Uncertainty appraisal provides useful information for the management of a manual grape harvest

Article

Accepted Version

Highlights

- We developed a model to define sensitive variables in manual grape harvest
- We characterise their uncertainty through Monte Carlo simulations
- Output uncertainty was apportioned onto inputs with a global sensitivity analysis
- Criticalities were identified and analysed through regional sensitivity analysis
- The approach proposed could support decision making in grape harvest management

Uncertainty appraisal provides useful information for the management of a manual grape harvest

Samuele Lo Piano[a)], Alessandro Parenti[b)], Lorenzo Guerrini[b),*]

a) School of the Built Environment, University of Reading, Reading, United Kingdom

b) Department of Agriculture, Food, Environment and Forestry, University of Florence, Florence, Italy

* [corresponding email]

Abstract

This contribution presents a novel approach to characterise uncertainty in the manual grape harvest of a winery in Tuscany (Italy). After identifying the potential sources of variability arising from randomness, weather, and management options, a model to define useful output variables is built. These output variables include the discrepancy in the harvest date of the vineyards (*harvest date discrepancy)*, the discrepancy in the required workforce across harvest dates (labour *discrepancy),* and, finally, the potential deficit of working hours throughout the grape harvest campaign (labour *deficit)*. The range spanned by these variables is first assessed through a Monte Carlo uncertainty analysis wherein the model is repeated approximately 16,000 times with variable combinations of the input parameters per their probability distribution. The assessed uncertainty is then apportioned to the input parameters through a global sensitivity analysis. In turn, a regional sensitivity analysis characterizes the circumstances producing a deficit of working hours, which corresponds to sufferance in the grape harvest campaign. The discussed approach could be implemented in a user-friendly decision-support tool for risk characterisation and efficient grape harvest management.

Keywords

Uncertainty analysis; global sensitivity analysis; risk; timeliness; agricultural management; harvest planning

Nomenclature

Abbreviation	*Description*
De_l	*Labour deficit (h)*
Di_{hd}	*Harvest date discrepancy (day)*
Di_l	*Labour discrepancy (h)*
$hd_{trigger}$	*Ideal harvest date binary trigger*
$h_{trigger}$	*Hours extra binary trigger (h)*
L	*Workers*
$p_{trigger}$	*Productivity binary trigger*
p_n	*Productivity normal distribution (t h^{-1})*
p_u	*Productivity uniform distribution (t h^{-1})*
$r_{trigger}$	*Rain binary trigger*
$rt_{trigger}$	*Rain threshold binary trigger*
SU	*Sundays (day)*
$SU_{trigger}$	*Sundays work*
S_i	*First-order Sobol' indices*
T_i	*Total order indices*
v	*Vineyard*
$y_{trigger}$	*Yield binary trigger*
$y_{n,v}$	*Yield normal distribution (ton)*
$y_{u,v}$	*Yield uniform distribution (ton)*

Introduction

Planning fieldwork days according to the workforce demand profile set by crop features is of the utmost importance for successful agricultural operations. However, the time available may become critical in matching the optimal time window for harvesting the crops, which highlights the importance of estimating the workload required (De Toro & Hansson, 2004; Maton, Bergez, & Leenhardt, 2007) and its entailed costs (Marinello, Yezekyan, Armentano, & Sartori, 2020).

Primary sources of temporal variability include adverse environmental conditions (e.g., due to rain), interactions between climatic events and soil (Obour et al., 2019; De Toro, 2005), and climate change (Kolberg, Persson, Mangerud, & Riley, 2019). Other crop characteristics (e.g., soil slope, the density of planted trees, and other agronomic parameters) may introduce further variability due to the interactions with pedo-climatic conditions on the one hand (Cogato et al., 2020), and the working capacity of machines and field workers on the other (Strub, Kurth, & Loose, 2021; Strub & Loose, 2021). Additional constraints may be set by field non-trafficability, non-working days, and festivities. The ratio between the crop surface on the performed operations, or the mass to be harvested, and the available working time defines the minimal working capacity targeted for the fieldwork. This threshold identifies the deployed minimum machine power or manual work in the case of agricultural mechanization or farm management, respectively (Rotz & Harrigan, 2005).

This research studies grape harvest for winemaking and an agricultural operation involving manual or mechanical means. The selection between these two options can influence wine quality (Guerrini et al., 2018). Mechanical grape harvesters have dramatically advanced in recent years, rendering the quality of a mechanical grape harvest practically indistinguishable from manual harvesting (Parenti et al., 2015) in terms of minor sensory changes (Hendrickson et al., 2016). Additionally, mechanical harvesting offers several advantages in terms of costs, time, management, timeliness, and quick harvesting in adverse conditions (Ferrera et al., 2008, Parenti et al., 2015; Hendrickson et al., 2016). However, manual grape harvest remains popular among consumers, many of whom show a willingness to pay a higher price for wines made from hand-harvested grapes (Dominici, Boncinelli, Gerini, & Marone, 2019). Another advantage of the manual grape harvest is its robustness against seasonal variations, especially because terrain flooding may impede mechanical harvest. Other issues preventing the mechanization of grape harvest include excessive soil slope and the traditional training system (Cogato et al., 2020).

A crucial shortcoming of the manual grape harvest is its poor timeliness, which can significantly impact high-quality wines made from grapes with a narrow optimal harvesting point. A quality loss function has been proposed in the literature to estimate the damage produced when deviating from this optimal harvest date: Ferrer, Mac Cawley, Maturana, Toloza, and Vera (2008) provided an estimate of two days for the optimal time range for the harvest of premium quality grapes. More recently, Varas, Basso, Maturana, Osorio, and Pezoa (2020) proposed an even more conservative estimate of just one day. This picture encounters further complications in the typical settings of winemakers, many of whom own multiple vineyards with scattered features in terms of grape cultivar, soil, training systems, management operations, and density of planted trees. Each of these vineyards has a highly variable optimal harvest date and yield on a yearly basis. This aspect translates into high variability within and across vintages, which affects the required working capacity. Additionally, labour productivity constitutes a highly uncertain parameter due to its influence by different variability sources related to workers and environmental conditions, including the vineyard block slope (Bohle, Maturana, & Vera, 2010).

Hence, effective management of grape harvest operations remains impossible without a careful examination of the large spectrum of potential circumstances at play. This development can be achieved by acknowledging the large variabilities in terms of weather, operations, and randomness, as discussed above. A natural option to capture these settings involves performing numerous Monte Carlo simulations, each of which runs with an individual random combination of the variable factors acknowledged, sampled from their input distributions. Doing so helps characterize the level of uncertainty (uncertainty analysis) entailed by the simulations and, in turn, apportions it to the modelling hypotheses and factors through sensitivity analysis. Finally, the lesson learned can translate into decision-making through the management options per the identified constraints and criticalities.

The next section illustrates this methodology, followed by a discussion of the results and, finally, a presentation of the conclusions on the lesson learned from this study on manual grape harvesting.

Data and Methods

This section describes the collected data and the adopted methodology. The script and data used are available from a *GitHub* repository.

Data and modelling assumptions

The *Pietro Beconcini Agricola* winery, located in a hilly area in San Miniato (43° 41’ 16.1’’ N, 10.52’ 41.9’’ E), 30 km from Florence, Tuscany (Italy), represents the focus of our case study. Data refer to the 2018, 2019, and 2020 harvest campaigns. For each vineyard block (v, 1 - 19), the yearly figures for the grape harvests in terms of grape yields (y, *ton*), productivity per hour of work (p, $t\ h^{-1}$), and harvest date (hd) were recorded. The statistical properties of these variables were garnered under the assumptions of normality or uniformity of their probability distributions. The introduction of these contrasting distribution shapes compensated for the limited number of years from which they were drawn. Additionally, the impact of the distribution-shape assumption on output uncertainty underwent testing using sensitivity analysis through a binary trigger. Normal distributions were truncated at 1.2 standard deviations, which correspond to the largest variation over the years documented across the vineyards in the field. The time frame (t) for the possible harvest dates was defined between the lower threshold of the 229th day of the year (August 16 or, on leap years, August 15) and the upper threshold of the 305th day of the year (October 31 or, on leap years, November 1).

Dividing the sum of the required hours by the available time provided a gross estimate of the required workforce. The workforce that harvests the vineyard is usually hired before the vintage and is used over the whole harvest season with a certain level of flexibility. The payment that the workforce receives for its service is proportional to the working time, covering only the days of harvesting fieldwork.

The viable period to harvest each vineyard falls within one to two days (Ferrer et al., 2008; Varas et al., 2020). The optimal harvest day is also highly variable for each vineyard, as discussed in the introduction. Hence, for the trial-hosting company (Piero Beconcini Agricola), the vineyard's harvest must consist of 19 punctual events, each with its own optimal period. This necessarily results in peaks and troughs in the fieldwork. The number of workers was selected on the basis of the previous year's figures, based on an interview with the winery's owner. Resorting to extra work hours is controlled by a trigger ($h_{trigger}$), as is working on Sundays ($SU_{trigger}$). Sundays (*SU*) are randomly selected across the pool of inquired dates per a 229-235 trigger, which chooses the calendar position of the first Sunday and all subsequent Sundays consistently at a seven-day distance.

The available time for the entire harvest is further limited by the weather. In the CIOSTA method (Reith et al., 2017), this is accounted for by multiplying for a coefficient between 0 and 1, where the coefficient represents the probability of working on a given day per weather conditions. After an interview with the viticulturer, two thresholds were identified (i.e., 5mm and 15mm of rain). These thresholds were modelled into probabilities according to historical weather data. The amount of daily precipitation over the years 2003-2020 was retrieved from the San Miniato weather station located approximately 1 km from the winery. For each of the days investigated, the number of years for which precipitation exceeded the 5 mm or 15 mm threshold, divided by the total number of recorded occurrences, was recorded. This fraction resulted in the probability of precipitation above this given threshold. Triggers selected whether, in a given simulation, it would rain on a particular day ($r_{trigger}$) (where the random number extracted is higher than the probability of having 5 or 15 mm of rain on that day) and the threshold selected ($rt_{trigger}$) (i.e., 5 mm vs. 15 mm of rain). The latter trigger defined the risk propensity of the winery manager in terms of the required amount of precipitation to call off a harvest day (binary: 5 mm vs. 15 mm).

Table 1 reports the probability distributions for the uncertain parameters. For parameters with more than one distribution shape available, the triggers activate either one or the other based on the extracted random binary value in a specific simulation.

Table 1 Summary of the parameters and their distribution. D stands for discrete, U for uniform, N for normal, and DU for discrete uniform. The statistical moments for yield, productivity, and ideal harvest date are reported in Table 2 for clarity.

Parameter	Description	Distribution
$y_{n,v}\ (t)$	Yield normal	$\mathrm{N}(mean_w, std_w)$
$y_{u,v}\ (t)$	Yield uniform	$\mathrm{U}(min_w, max_w)$
$p_n\ (t\ h^{-1})$	Productivity normal	$\mathrm{N}(mean, std)$
$p_u\ (t\ h^{-1})$	Productivity uniform	$\mathrm{U}(min, max)$
$hd_{n,v}(d)$	Ideal harvest date normal	$\mathrm{N}(mean_w, std_w)$
$hd_{u,v}(d)$	Ideal harvest date uniform	$\mathrm{U}(min_w, max_w)$
$y_{trigger}$	Yield binary trigger	$\mathrm{DU}(0,1)$
$p_{trigger}$	Productivity binary trigger	$\mathrm{DU}(0,1)$
$hd_{trigger}$	Ideal harvest date binary trigger	$\mathrm{DU}(0,1)$
l	Workers	$\mathrm{DU}(8,17)$
$h_{trigger}(h)$	Hours extra binary trigger	$\mathrm{DU}(8,10)$
$SU(day)$	Sundays	$\mathrm{DU}(229,235)$
$SU_{trigger}$	Sundays work	$\mathrm{DU}(0,1)$
$r_{trigger}$	Rain binary trigger	$\mathrm{DU}(0,1)$
$rt_{trigger}$	Rain threshold binary trigger	$\mathrm{DU}(0,1)$

Table 2 reports the data collected for the individual vineyards. These data were gathered straightforwardly, yet they are an effective proxy to represent the winery arrangements in terms of fieldwork organization. These features make the variables crucial for simulating and modelling a manual grape harvest functional to high-quality winemaking and the characterization of the variability of this stage. The vineyard's average harvest dates vary from day 251 to day 285 (an interval of 35 days). Vineyard blocks three and 11 are expected to be harvested on average on day 251, whereas the others in roughly ten more days. Thus, all the remaining vineyards should be harvested within 25 days (specifically, four vineyards from day 260 to day 262 and nine vineyards from day 267 to day 271). Hence, 13 out of 19 vineyards will be harvested in a range of 11 days, which represents the vintage's critical phase. The number of working hours required to harvest the vineyards ranged from 17±5 h needed for vineyard block 17 to the 110±35 h required for vineyard block 12. The yields vary highly across vineyards, and where they depend primarily on the viticulturer's agronomical choices and oenological targets.

Table 2 Recorded data for vineyard blocks showing surface, cultivar, ideal harvest date, work hours, and yields (mean ± standard deviation).

Vineyard	Surface (m^2)	Cultivar	Harvest date	Work hours (h)	Yield (kg)
1	3000	Tempranillo	262±11	37±14	3445±78
2	6000	Colorino	270±12	73±7	4025±177
3	2500	Tempranillo	251±11	43±14	2825±106
4	3000	Merlot	268±24	69±11	3575±177
5	7000	Sangiovese	278±11	49±21	5100±424
6	4000	Trebbiano	271±16	47±10	3350±212
7	8000	Petit manseng	260±15	81±12	4250±71
8	6000	Malvasia, Bianca Lunga	267±13	53±16	6350±636
9	4000	Tempranillo	273±16	28±7	3550±71
10	2500	Merlot	277±21	28±11	1400±141
11	3500	Sangiovese	251±16	45±12	3600±283
12	13000	Sangiovese	260±14	110±35	4650±71
13	9000	Tempranillo	261±4	53±11	5855±346
14	6000	Tempranillo	267±11	73±26	5000±283
15	3000	Sangiovese	285±16	41±8	1000±283
16	10000	Sangiovese	270±11	77±25	4350±212
17	4000	Sangiovese	271±12	17±5	950±71
18	5000	Petit manseng	268±12	35±18	3375±247
19	14000	Sangiovese	268±4	81±6	3250±354

Model output variables

Each simulation in this research produced three output variables: *harvest date discrepancy*, the labour *discrepancy,* and the labour *deficit.*

The output variable *harvest date discrepancy* (D_{hd}, *d*) is the average time lag in harvest dates across the vineyards in the winery, as shown in Eq. 1. In this equation, $1 \le n \le v$ represents the number of harvest date occurrences, and *k* is the index running over the population of *n*. The theoretical minimum for *n* is 1, which corresponds to the unlikely event all vineyards mature on the same day. Conversely, the theoretical maximum for *n* is 19, which corresponds to the case where no harvest date across the winery coincides. Ideal circumstances would correspond to a low figure for D_{hd}

(approximately 1), with distinct harvest dates that are temporally close enough to allow for the optimal use of the workforce available.

$$Di_{hd} = \frac{\sum_{k=1}^{k=n-1}(hd_{k+1} - hd_k)}{n-1} \quad (1)$$

The even allocation of the available workforce across the days is a crucial variable. To this end, one can define the labour discrepancy (Di_l, days) per Eq. 2. This output variable expresses the degree to which the labour requirements are scattered across the vineyards over their corresponding harvest dates. The lower the variable, the more regular the pattern of labour requirements.

$$Di_l = \frac{\sum_{k=1}^{k=n} |(\sum_v(\frac{y_v}{p}))_k - \frac{\sum_{k=1}^{k=n} |(\sum_v(\frac{y_v}{p}))_k}{n}|}{n-1} \quad (2)$$

Finally, the labour *deficit* (De_l, h) defined in Eq. 3 expresses the difference between the labour requirement from the field and the availability of labour, which depends on the following factors: number of workers hired, the decision to resort to extra daily hours, work shifts on Sundays, and the harvesting days after accounting for the working days lost due to rain. Values around zero correspond to adequate workforce availability, while negative values represent a situation of excess. Additionally, positive values highlight situations of deficiency whose criticality increases with the proxy value. This variable is the most crucial in determining the impact of managerial decisions on the harvest outcome.

$$De_l = \sum_{k=1}^{k=n}(\sum_v(\frac{y_v}{p}) - l * h)_k \quad \text{Eq. (3)}$$

Uncertainty analysis was run on the output variables. This analysis comprised 16,383 (2^{14} - 1) Monte Carlo simulations performed on samples drawn from the parameters' probability distributions per the specifics detailed in Table 1. The low-discrepancy Sobol' sequence of quasi-random numbers (Bratley & Fox, 1988) drew these samples, the rationale being that it converges faster than pure random numbers. The quasi-random numbers were transformed into instances of the sample probability distributions through an inverse transformation of the cumulative probability figures for each of the input variables.

Finally, the variance of the output variable De_l was apportioned to the input parameters through global sensitivity analysis (Saltelli, Ratto, et al., 2008), whereby all the parameters were varied simultaneously within their uncertainty range. Doing so made it possible to fully characterize the output variability by including the part caused by interactions among parameters. This is the typical case of non-additive models (i.e., those in which mathematical relations among the uncertain parameters are beyond mere additions and subtractions). This paper discusses two metrics of sensitivity: firstly, the first-order Sobol' indices S_i, which estimate the contribution to the variance of individual parameters (Sobol', 1993); and secondly, the total-order indices T_i, which also quantify the contribution of the parameters through their interaction with other parameters (Homma and Saltelli, 1996). These indices are included in the range (*0, 1*) for independent input parameters and represent a convenient way to communicate the importance of the input parameters' contribution to the output uncertainty. For a given input parameter *i*, it is always valid that $S_i \leq T_i$. The Saltelli and Jansen estimators for *Si* and T_i were used, respectively (Saltelli, Annoni, et al., 2010). Additionally, 1,000 bootstrap replicas of the Monte Carlo simulations with replacements were generated to strengthen the estimations of the sensitivity indices. Finally, *regional sensitivity analysis* (Saltelli, Ratto, et al., 2008) was adopted to understand the range of input parameter values responsible for a given output range (e.g., in the case of De_l) and in terms of the direction of change (Deza and Deza, 2013).

Results and Discussion

Figure 1 shows plots for the output variable *harvest date discrepancy* (Di_{hd}). Most of the values cluster around an average discrepancy of 2.5 days (Table 3). Approximately 98% of the simulations produced $2 < Di_{hd} < 3$, while only six simulations resulted in an output larger than four days; this is a negligible figure over the full pool of approximately 16,000 simulations. The largest value of seven days was obtained from a simulation in which vineyards harvest days clustered over five days (specifically, days

276, 284, 291, 301, and 304). The simulations producing 2 < Di_{hd} < 3 resulted from an average of 15 harvest dates over the 19 vineyards, with 267 as the average harvest date and an average spacing of 1.2 days across the vineyards.

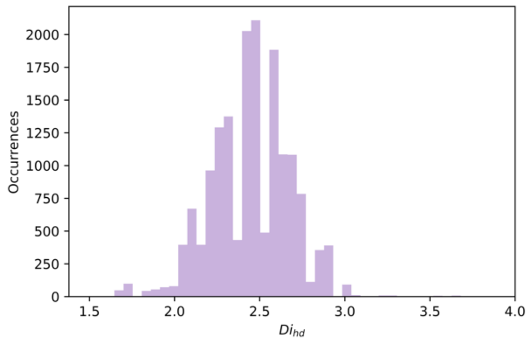

Fig. 1 Number of occurrences of Di_{hd} (d) in a pool of simulations (truncated at four d).

Table 3 Statistical properties of output distribution.

	Di_{hd}(d)	Di_l(h)	De_l (h)
mean	2.45	34	~ −330
std	0.24	11	~ 520
min	1.65	16	~ −2,200
25%	2.29	27	~ −680
50%	2.47	31	~ −330
75%	2.58	38	2.0
max	7.00	~350	~ 1,700

The observed trends can be understood by capturing the seasonal variability in the different vineyards, as showcased in Fig. 2. In particular, this figure illustrates the population of harvest dates obtained from Monte Carlo simulations on the basis of the assumed distribution shapes (i.e., normal and uniform) for the experimental data. Vineyards characterized by a narrower date range around the harvest date are less sensitive to the specifics of the vintage of a particular year than vineyards with a more extended date range. The ideal managerial situation would correspond to vineyards with a narrow distribution of harvest dates and harvest dates that are poorly spread within each vintage, lacking overlaps across vineyards. To this end, one could resort to canopy management techniques to delay or anticipate the harvest day for the vineyards on the most critical days. The same result could be achieved by planting a widely scattered pool of grape varieties and cultivars expected to mature on less critical dates during vintages.

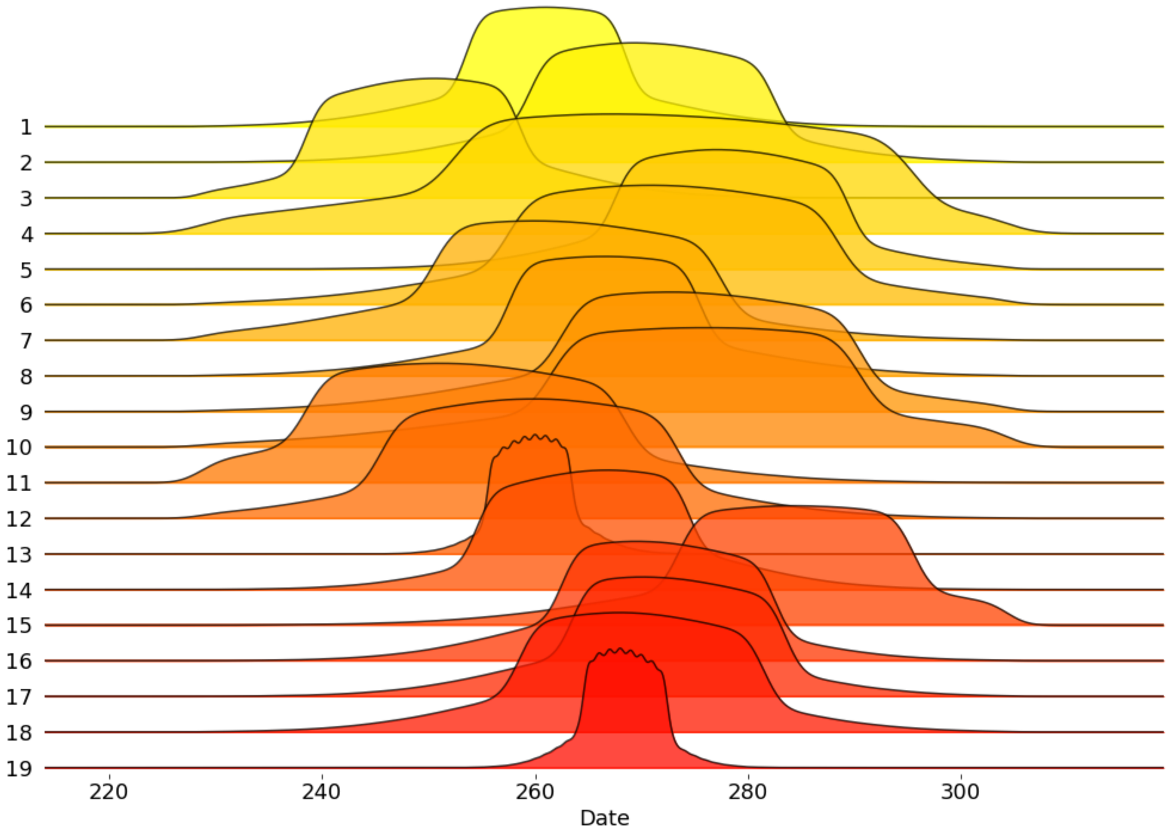

Fig. 2 Visualisation of the distribution of the ideal harvest dates across vineyards as overlayed sampled distributions.

The output variable *harvest date discrepancy* (Di_l) shows some level of right skewness, as illustrated in Fig. 3. In total, nine simulations produce a discrepancy larger than 100 h, with the maximum at approximately 350 h (Table 3). The latter value was obtained from the same simulation, producing the largest Di_{hd}, for which daily labour requirements on the harvest days were as follows: 60, 81, 54, 43, and 929 h. The outlier on day 304 (929 h) was responsible for this Di_{hd} figure. This variable can be regarded as a proxy that summarises the flexibility in the required workforce for the harvest. Given that the hiring of workers needs to occur before the harvest when the ideal harvest date is not known, this figure can inform the level of flexibility when contracting the workforce. Additionally, the company manager can also minimize the variability of this output variable with appropriate agronomical choices. These choices include calibrating the extension of the fields allocated to a variety of grapes typically scattered and having harvest dates well-spaced in time in seeking a constant workforce demand across harvest dates.

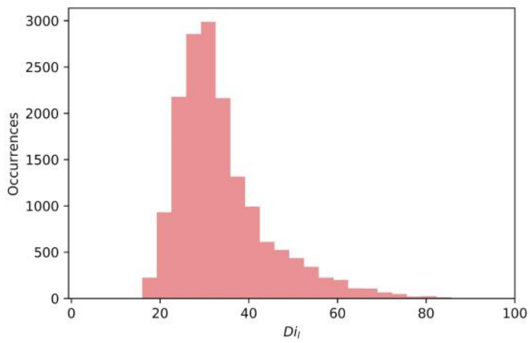

Fig. 3 Number of occurrences of Di_l (h) in the pool of simulations run truncated at 100 h.

Finally, Fig. 4 shows the distribution of the output variable labour *deficit* (De_l). Three-quarters of the simulations produce negative figures, which reflect situations in which a sufficient workforce has been employed for the harvest. The maximum deficit of more than 1,600 h (Table 3) was produced in correspondence with profound mismatches between the harvest dates and the days on which the workforce was available. This situation occurred in simulation 15,684, for which the vineyards' requirements over 13 different harvest dates were satisfied only once. The causes of this pattern will be discussed later. All the simulations with large deficits of hours will most likely force the company to harvest on less preferable dates, eventually negatively impacting the quality of the wine produced.

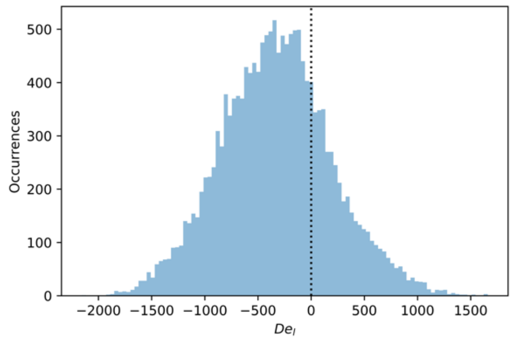

Fig. 4 Number of occurrences of De_l (h) in the pool of simulations.

A global sensitivity analysis identified the parameters primarily responsible for the observed trend. Only the sensitivity analysis on the output variable De_l is shown here because it is the output for which the input parameters potentially affect variability (Fig. 5). The narrow whisker-box plots for the bootstrapped samples confirm the stability of the estimates. The number of workers was the most influential parameter (S_l = 0:29 -0:31) followed by the selection of rainy days (S_r = 0:28 - 0:30). Neither the triggers selecting the distribution shapes nor the individual yields showed any significant impact on the output ($S_{ytrigger} \approx 0$), most likely attributable to their low standard deviation (approximately 5% of the mean) across simulations. Conversely, the triggers for the harvest date and productivity distribution had an effect through interaction with other parameters ($S_{hd,trigger} \approx 0$, $T_{hd,trigger} > 0$; $S_{p,trigger} \approx 0$, $T_{p,trigger} > 0$). The sum of all first-order terms only explained 75-86% of the output variance, which means that the reminder occurred through interactions between pairs or larger groups of parameters.

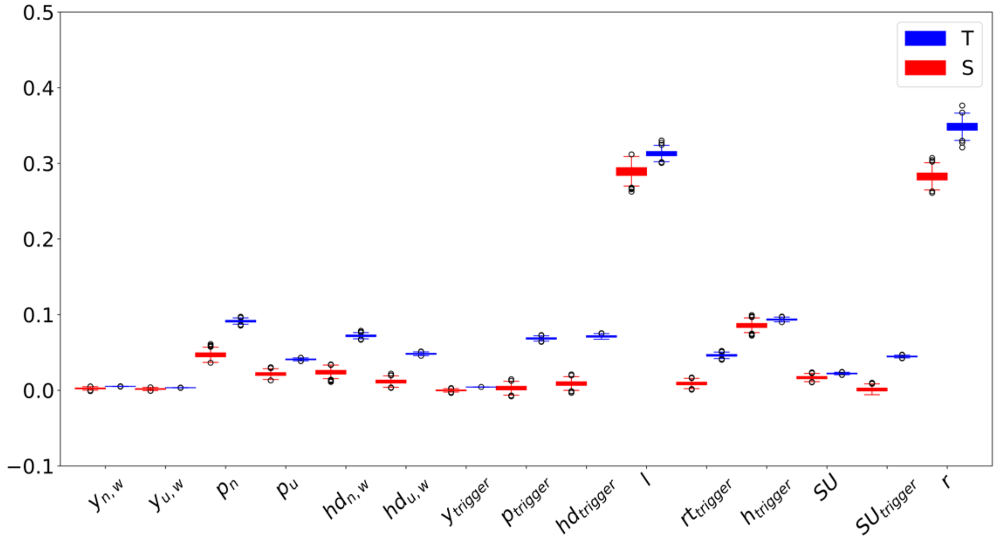

Fig. 5 Sensitivity indices for uncertain input variables for De_l in a pool of simulations. The whisker box plots produced over 1,000 bootstrap replicas with replacements.

Managing the manual grape harvest of the different vineyards was confronted with four crucial sources of variability, for which the scope for management control is limited when the harvest date approaches. These sources of variability were the ideal harvest date, the vineyard yield, the workers' productivity, and the weather. Hence, quantifying the contribution of these sources of variability to the final uncertainty of the output variables represents a method of understanding to what extent management choices may influence grape harvest criticalities. To do so, the input parameters were clustered into three groups: firstly, natural (which includes the natural variability, including yields, harvest days, and rainy days); secondly, modelling assumptions and exogenous variables (triggers for the distribution shapes adopted, productivity, and the trigger for Sundays); and thirdly, managerial variables (i.e., number of workers, triggers for resorting to extra hours and Sunday working shifts, and the rain threshold). The first group represents the variables minimally controllable by the winery management when the harvest date is approaching, the second represents those related to modelling and other assumptions, and the third captures the effect of management choices. Figure 6 shows the results for the output variable *De_l*.

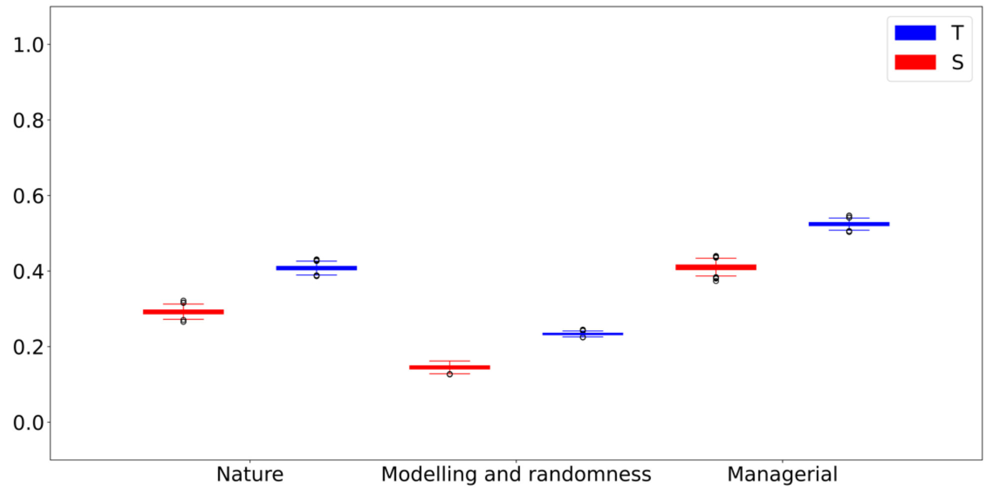

Fig. 6 Sensitivity indices for grouped uncertain input variables for De_l in a pool of simulations. The whisker box plots produced over 1,000 bootstrap replicas with replacements.

The managerial choices available are the variables that have the most substantial effect on output uncertainty, although they can only justify 37-44% of the output variance (i.e., output variance would reduce on average by this fraction by fixing this group of parameters). $T_{managerial}$ amounts to 50-55% when one acknowledges the interactions with the other group of variables. The sum of first-order effects $S_{group,i}$ apportions again only 85-90% of the output variance, showing that one still has interactions across the three parameter groups. This situation corresponds to irreducible output uncertainty, which is not uncommon when modelling the interface between human and natural systems (see, for example, Lo Piano, Saltelli, & van der Sluijs, (2019); Puy, Lo Piano, & Saltelli, (2020)).

Let us now address the question of which values of the uncertainty input variables will more likely lead to workforce deficits during the harvest through a regional sensitivity analysis. This discussion focuses on the two most crucial variables under management control - namely, l and $h_{trigger}$ - as the choice of an adequate number of workers is probably the most critical decision in managing the grape harvest. This choice should simultaneously rule out the risks of overinvestments (i.e., too many workers) and the risk of harvest dates mismatching the ideal situation due to employing too few workers (Allen & Schuster, 2004).

Figure 7 analyses the D_{el} distribution against these managerial choices.

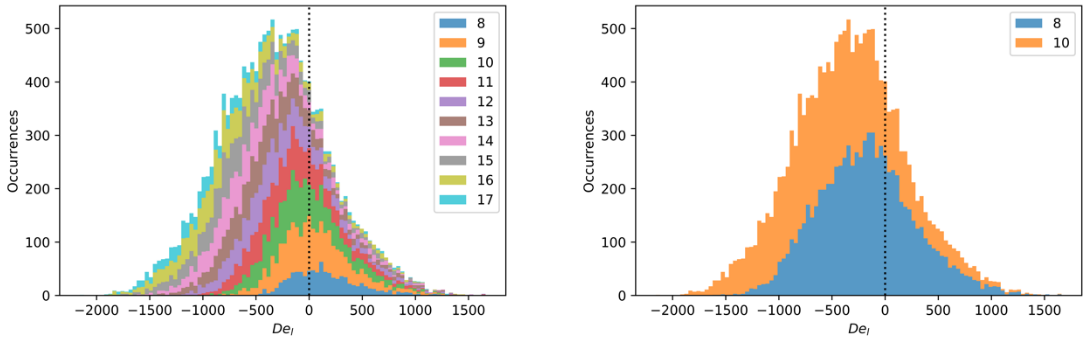

Fig. 7 Number of occurrences of De_l in a pool of simulations for a) the number of workers and b) the number of hours due to working extra hours (daily shifts of 10 h against 8 h).

Figure 7 shows that none of the possible combinations explored can rule out the risk of incurring deficit hours. Contracting more workers leads to less deficient simulations, yet one can result in a deficit even when hiring the maximum number of workers (i.e., 17 workers). This situation occurred in 84 simulations out of 910 with 17 workers (i.e., almost 10%). Considering the number of available working days for this setting against the whole pool of simulations (Table 4), one can understand the causes of this finding.

Table 4 Number of available working days in simulations with 17 workers leading to a deficit in De_l compared to those in the whole pool of simulations

	Working days (17 workers and deficit)	Working days (whole pool)
mean	27	61
std	14	14
min	11	11
25%	14	65
50%	26	66
75%	30	66
max	66	77

The deficit simulations suffer from a limited number of days of workforce availability (mean = 27), but a deficit harvest may be produced for a large number of working days (up to 66). This circumstance occurs from a simulation in which a substantial number of vineyards mature on days where the workforce is not available due to rainy days or non-working Sundays.

Even jointly resorting to the maximum number of workers (i.e., 17) and a daily extra-hour work shift produced 30 simulations with a deficit of hour work out of a total of 454 (6%). On the other hand, the choice of hiring an abundance of workers to limit the risk of deficit situations is not free from drawbacks because a significant excess of the workforce may lead to extra costs for the management and demotivate the engaged workforce due to the inadequateness of the tasks assigned. Notably, the same situation may also result under the condition of a large deficit in terms of the workforce.

Conclusions

This contribution presented a model that identifies features related to risk in a grape harvest campaign. The proposed variables *harvest date discrepancy*, labour *discrepancy,* and labour *deficit* proved effective proxies that can be calculated straightforwardly from limited and easily accessible data. Additionally, our contribution illustrated the usefulness of Monte Carlo-based uncertainty analysis and sensitivity analysis in estimating and characterising the main sources of risk in a grape harvest campaign.

The proposed approach can be escalated and replicated in other wineries to inform managers about the available options for mitigating potentially critical situations. Uncertainty analysis can help quantify the extent of these critical issues by evaluating a large number of potential combinations of input factors, where their specific impact on the output uncertainty can eventually be apportioned through global sensitivity analysis. Another valuable piece of information is the amount of variability under the control of the viticulturer through their management choices, which can eventually lead to sound estimations of the costs and the level of risk one wishes to embrace. In our case study, only around 40% of the variance of the labour deficit depended on parameters under the control of the management.

The model and the approach elaborated in the present study may fruitfully serve as the backbone of a user-friendly decision-support tool that can help winemakers readily explore a set of assumptions and produce inferences about the consequence of their management choices. The approach could be further refined by including monetary proxies and penalty functions dependent on the temporal mismatch between the actual and ideal harvest dates for the vineyards blocks.

Acknowledgments

We thank the Beconcini winery and Lorenzo Cecchini for their collaboration in developing this manuscript. We also thank the two anonymous reviewers for their helpful comments. The responsibility for any error remains solely with us. This study did not receive any public or private funding.

References

Allen, S. J., & Schuster, E. W. (2004). Controlling the Risk for an Agricultural Harvest. Manufacturing & Service Operation Management, 6(3), 225–236. https://doi.org/10.1287/msom.1040.0035

Bohle, C., Maturana, S., & Vera, J. (2010). A robust optimization approach to wine grape harvesting scheduling. European Journal of Operational Research, 200(1), 245–252. https://doi.org/10.1016/j.ejor.2008.12.003

Bratley, P., & Fox, B. L. (1988). Algorithm 659: Implementing Sobol’s quasi-random sequence generator. ACM Transactions on Mathematical Software, 14(1), 88–100. https://doi.org/10.1145/42288.214372

C. A. Rotz, & T. M. Harrigan. (2005). Predicting Suitable Days for Field Machinery Operations in a Whole Farm Simulation. Applied Engineering in Agriculture, 21(4), 563–571. https://doi.org/10.13031/2013.18563

Cogato, A., Pezzuolo, A., Sørensen, C. G., De Bei, R., Sozzi, M., & Marinello, F. (2020). A GIS-based multicriteria index to evaluate the mechanisability potential of Italian vineyard area. Land, 9(11), 1–17. https://doi.org/10.3390/land9110469

Cogato, A., Pezzuolo, A., Sozzi, M., & Marinello, F. (2020). A sample of Italian vineyards: Landscape and management parameters dataset. Data in brief, 33, 106589. htpps://doi.org/10.1016/j.dib.2020.106589

De Toro, A. (2005). Influences on timeliness costs and their variability on arable farms. Biosystems Engineering, 92(1), 1–13. https://doi.org/10.1016/j.biosystemseng.2005.06.007

De Toro, A., & Hansson, P. A. (2004). Analysis of field machinery performance based on daily soil workability status using discrete event simulation or on average workday probability. Agricultural Systems, 79(1), 109–129. https://doi.org/10.1016/S0308-521X(03)00073-8

Deza, M. M., & Deza, E. (2014). Distances in probability theory. In Encyclopedia of Distances (pp. 257-272). Springer, Berlin, Heidelberg.

Dominici, A., Boncinelli, F., Gerini, F., & Marone, E. (2019). Consumer preference for wine from hand-harvested grapes. British Food Journal, 122(8), 2551–2567. https://doi.org/10.1108/BFJ-04-2019-0301

Ferrer, J.-C., Mac Cawley, A., Maturana, S., Toloza, S., & Vera, J. (2008). An optimization approach for scheduling wine grape harvest operations. International Journal of Production Economics, 112(2), 985–999. https://doi.org/10.1016/j.ijpe.2007.05.020

Guerrini, L., Masella, P., Angeloni, G., Calamai, L., Spinelli, S., Di Blasi, S., & Parenti, A. (2018). Harvest of Sangiovese grapes: The influence of material other than grape and unripe berries on wine quality. European Food Research and Technology. https://doi.org/10.1007/s00217-018-3063-y

Kolberg, D., Persson, T., Mangerud, K., & Riley, H. (2019). Impact of projected climate change on workability, attainable yield, profitability and farm mechanization in Norwegian spring cereals. Soil and Tillage Research, 185(1432), 122–138. https://doi.org/10.1016/j.still.2018.09.002

Hendrickson, D. A., Lerno, L. A., Hjelmeland, A. K., Ebeler, S. E., Heymann, H., Hopfer, H., Block, K. L., Brenneman, C. A., Oberholster A. (2016). Impact of Mechanical Harvesting and Optical Berry Sorting on Grape and Wine Composition. American Journal of Oenology and Viticulture, 67, 385-397. https://doi.org/10.5344/ajev.2016.14132

Lo Piano, S., Saltelli, A., & van der Sluijs, J. P. (2019). Silver as a Constraint for a Large-Scale Development of Solar Photovoltaics? Scenario-Making to the Year 2050 Supported by Expert Engagement and Global Sensitivity Analysis. Frontiers in Energy Research, 7. https://doi.org/10.3389/fenrg.2019.00056

Marinello, F., Yezekyan, T., Armentano, G., & Sartori, L. (2020). Modelling of Agricultural Machinery Trends for Power, Mass, Working Width and Price. In Lecture Notes in Civil Engineering (Vol. 67). Springer International Publishing. https://doi.org/10.1007/978-3-030-39299-4_54

Maton, L., Bergez, J. E., & Leenhardt, D. (2007). Modelling the days which are agronomically suitable for sowing maize. European Journal of Agronomy, 27(1), 123–129. https://doi.org/10.1016/j.eja.2007.02.007

Obour, P. B., Keller, T., Jensen, J. L., Edwards, G., Lamandé, M., Watts, C. W., … Munkholm, L. J. (2019). Soil water contents for tillage: A comparison of approaches and consequences for the number of workable days. Soil and Tillage Research, 195(January), 104384. https://doi.org/10.1016/j.still.2019.104384

Parenti, A., Spugnoli, P., Masella, P., Guerrini, L., Benedettelli, S., & Di Blasi, S. (2015). Comparison of grape harvesting and sorting methods on factors affecting the must quality. Journal of Agricultural Engineering, 46(1). https://doi.org/10.4081/jae.2015.456

Puy, A., Lo Piano, S., & Saltelli, A. (2020). Current Models Underestimate Future Irrigated Areas. Geophysical Research Letters, 47(8), e2020GL087360. https://doi.org/10.1029/2020GL087360

Reith, S., Frisch, J. and Winkler, B. (2017) 'Revision of the working time classification to optimize work processes in modern agriculture', Chemical Engineering Transactions, 58(2006), pp. 121–126. doi: 10.3303/CET1758021.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., ... & Tarantola, S. (2008). Global sensitivity analysis: the primer. John Wiley & Sons.

Strub, L., Kurth, A., & Loose, S. M. (2021). Effects of viticultural mechanization on working time requirements and production costs. American Journal of Enology and Viticulture, 72(1), 46–55. https://doi.org/10.5344/ajev.2020.20027

Strub, L., & Loose, S. M. (2021). The cost disadvantage of steep slope viticulture and strategies for its preservation. Oeno One, 55(1), 49–68. https://doi.org/10.20870/OENO-ONE.2021.55.1.4494

Varas, M., Basso, F., Maturana, S., Osorio, D., & Pezoa, R. (2020). A multi-objective approach for supporting wine grape harvest operations. Computers and Industrial Engineering, 145(May 2019), 106497. https://doi.org/10.1016/j.cie.2020.106497

Figures caption

Fig. 1 Number of occurrences of Di_{hd} (d) in the pool of simulations run truncated at 4 d.

Fig. 2 Visualisation of distribution of the ideal harvest dates across vineyards as overlayed sampled distributions.

Fig. 3 Number of occurrences of Di_l (h) in the pool of simulations run truncated at 100 h.

Fig. 4 Number of occurrences of De_l (h) in the pool of simulations.

Fig. 5 Sensitivity indices for the uncertain input variables for De_l in the pool of simulations. The whisker box plots have been produced over 1,000 bootstrap replicas with replacement.

Fig. 6 Sensitivity indices for the grouped uncertain input variables for De_l in the pool of simulations. The whisker box plots have been produced over 1,000 bootstrap replicas with replacement.

Fig. 7 Number of occurrences of De_l in the pool of simulations for a variable a) number of workers; b) number of hours due to working extra hours (daily shifts of 10 against 8 h.).

Tables caption

Table 1 Summary of the parameters and their distribution. D stands for discrete, U for uniform, N for normal and DU for discrete uniform. The statistical moments for yield, productivity and ideal harvest date are reported in Table 2 for clarity.

Table 2 Recorded vineyard blocks data: surface, cultivar, ideal harvest date, work hours and yields (mean ± standard deviation).

Table 3 Statistical properties of the output distributions.

Table 1 Number of available working days in simulations with seventeen workers leading to a deficit in De_l against those in the whole pool of simulations.