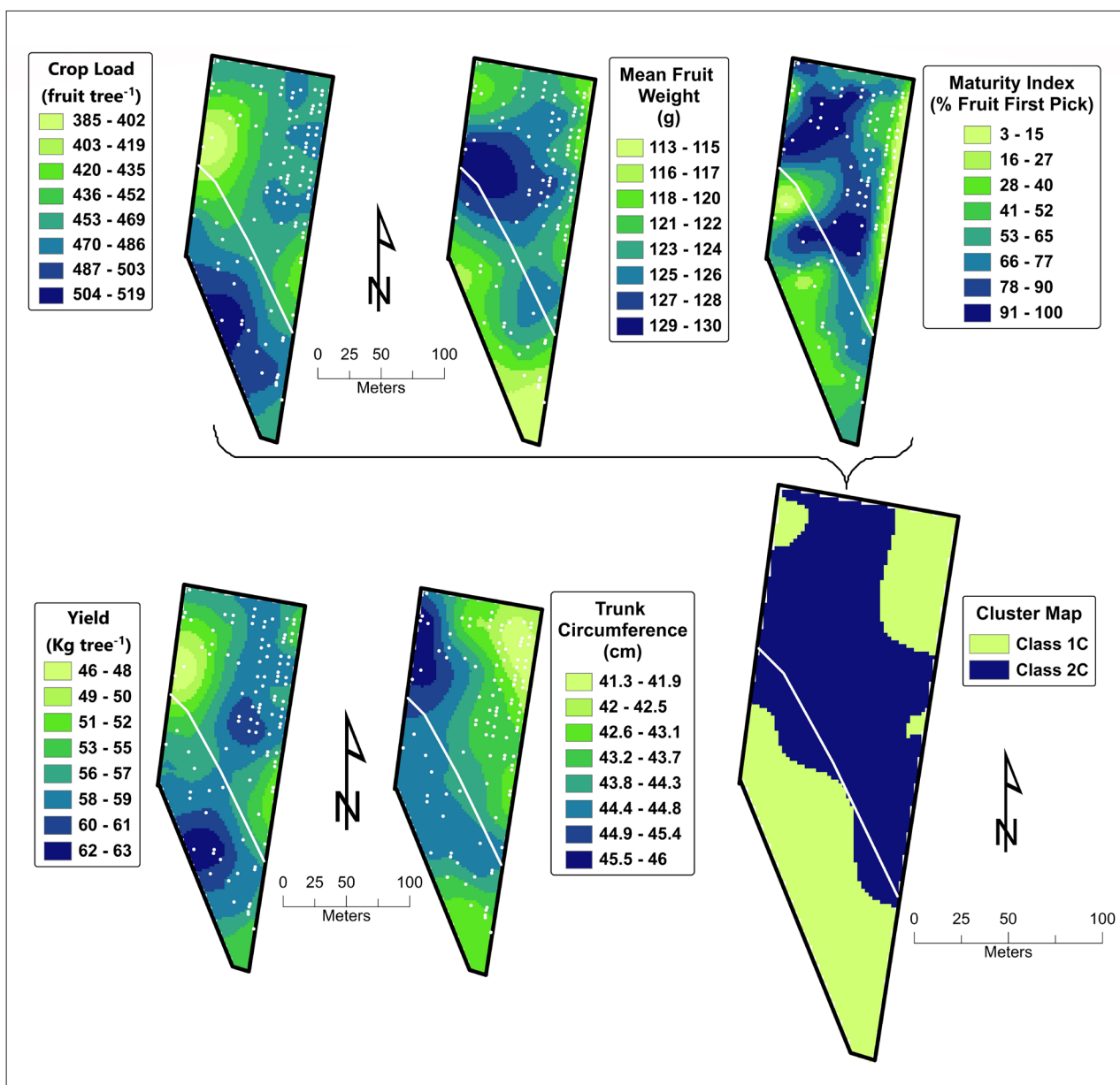
Maps of crop production parameters were generated by punctual kriging with a global variogram using the Vesper freeware (Minasny et al., 2005). The interpolation was performed onto the same grid used for the environmental parameters. Surfaces of FN, FW, MFW and TC were kriged using all 104 points while FP% was kriged using only the 85 points, where FP% was recorded. Maps were created in ArcGIS (ESRI, Redlands, CA, USA).

#### Data analysis

The primary hypothesis testing revolved around whether or not production variation is driven by environmental variation. As a preliminary investigation, Pearson's correlation analysis was performed between yield component, FP%, TC and  $EC_a$ , elevation and NR. With a fairly dense spatial data set (104 pts over 2.2 ha), autocorrelation is expected to affect these data (Taylor and Bates, 2013). Therefore, the method of Dutilleul (1993) to adjust the sample size to account for any auto-correlation in the processes was used to assess significance ( $p < 0.1$ ) with the correlation analysis. The correlation analysis was performed using the PASSaGE v.2 shareware suite (Rosenberg and Anderson, 2011).

#### MU Analysis 1 – Harvest data as the independent variable(s)

The crop parameter sampling scheme was based on the environmentally derived management units. An unbalanced ANOVA ( $n = 41$  and  $63$  in Class 1<sub>E</sub> and Class 2<sub>E</sub>, respectively) was performed in JMP 6.0 (SAS Institute) on the point production data (FN, FW, MFW) ( $n = 104$ ) using the two 'E' classes as treatment effects and Tukey-Kramer's means test. This investigates how the environmental variables explain the variation observed in the crop data. It is expected that the local environment will impart some influence on plant productivity at each individual site. The ANOVA yielded no



**FIGURE 2.** Production maps derived from the manual sampling. From top left maps are crop load (fruit tree<sup>-1</sup>), mean fruit weight (g), maturity index (% of fruit harvested at the first pick), yield (kg tree<sup>-1</sup>), trunk circumference (cm) and the Cluster<sub>C</sub> map derived from the MFW, FP% and FN data. The Class 1<sub>C</sub> consists of 4 discrete zones, with two of the zones large enough to manage – the North-East Zone and the Southern Zone. Class 2<sub>C</sub> is one contiguous ‘center’ zone.

significant difference in harvest parameters between the two classes (see results for full analysis). A three-cluster approach was also tried (map and results not shown) but again showed no significant difference in harvest parameters between classes.

**MU Analysis 2 – Environmental data as the independent variable(s)**

As an alternative to MU Analysis 1, and to further interrogate the data, a backwards modeling approach was also performed, i.e., do crop variables explain environmental variation? It is not expected that crop variables will influence the local environment (at least not in a relatively young production system). However, a ‘backwards’ approach may provide alternative (different) information than MU Analysis 1, which may help explain the spatial response in the orchard. The backward management class modeling was performed using *k*-means clustering on the interpolated crop data (FN,

MFW and FP%), again using the protocol of Taylor et al. (2007). These management classes were labeled with the subscript ‘C’ denoting that they are derived from crop parameters, i.e., Class 1<sub>C</sub> to differentiate them from the ‘E’ classes. Again, a range of *k*-values was tested, and the 2-means result was again determined to be the best model using the same criteria that was used on the ‘E’ classes. The 2-means ‘C’ cluster map is shown in Figure 2. Standard comparisons of class mean, such as the Tukey-Kramer test, cannot be used on spatial data, especially interpolated data, where the assumption of independent errors is not valid. In high density spatial data sets, neighboring points are generally auto-correlated, thus the number of independent points does not equal the number of points. Instead, means comparison was assessed using the modified confidence interval of Cupitt and Whelan (2001) (Eq. 1). This statistic determines if the differences between class means is greater (or less than) the mean error associated with interpolation. If the error is greater than it

is possible for the difference in classes to be an artifact of the interpolation procedure. Means comparisons could not be performed with the modeled NR data as it does not have a kriging variance, a limitation to the statistic of Cupitt and Whelan (2001).

The modified C.I. of Cupitt and Whelan (2001) is calculated as (after Taylor et al., 2007b):

$$95\% \text{ C.I.} = \mu \pm \left( \sqrt{\sigma_{\text{krig}}^2} \cdot 1.96 \right) \quad (\text{Eq. 1})$$

Where  $\sigma_{\text{krig}}^2$  is the median kriging variance, and for statistically different classes

$$|Y_{\text{zone}_x} - Y_{\text{zone}_w}| (\sigma_{\text{krig}} 1.96) \cdot 2 \quad (\text{Eq. 2})$$

where  $Y_{\text{zone}_x}$  is the mean of Class/Zone  $x$ ,  $Y_{\text{zone}_w}$  is the mean of Class/Zone  $w$  and  $\sigma_{\text{krig}}$  is the median kriging standard deviation.

## Results

### Preliminary data analysis

The results from the multivariate correlation analysis using Dutilleul's correction are shown in Table 1. Yield components (FN, FW and MFW) showed no correlation with the individual environmental parameters. The maturity index (FP%) was negatively correlated with  $EC_a$  ( $r = -0.49$ ) and positively correlated with net radiation ( $r = 0.44$ ). There was a negative correlation between  $EC_a$  and trunk circumference ( $r = -0.38$ ).

### Description of the crop attribute maps

The kriged maps for the measured crop parameters and TC are represented in Figure 2 and showed an independent pattern to the initial environmental clustering map (Figure 1). A high productive area ( $FW > 60 \text{ kg tree}^{-1}$ ) is visible in the southwestern side of the orchard while a lower productive area ( $FW < 50 \text{ kg tree}^{-1}$ ) is located in the central northwestern section of the orchard (Figure 2). The FN map followed a similar pattern. The MFW map (Figure 2) showed the same pattern, but inverted, with larger fruit ( $> 126 \text{ g}$ ) in the areas

with lower FW and FN and smaller fruit ( $< 115 \text{ g}$ ) where FW and FN were higher.

To evaluate changes in the maturity patterns, a map of FP% was produced. The FP% map showed distinct spatial patterns that differed from the other yield components maps. There was a slow maturing fruit (low FP%) area along the eastern edge. The southwestern section of the orchard was also slow maturing and aligned with the higher FN and FW area. The TC map also described a clear pattern: trees on the north-west side of the orchard showed a larger TC while the north-east and the south seems to have a smaller size. The trend in TC in the northern section of the orchard follows the spatial trend exhibit by the  $EC_a$  and elevation maps (Figure 1). The NR surface (Figure 1) showed less radiation interception in the eastern and northern sections of the block.

### Results of Approach 1 – Effectiveness of MUs based on environmental variables

The Cluster<sub>E</sub> map, derived from analysis of the environmental data, is shown in Figure 1 and the mean response of both the environmental data and crop yield parameters for the two clusters are shown in Table 2. The clustering performed on the 'E' data (Figure 1) divided the orchard into two blocks corresponding to high/low  $EC_a$  and high/low elevation. This also shows the breakdown into two main regions: North-East and South (Figure 1). The ANOVA showed that there was no statistical difference in mean crop production responses between these two 'E' classes. The harvest parameter maps (Figure 2) showed strong spatial patterns but these patterns did not visually correlate with the Cluster<sub>E</sub> map (Figure 1).

### Results of Approach 2 – The backward modeling approach

The Cluster<sub>C</sub> map, derived from the cluster analysis of the FN, FW and FP% parameters, is shown in Figure 2. In the Cluster<sub>C</sub> map, Class 1<sub>C</sub> is separated into 4 discrete zones whilst Class 2<sub>C</sub> is contiguous. The class means and comparisons are given in Table 3. The two smallest zones in Class 1<sub>C</sub> (mid-eastern and north-western corner) contained no or few sample points and are probably too small to be considered

**TABLE 1.** Correlations between environmental parameters and crop production parameters from the 104 sample sites assessed. Significant relationships indicated in bold.

Environmental parameter	Trunk circumference (cm)	Yield (kg tree <sup>-1</sup> )	Maturity index (% fruit pick first harvest)*	Crop load (fruit tree <sup>-1</sup> )	Mean fruit weight (g)
Elevation (m)	-0.27	0.05	-0.03	0.01	0.15
$EC_a$ (mS m <sup>-1</sup> )	<b>-0.37</b>	0.01	-0.24	0.01	0.01
Net radiation (W m <sup>-2</sup> )	0.21	0.04	<b>0.37</b>	0.04	<b>0.34</b>

\* Sample size ( $n$ ) = 85 for the FP% data.

**TABLE 2.** The mean response per class for the environmental variates used in the initial cluster analysis and the corresponding mean crop production parameters based on hand sampling within the two clusters (zones).

	Environmental data <sup>†</sup>			Crop production data* <sup>†</sup>		
	Elevation (m)	$EC_a$ (mS m <sup>-1</sup> )	Net radiation (W m <sup>-2</sup> )	Yield (kg tree <sup>-1</sup> )	Crop load (fruit tree <sup>-1</sup> )	Mean fruit weight (g)
Class 1 <sub>E</sub> (N=41)	397.09	48.54	103.99	55.90	461.98	121.95
Class 2 <sub>E</sub> (N=63)	394.38	33.99	105.32	56.35	465.65	121.64
Significance	Yes	Yes		No	No	No

<sup>†</sup> Significance between means determined using Cupitt and Whelan's statistic.

\* Significant differences between the crop production means was assessed by ANOVA using Tukey's means comparison with  $p < 0.05$ .

**TABLE 3.** The mean class and zone response of the interpolated crop production environmental data from the Cluster<sub>c</sub> analysis. The first two rows represent the class responses while the bottom three rows represent the zone responses where Class 1<sub>c</sub> has been split into two zones.

ID	N	Crop production data <sup>†</sup>			Environmental data <sup>*</sup>		
		Crop load (fruit tree <sup>-1</sup> )	Mean fruit weight (g)	Maturity index (% fruit picked first harvest)	EC <sub>a</sub> (mS m <sup>-1</sup> )	Elevation (m)	Net radiation (W m <sup>-2</sup> )
Class 1 <sub>c</sub>	51	487.2 <sup>1</sup>	118.7 <sup>1</sup>	60.2 <sup>1</sup>	36.49 <sup>1</sup>	394.26 <sup>1</sup>	106.96
Class 1 <sub>c</sub> North-East	27	468.6 <sup>ab</sup>	121.7 <sup>ab</sup>	65.1 <sup>b</sup>	50.05 <sup>a</sup>	397.45 <sup>a</sup>	104.36
Class 1 <sub>c</sub> South	21	485.2 <sup>a</sup>	119.2 <sup>b</sup>	64.6 <sup>b</sup>	32.28 <sup>b</sup>	393.23 <sup>b</sup>	107.91
Class 2 <sub>c</sub>	53	442.1 <sup>2,b</sup>	124.7 <sup>2,a</sup>	81.8 <sup>2,a</sup>	34.32 <sup>1,b</sup>	394.93 <sup>1,b</sup>	107.80

Different superscript numbers indicate significant differences between classes whilst different superscript letters indicate significant differences between the two zones in Class 1<sub>c</sub> and Class 2<sub>c</sub> (contiguous center zone in Figure 2).

<sup>†</sup> Significant differences between the crop production means were assessed by Tukey means comparison with  $p < 0.05$ .

<sup>\*</sup> The confidence interval of Cupitt and Whelan was used to determine significance with the environmental data. C.I. for EC<sub>a</sub> and Elevation were 3.94 and 0.94, respectively.

as management units. It is ignored for subsequent analysis and discussion. The 2-class outcome did not yield significant differences between classes in the environmental data response (Table 3).

When Class 1<sub>c</sub> was deconstructed into the two dominant zones – a north-east zone and a southern zone – a different picture emerged. The means comparison showed no significant differences in crop production between the north-east and southern zones of Class 1<sub>c</sub> (as can be expected given that classification was based on this data). However, the environmental data was significantly different. The southern zone of Class 1<sub>c</sub> has a lower EC<sub>a</sub> and elevation than the north-east zone. There was no significant difference in EC<sub>a</sub> and elevation between this southern zone of Class 1<sub>c</sub> and Class 2<sub>c</sub>, but a significant difference in all crop production parameters (Table 3).

## Discussion

### Discussion on preliminary data analysis and crop attribute maps

The multivariate correlation analyses (Table 1) only showed a few significant relationships between the environmental parameters (elevation, EC<sub>a</sub> and Net Radiation) and tree and crop production parameters. The EC<sub>a</sub> was correlated to trunk circumference, indicating that the long-term development of trees was linked to soil variability. However, the in-season correlations between EC<sub>a</sub> and crop parameters were non-significant and close to 0 for yield, crop load, and mean fruit weight. The NR was significantly correlated to MFW and FP% indicating that there were in-season spatial effects of net radiation on production. Visually, this can be seen in the maturity patterns displayed in the FP% map (Figure 2) that show a slow maturing fruit zone along the eastern edge. This is hypothesized to be associated with morning shading due to the Eucalyptus forest along the northern and eastern side of the block. The maps show spatial relationships between crop parameters that are to be expected from a physiological perspective (Palmer et al., 1991; Wünsche et al., 2000; Lopez et al., 2018). There was a high-yield area due to a high crop load in the southwestern side of the orchard and a lower productive zone, associated with a low FN, located in the central northeastern section of the orchard. The MFW map followed a similar but inverted pattern, reflecting a probable result of managerial effect at the within-orchard level. In addition, the southwestern section of the orchard also had slow maturing fruit and aligned with the higher FN and FW area. Late maturity in this area was likely due to

the larger ripening load placed on the trees as described by Wünsche et al. (2000).

### Discussion of Approach 1 and 2

In Approach 1, the Cluster<sub>E</sub> map derived from analysis of the environmental data clearly shows a delineation between the high elevation and high EC<sub>a</sub> in the northeast corner of the block and the remainder of the block. It was hypothesized that crop production would be influenced by this pattern, but the seasonal data did not support this. The backward model (Approach 2), derived from clustering the FN, FW and FP% parameters (Figure 2), produced more segmented classes (Cluster<sub>c</sub> map). Class 1<sub>c</sub> was separated into 4 discrete zones, of which 3 can be considered manageable, whilst Class 2<sub>c</sub> was contiguous. This contrasted with the Cluster<sub>E</sub> map, which had 2 classes and 2 zones. The first obvious feature of the Cluster<sub>c</sub> map was the zone in the north-east corner that corresponded to the area of high EC<sub>a</sub> response (and Class 1<sub>E</sub>). The second is that the large south zone in Class 1<sub>c</sub> encompassed the area in the block south of the internal machinery track.

With Approach 2, the classes derived from the clustering of crop production data (Table 3) did not yield significant differences in the environmental data response. This reflected the results found in Approach 1. However, deconstructing Class 1<sub>c</sub> into its main zones did show significant differences (Table 3). The EC<sub>a</sub> and NR response of the south zone of Class 1<sub>c</sub> was not significantly different from the center zone (Class 2<sub>c</sub>) (Figure 2). However, the crop response between the centre and south zones is significantly different for all three parameters (FN, MFW and FP%) (Table 3). There is an unidentified effect(s) in the centre zone that reduced FN and produced larger earlier ripening fruit. While this effect may be environmental, an observation of the patterning in the cluster map indicated that it is more likely to be managerial. The boundary between the centre and south zones followed the internal track. This can be considered a strong fit given that the cluster map (and backwards modeling approach) was derived from the low-density 104 sampling points, making boundary locations uncertain.

Sources of the managerial effects on production may be diverse and could range from different treatments between the two sections during block establishment, different management within this season or possibly different management in previous seasons. With perennial systems it is possible for management effects to linger for several years, for example, a different pruning strategy in year 1 will affect canopy and fruit development in year 2. This in turn influ-

ences pruning at the end of year 2, which could influence fruit development in year 3, and so on. Although the orchard manager believed that the two sections were uniformly managed, further investigation and discussion are needed to identify likely sources of the spatial pattern in the block. This is a key part of any spatial analysis and the next step in developing a spatial management strategy for the orchard, such as targeting differential, spatial pruning, thinning and/or harvest strategies based on variable fruit set and maturity patterns to meet market targets. In this specific case, the NR area that showed low radiation interception in the eastern and northern sections of the block, could be heavily thinned to overcome the late maturity and obtain larger fruit. The same could be said for the Class 1<sub>c</sub> area where a combination between crop load and tree vigor probably did not permit production to achieve a higher-class fruit size. However, the size, density and interactions within spatial crop data often hides these spatial relationships.

The environmental differences between the north-east and south zones offset each other and hide the differences between Class 1<sub>c</sub> and Class 2<sub>c</sub>. The backwards Cluster<sub>c</sub> modeling approach yielded a different spatial structure, which, on first analysis again, yielded no statistical differences. However, interpretation of the results on a zonal, rather than class, basis indicated that the management classes (zones) are segregating crop variation and may be useful as a means for future investigations. If the analysis was restricted to the preliminary multivariate correlation analysis and management class model, then the conclusion for this block would be that EC<sub>0</sub>, elevation and NR are not indicators (or drivers) of production variation. The backwards modeling revealed that the environmental parameters do affect production, however the relationship is complex and probably influenced by management. To understand these spatial relationships better, further information on crop management is needed.

Over time, temporal and spatial changes could provide effective feedback for precision horticulture and for growers on how persistent and responsive the patterns within an orchard are depending on management treatments and environmental conditions. If normal treatments are planned as homogeneous, it is rather common to find zonal pattern in a field. That may indicate an environmental implication or inconsistent managing treatments.

## Conclusions

Different approaches are often needed to uncover relationships between environmental and crop parameters. In this study a backwards modeling of management classes is used as a tool for achieving this result. A typical non-spatial analysis was undertaken but showed low or no correlation between environmental and production/quality parameters within the 'Gala' orchard. A conventional (forward) management class model also showed no spatial relationships. However, a backwards modeling approach, coupled with some local knowledge, identified relationships between the production, managerial and environmental conditions.

This result underlined that spatial variation does exist in apple orchards and there may be differential management options (i.e., differential management operation such as picking, thinning and pruning) available to growers who are aware of this variation, to increase their potential to respond to market demands for higher quality fruit without additional cost. Advisors and growers should take the time to apply various data mining and modeling techniques to gain the best understanding of the data possible.

## References

- Aggelopoulou, K.D., Wulfohn, D., Fountas, S., Gemtos, T.A., Nanos, G.D., and Blackmore, S. (2010). Spatial variation in yield and quality in a small apple orchard. *Precision Agric.* *11*, 538–556. <https://doi.org/10.1007/s11119-009-9146-9>.
- Aggelopoulou, K., Castrignanò, A., Gemtos, T., and De Benedetto, D. (2013). Delineation of management zones in an apple orchard in Greece using a multivariate approach. *Comput. Electron. Agric.* *90*, 119–130. <https://doi.org/10.1016/j.compag.2012.09.009>.
- Boini, A., Manfrini, L., Bortolotti, G., Corelli-Grappadelli, L., and Morandi, B. (2019). Monitoring fruit daily growth indicates the onset of mild drought stress in apple. *Sci. Hortic.* *256*, 108520. doi:10.1016/j.scienta.2019.05.047.
- Bresilla, K., Perulli, G.D., Boini, A., Morandi, B., Corelli Grappadelli, L., and Manfrini, L. (2019). Single-shot convolution neural networks for real-time fruit detection within the tree. *Front. Plant Sci.* *10*, 1–12. doi:10.3389/fpls.2019.00611.
- Cupitt, J., and Whelan, B.M. (2001). Determining potential within-field crop management zones. In *ECPA 2001 – Third European Conference on Precision Agriculture* (Montpellier: Ecole Nationale Supérieure Agronomique de Montpellier, France). Vol. 1, p. 7–12.
- Dutilleul, P. (1993). Modifying the t test for assessing the correlation between two spatial processes. *Biometrics* *49*, 305–314. <https://doi.org/10.2307/2532625>.
- Florin, M., McBratney, A.B., and Whelan, B.M. (2008). Inverse meta-modelling of yield monitor data for estimation of soil available water capacities at a spatial resolution of 10 metres across farms. In *Proceedings of the 1<sup>st</sup> Global Workshop on High Resolution Digital Soil Sensing and Mapping* (Sydney, Australia: ACPA), p. 10.
- Gemtos, T., Fountas, S., Tagarakis, A., and Liakos, V. (2013). Precision agriculture application in fruit crops: Experience in handpicked fruits. *Procedia Technol.* *8*, 324–332. <https://doi.org/10.1016/j.protcy.2013.11.043>.
- Greiner, G., and Blackmore, S., eds. (2001). *Proceedings of the Third European Conference on Precision Agriculture*, June 18–20 (France: Montpellier), p. 7–12.
- Käthner, J., Ben-Gal, A., Gebbers, R., Peeters, A., Herppich, W.B., and Zude-Sasse, M. (2017). Evaluating spatially resolved influence of soil and tree water status on quality of European plum grown in semi-humid climate. *Front. Plant Sci.* *8*, 1–10. <https://doi.org/10.3389/fpls.2017.01053>.
- Kitchen, N.R., Drummond, S.T., Lund, E.D., Sudduth, K.A., and Buchleiter, G.W. (2003). Soil electrical conductivity and topography related to yield for three contrasting soil-crop systems. *Agron. J.* *95*, 483–495. <https://doi.org/10.2134/agronj2003.4830>.
- Liakos, V., Tagarakis, A., Fountas, S., Nanos, G.D., Tsiropoulos, Z., and Gemtos, T. (2015). Use of NDVI to predict yield variability in a commercial apple orchard. *Precision Agric.* *15*, 188–197. [https://doi.org/10.3920/978-90-8686-814-8\\_69](https://doi.org/10.3920/978-90-8686-814-8_69).
- Lopez, G., Boini, A., Manfrini, L., Torres-Ruiz, J.M., Pierpaoli, E., Zibordi, M., Losciale, P., Morandi, B., and Corelli-Grappadelli, L. (2018). Effect of shading and water stress on light interception, physiology and yield of apple trees. *Agric. Water Manag.* *210*, 140–148. doi:10.1016/j.agwat.2018.08.015.
- Losciale, P., Manfrini, L., Morandi, B., Pierpaoli, E., Zibordi, M., Stellacci, A.M., Salvati, L. and Corelli Grappadelli, L. (2015). A multivariate approach for assessing leaf photo-assimilation performance using the IPL index. *Physiol. Plant.* *154*, 609–620. doi:10.1111/ppl.12328.
- Manfrini, L., Taylor, J.A., and Corelli Grappadelli, L. (2009). Spatial analysis of the effect of fruit thinning on apple crop load. *Eur. J. Hortic. Sci.* *74*, 54–60.

- Manfrini, L., Pierpaoli, E., Taylor, J.A., Morandi, B., Losciale, P., Zibordi, M., Corelli Grappadelli, L., and Bastías, R.M. (2012). Precision fruit growing: How to collect and interpret data on seasonal variation in apple orchards. *Acta Hort.* 932, 461–470. <https://doi.org/10.17660/ActaHortic.2012.932.67>.
- Manfrini, L.L., Pierpaoli, E.E., Zibordi, M.M., Morandi, B.B., Muzzi, E.E., Losciale, P.P., and Corelli Grappadelli, L.L. (2015). Monitoring strategies for precise production of high quality fruit and yield in apple in Emilia-Romagna. In *Chem. Engin. Transactions* (Vol. 44). <https://doi.org/10.3303/CET1544051>.
- McKenney, D.W., Mackey, B.G., and Zavitz, B.L. (1999). Calibration and sensitivity analysis of a spatially-distributed solar radiation model. *Int. J. Geo. Inf. Sci.* 13, 49–65. <https://doi.org/10.1080/136588199241454>.
- Minasny, B., McBratney, A.B., and Whelan, B.M. (2005). VESPER version 1.62. Australian Centre for Precision Agriculture, University of Sydney, NSW, Australia. <https://doi.org/10.1007/s11119-005-0681-8>.
- Miranda, C., Santesteban, L.G., Urrestarazu, J., Loidi, M., and Royo, J.B. (2018). Sampling stratification using aerial imagery to estimate fruit load in peach tree orchards. *Agriculture* 8(6), 78. <https://doi.org/10.3390/agriculture8060078>.
- Palmer, J.W., Cai, Y.L., and Edjamo, Y. (1991). Effect of part-tree flower thinning on fruiting, vegetative growth and leaf photosynthesis in "Cox's Orange Pippin" apple. *J. Hortic. Sci.* 66(3), 319–325. <https://doi.org/10.1080/00221589.1991.11516159>.
- Paoli, J.N., Strauss, O., Tisseyre, B., Roger, J.M., and Guillaume, S. (2007). Spatial data fusion for qualitative estimation of fuzzy request zones: Application on precision viticulture. *Fuzzy Sets Sys.* 158(5), 535–554. <https://doi.org/10.1016/j.fss.2006.10.019>.
- Peeters, A., Zude, M., Kathner, J., Unlu, M., Kanber, R., Hetzroni, A., Gebbers, R., and Ben-Gal, A. (2015). Getis-Ord's hot- and cold-spot statistics as a basis for multivariate spatial clustering of orchard tree data. *Comput. Electron. Agric.* 111, 140–150. <https://doi.org/10.1016/j.compag.2014.12.011>.
- Rosenberg, M.S., and Anderson, C.D. (2011). PASSaGE: Pattern Analysis, Spatial Statistics and Geographic Exegesis, V. 2. *Methods Ecol. Evol.* 2, 229–232. <https://doi.org/10.1111/j.2041-210X.2010.00081.x>.
- Taylor, J.A., and Whelan, B.M. (2011). Selection of ancillary data to derive production management units in sweetcorn (*Zea Mays* var. *rugosa*) using MANOVA and an information criterion. *Precision Agric.* 12(4), 519–533. <https://doi.org/10.1007/s11119-010-9195-0>.
- Taylor, J.A., Praat, J.P., and Bollen, A.F. (2007a). Spatial variability of kiwifruit quality in orchards and its implications for sampling and mapping. *HortSci.* 42, 246–250. <https://doi.org/10.21273/HORTSCI.42.2.246>.
- Taylor, J.A., McBratney, A.B., and Whelan, B.M. (2007b). Establishing management classes for broadacre agricultural production. *Agron. J.* 99, 1366–1376. <https://doi.org/10.2134/agronj2007.0070>.
- Taylor, J.A., and Bates, T.R. (2013). A discussion on the significance associated with Pearson's correlation in precision agriculture studies. *Precision Agric.* 14, 558. <https://doi.org/10.1007/s11119-013-9314-9>.
- Usha, K., and Singh, B. (2013). Potential applications of remote sensing in horticulture – A review. *Sci. Hortic. (Amsterdam)* 153, 71–83. <https://doi.org/10.1016/j.scienta.2013.01.008>.
- Wünsche, J.N., Palmer, J.W., and Greer, D.H. (2000). Effects of crop load on fruiting and gas-exchange characteristics of 'Braeburn'/M.26 apple trees at full canopy. *J. Am. Soc. Hortic. Sci.* 125, 93–99. <https://doi.org/10.21273/JASHS.125.1.93>.
- Zude-Sasse, M., Fountas, S., Gemtos, T.A., and Abu-Khalaf, N. (2016). Applications of precision agriculture in horticultural crops. *Eur. J. Hortic. Sci.* 81, 78–90. <https://doi.org/10.17660/eJHS.2016/81.2.2>.

Received: Apr. 24, 2018

Accepted: Jul. 26, 2019

Addresses of authors:

Luigi Manfrini<sup>1,\*</sup>, Luca Corelli Grappadelli<sup>1</sup>, Brunella Morandi<sup>1</sup>, Pasquale Losciale<sup>2</sup> and James Arnold Taylor<sup>3</sup><sup>1</sup> Dept. of Agricultural and Food Sciences, University of Bologna, Viale Fanin 46, 40127, Bologna, Italy<sup>2</sup> Department of Soil, Plant and Food Sciences – DiSSPA, University of Bari "Aldo Moro", Via Amendola 165/A, 70126 Bari (Italy)<sup>3</sup> ITAP, INRAE, University of Montpellier, Montpellier, 34000, France

\* Corresponding author; E-mail: luigi.manfrini@unibo.it