Project Number: 774571
Start Date of Project: 2017/11/01
Duration: 48 months

Type of document 5.5 – V1.0

Fruit development and production monitoring

Dissemination level	PU
Submission Date	2021-08-31
Work Package	WP5
Task	T 5.6
Type	Report
Version	1.0
Author	Valerio Cristofori, Cristian Silvestri, Marco Paolocci, Emanuele Garone, Nico Bono Rossello, Martina Lippi
Approved by	Andrea Gasparri + PMC

Executive Summary

This document treats the agronomic approach at field-scale to allow the development of an automated protocol for fruit detection in hazelnut plants and orchards. The model developed, to be applied in large-scale hazelnut orchards, will allow technicians, growers, and traders to predict the expected yearly yield during the growing season, when the nuts in the clusters are fully developed and when the in-shell nuts and clusters have completed their early drop-out due to lack of fruit set or caused by pathological injuries. The fruit detection model could also be a useful tool to estimate the hazelnut losses due to pests and diseases or adverse climatic conditions (strong wind or hailstorms) in early developing stages of the nuts.

The following sub-tasks have been analysed:

1. Analysis of hazelnut world production consistency, mainly focused on suited areas for hazelnut cultivation and analysis of “alternate bearing” effects;
2. Description of the hazelnut phenology, growth and fruit development to identify the best moment for nuts detection and estimation at field-scale;
3. Trials description for developing the proper procedure to detect nuts, clusters and yield at field-scale;
4. Nut and kernel traits and nut and kernel defects evaluation to clean up the yield estimation from the portion of defective nuts;
5. Development of mathematical model for early detection of clusters, nuts and for orchard yield estimation.

Table of Content

Abbreviations and Acronyms

PA	Precision Agriculture
3D	Three-Dimensional
UGV	Unmanned Ground Vehicle
UAV	Unmanned Aerial Vehicle
Yo	Young Tree
Ad	Adult Tree
ID	Identification Code
cv	Cultivar
F_n	Fruit Detection Sample
TCSA	Trunk-Cross Sectional Area
YE	Yield Efficiency
LTV	Linear Time-Variant
SD	Standard Deviation
SSE	Sum of Squares
RMSE	Root Mean Square Error

1 World hazelnut production consistency

1.1 World hazelnut cultivation and production

Hazelnut is a perennial temperate nut crop spread from Anatolian peninsula and Caucasus region to Europe and Nord Africa, and their wilds are distributed between 40th and 45th North latitude. This latitudinal range also represents where the most important hazelnut production districts in the world are located [1].

The most relevant countries for hazelnut cultivation include Turkey, Italy, USA, Georgia, Azerbaijan, and Spain and, increasingly, the hazelnut cultivation is also moving in other countries such as Chile, South Africa, China, and Australia [2].

The world hazelnut production was about 512,000 tons of shelled nuts (kernel) in 2020 - 2021; more than 60% of the total production was reached in Turkey (Figure 1), followed by Italy which reached 75,000 tons of shelled nuts [3]. Similar trends were also recorded in the USA and Chile, where the hazelnut production have risen considerably from the previous season. Anyway, due to the biennial bearing aptitude to the species, physiological disorders accentuated when plants are aged and mismanaged, the year-to-year production varies significantly [4]. Furthermore, the world hazelnut cultivation areas are showing nowadays a marked dynamism: for instance, in Italy the increasing interest for this nut crop is promoting its cultivation also in non-traditional or unsuited areas.

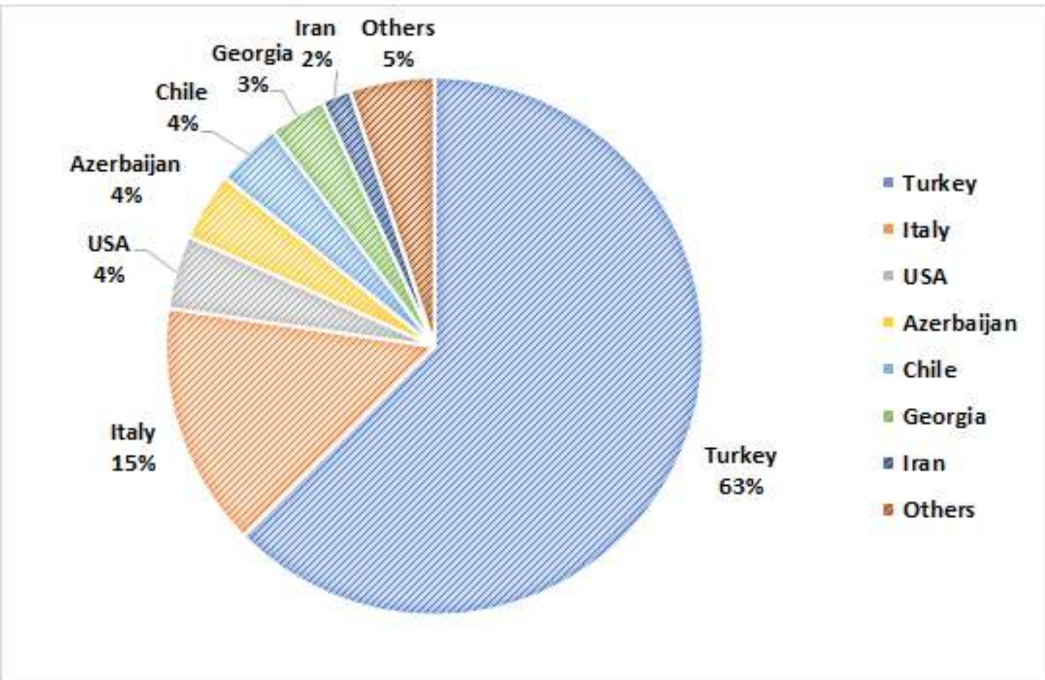

Figure 1 - World hazelnut production, expressed as percentage of the total annual production of shelled nuts (kernel). Growing season 2020 - 2021.

1.2 Trend of hazelnut cultivation and production in Italian suited areas

A “case of study” is briefly reported in the following for better explain the increase in hazelnut cultivation worldwide and its relations with the long unproductive period during the plant juvenility and its biennial bearing aptitude to the mature orchards. The case of study focuses on the hazelnut consistency in the most important Italian provinces for hazelnut cultivation: Cuneo (Piedmont) and Viterbo (Latium), respectively (Figure 2).

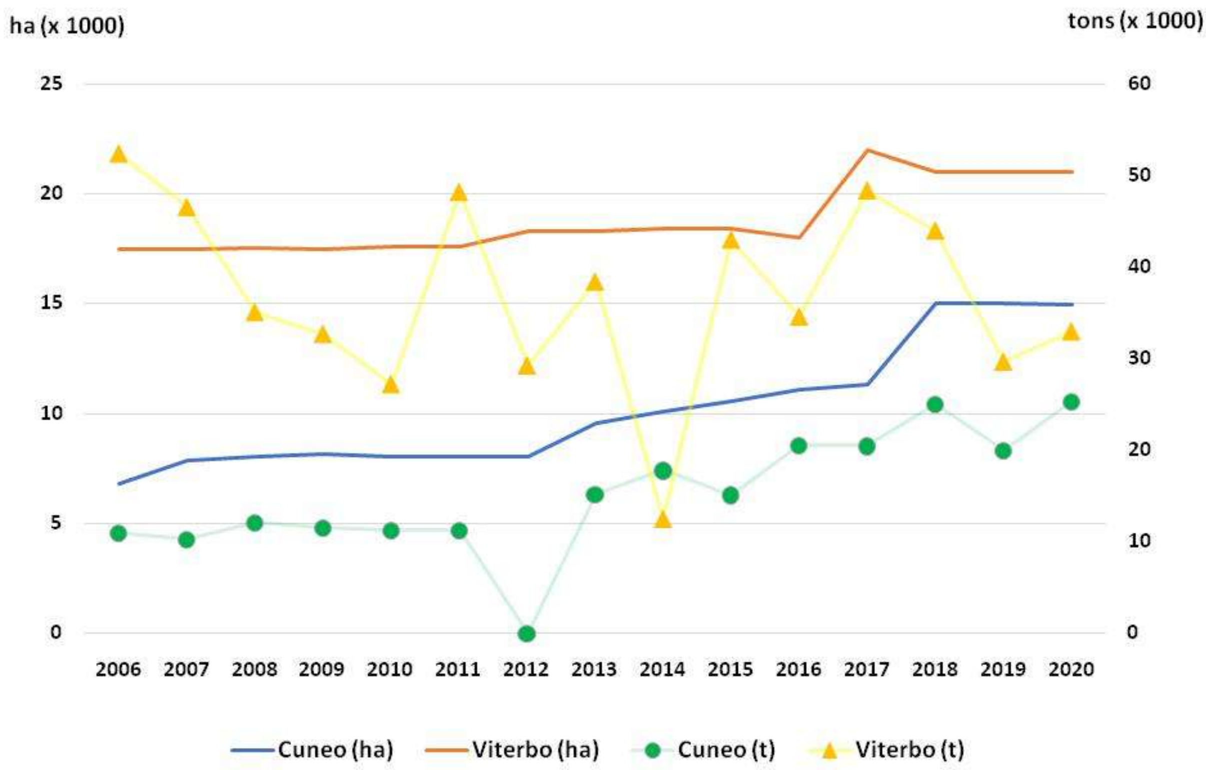

Figure 2 – Trends in total production (in-shell nuts) and hazelnut cultivated areas in Cuneo and Viterbo provinces, the most important productive districts of Italy. Tons of produced nuts (in green for Cuneo and in yellow for Viterbo) as well as considered hectares (in blue for Cuneo and in red for Viterbo) are reported.

More than the 50% of the total hazelnut cultivation of the Piedmont region is located in the province of Cuneo. For this area, a huge growth of hazelnut cultivation has been recorded in the last ten years, with an increase of +87% of hectares devoted to this crop.

Similarly, a recent spread of hazelnut cultivation in Latium region has been observed and, currently the surface dedicated to this crop is almost 24,000 hectares. However, the increase in hazelnut cultivated areas does not always match with a steady increase in total yield. This “paradox” is triggered by many factors including the effect of climate change: despite the presence of abundant blooming during the plant winter rest and high levels of fruit-set, often the nuts production is limited with losses of 50-60% of in-shell nuts per plant at the end of growing season.

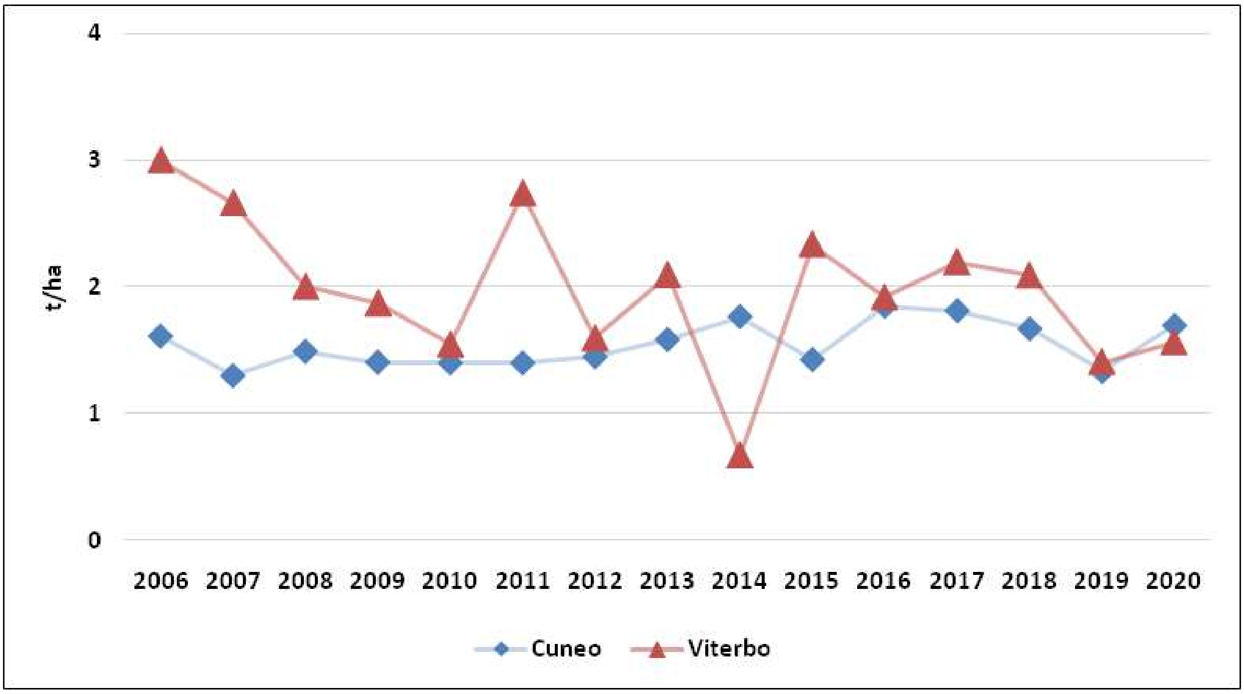

Figure 3 - Yearly average yield per hectare (t ha-1) of in-shell hazelnuts recorded for Cuneo and Viterbo provinces.

The average yield per hectare of in-shell hazelnuts obtained from mature orchards of Cuneo and Viterbo provinces (Figure 3) shows a peculiar trend characterized by large amount of nuts in on-crop years, followed by low amount of nuts in off-crop years. This is mainly due to the biennial bearing and to the high portion of senescent orchards present in oldest growing areas, where orchards are more than 50 years old.

This physiological phenomenon, commonly known as “alternate bearing”, is highly genotype-dependent and can be strongly influenced by many factors, including agronomic interventions such as irrigation, nutrition, and pruning [5]. Especially light management in the hazelnut orchard is one of the key components to ensure more consistent annual production. Reduced light levels from canopy shading reduces yield, nut quality and female flower and catkin density in the following year, and the time of shading also increases the number of poorly filled, shriveled and moldy nuts [6].

Furthermore, not suitable weather conditions deeply impact on fruit-set and fruit development increasing the tendency to alternate bearing of the trees. For instance, a frost during the fruit-set could significantly reduce the number of clusters (sink for carbohydrates) determining their reduced use of the stored carbohydrates in the lignified plant organs; as a consequence, during the following year, the twigs will produce high amount of flowers that will generate high fruit-set, establishing the consequent “alternate bearing” pattern.

2 Plant phenology and fruit development

This chapter discusses the seasonal phenology and fruit development of the European hazelnut as a physiological tool to identify the best times to carry yield estimation activities.

European hazelnut is a polygenic and polymorphic species, exhibiting a wide diversity in morphological characteristics such as plant size, growth habit, nut size, nut shape and husk length [7]. It is a monoecious, wind-pollinated, dichogamous species, characterized by a sporophytic incompatibility [8]. Its seasonal phenology is characterized by an early flowering and pollination that occurs during the winter (December – February in Europe), when the plant is still dormant. Vegetative bud-break and leaf emergence occurs in late February-March and this phenological stage is followed by fertile and vegetative shoot development. Ovule development in the female flowers occurs in early spring, and its fertilization takes place at the beginning of June, followed by the clusters and nuts development. The develop of kernel into the nut tarts in early July and may continue until the first half of August depending on the cultivar earliness (Figure 4). The ripe nuts start to naturally drop on the ground at the end of August and continues for the next three-four weeks according to the cultivar and the climatic conditions.

Referring to floral induction, which ensures flowering during the following growing season, catkin differentiation and development start in May - June, meanwhile the female flower differentiation starts later, in mid-June, and the stylar development occurs in autumn (September - November).

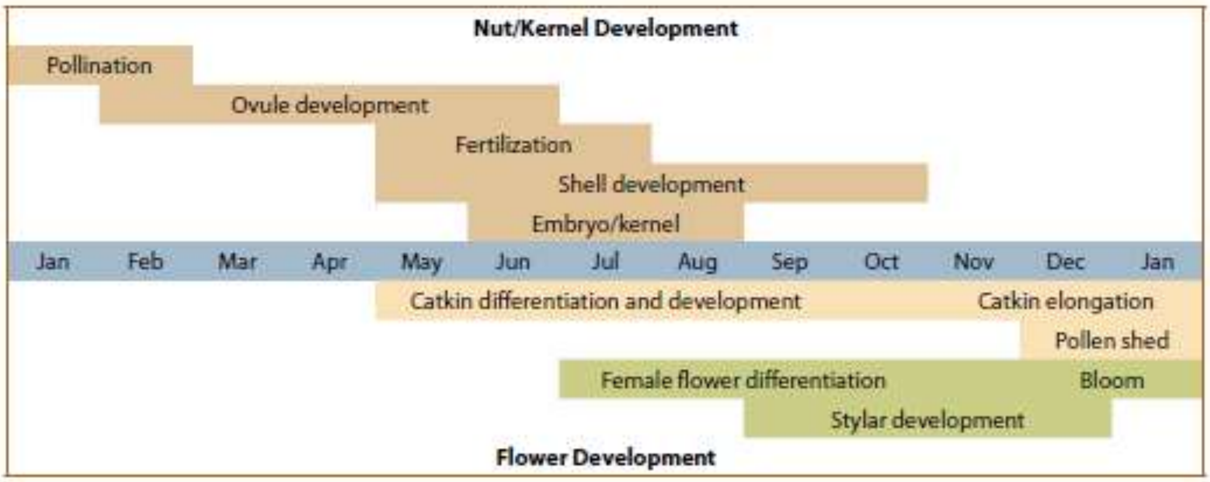

Figure 4 - Eco-physiology of European hazelnut during the growing season: reproductive phase.

In pollinated flowers, the ovules slowly develop and produce megaspores. The pollen tubes grown again, and fertilization occurs about at the beginning of June, depending on the cultivar and the growing environment. The most developed between the two initial ovules in the ovary is fertilized producing only one seed (kernel) per nut. Although, in some cultivars also both ovules can be fertilized, producing two seeds per nut (double kernel) that it is considered a defect [9]. During the following 3-4 weeks after fertilization, the embryo grows very slowly (3-5% of the final volume) and when the shell begins to lignify (early mid-July), the embryo grows very quickly and fills the nut (Figure 5).

Nut clusters that fail to develop before fertilization, abscise in late spring – early summer. This loss is due to the competition between flowers and apical dominance of the apical shoots. Sometimes the kernel does not develop determining the blank incidence (empty nuts) and in some cases it could represent a significant loss of yield. Many environmental factors can adversely affect the development of kernel, including environmental stressors as low temperature during fertilization and inadequate nutrition [10]. Furthermore, blank nuts occur more frequently in some cultivars, such as 'Barcelona' and 'Tonda di Giffoni', for their "genotype influence" [11].
The nuts are normally grouped in clusters (from 1 to 7 or more nuts per cluster according to the cultivar), and sometimes single nuts are inserted into their single husk, that which ripen and turn brown at the end of fruit maturation. The descriptors reported by many authors all over the world [12, 13] indicated that there are large differences among cultivars in terms of number of female inflorescences and, consequently, in the terms of nuts per cluster, with a variation not only ascribed to the varieties (genotype-dependent), but also related to other factors such as environmental once (seasonal conditions, management and agricultural practices).

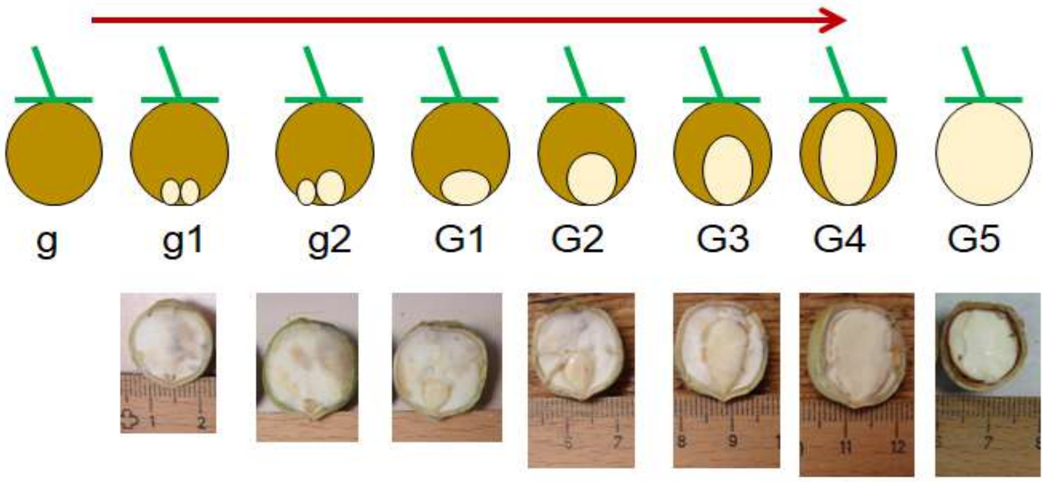

Figure 5 - European hazelnut: ovule fertilization and kernel development stages (g - no eggs observed in the medullary tissue of the ovary; g1 - two ovules of the same size are observed; g2 - one ovule more developed than the other is observed: beginning of grow

To find the best times to perform cluster and nut counts in relation to their development, maturation and early dropping, nut and kernel development monitoring was conducted during the growing season 2018 in mature trees selected in the fields 18 and 21 on cultivar 'Tonda Gentile Romana'. The observations were done weekly starting to the bud break stage until the harvest time. Table 1 reports the assigned alphanumeric phenological codes of the reproductive stages of hazelnut for modeling an accurate yield prediction time and procedure. For each representative phenological stage of fruit development, a proper phenological code has been assigned and associated to a representative picture collected during the growing season, as reported in Figure 6 - Figure 7 - Figure 8.

Observation date	Phenological code	Description of phenological stage
20/04/2018	F04.1	Beginning of ovaries enlargement (beginning glomerules appearance)
27/04/2018	F04.1	Beginning of ovaries enlargement (beginning glomerules appearance)
04/05/2018	F04.2	Ovaries enlargement
11/05/2018	F04.2	Ovaries enlargement
18/05/2018	F04.2	Ovaries enlargement
25/05/2018	F05.1	Beginning fruit development
01/06/2018	F05.1.2	Fruit development
08/06/2018	F05.2	Fruit development
15/06/2018	F05.2.3	Fruit development/end of fruit development
22/06/2018	F05.3-F05.2.1	End of fruit development/beginning of seed development
29/06/2018	F05.3-F05.3.1	End of fruit development/ seed development
06/07/2018	F05.3-F05.3.27F05.3.3	End of fruit development/ seed development/ end of seed development
13/07/2018	F05.3-F05.3.3	End of fruit development/ seed development
20/07/2018	F05.3-F05.3.3	End of fruit development/ seed development
27/07/2018	F05.3-F05.3.3	End of fruit development/ seed development
03/08/2018	F05.3-F06.1	End of fruit development/ beginning of shell brown
10/08/2018	F05.3(F06.1-F06.2)	End of fruit development/ beginning of shell brown/shell brown
17/08/2018	F05.3-F06.3-F07.1	End of fruit development/ end of shell brown/ fruit maturation
24/08/2018	F05.3-F08	End of fruit development/ Beginning of fruit dropping
31/08/2018	F05.3-F07.2/F08	End of fruit development/ Fruit maturation/ Beginning of fruit dropping
07/09/2018	F05.3-F08	End of fruit development/end of the fruit dropping
14/09/2018	F05.3-F08	End of fruit development/end of the fruit dropping
21/09/2018	F05.3-F08	End of fruit development/end of the fruit dropping
28/09/2018	F05.3-F08	End of fruit development/end of the fruit dropping

Table 1 - Alphanumeric phenological codes associated to the reproductive stages of hazelnut (cv 'Tonda Gentile Romana'). The observations were carried out during the growing season 2018 on mature trees selected in the fields 18 and 21.

Figure 6 - Evidence of seasonal development of clusters, husks, and nuts in European hazelnut, cv 'Tonda Gentile Romana', during the growing season 2018. a) 18 May; b) 25 May; c) 8 June.

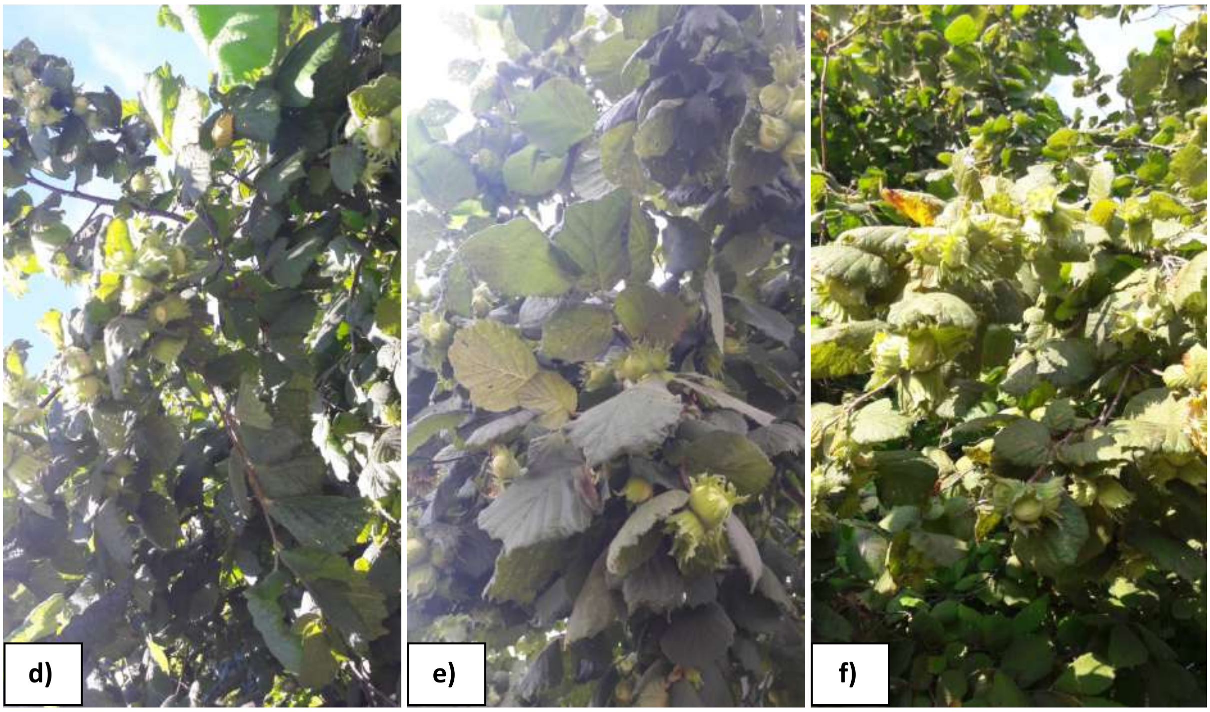

Figure 7 - Evidence of seasonal development of clusters, husks and nuts in European hazelnut, cv 'Tonda Gentile Romana', during the growing season 2018. d) 15 June; e) 29 June; f) 13 July.

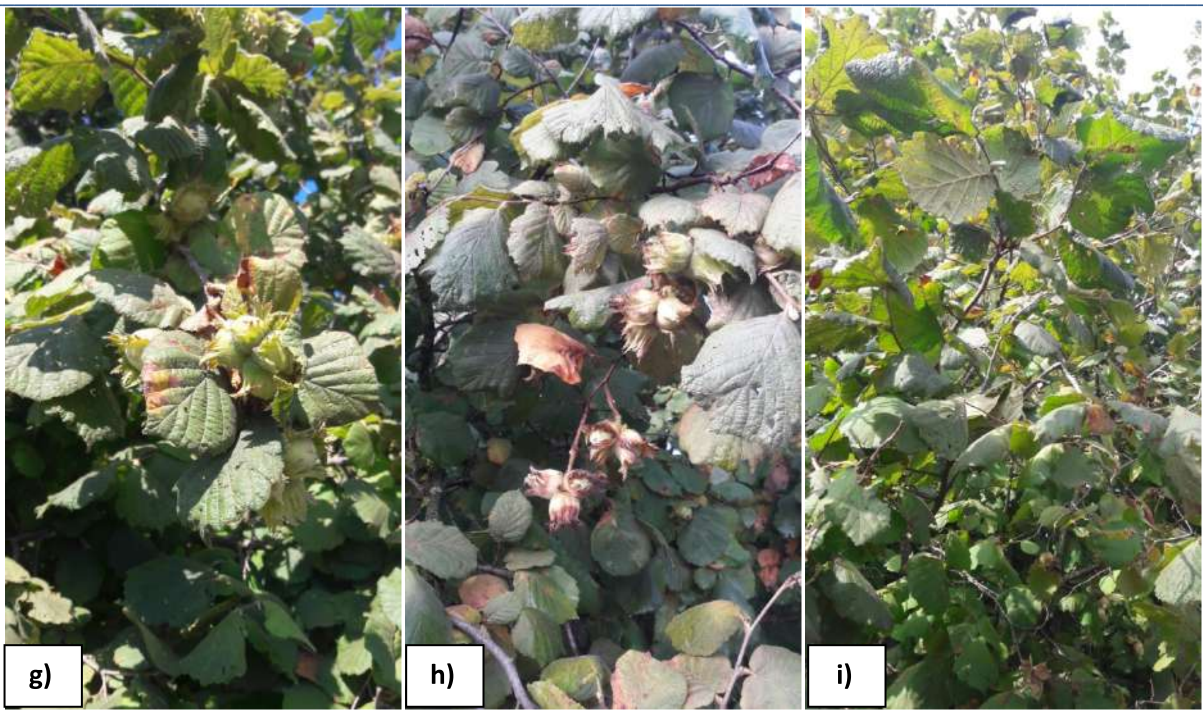

Figure 8 - Evidence of seasonal development of clusters, husks, and nuts in European hazelnut, cv 'Tonda Gentile Romana', during the growing season 2018. g) 27 July; h) 10 August; i) 24 August.

3 Nut and kernel traits of the main hazelnut cultivars

To properly adapt the yield estimation approach to different hazelnut cultivars, it is considered useful to briefly describe the differences in nut traits (shape, percent kernel and commercial uses) of the main cultivars grown in the most important hazelnut producing areas, as listed in Table 2.

The commercial hazelnut varieties grown around the world have been selected in the past to guarantee high-quality nuts, distinguishable by low incidence of defects, high kernel/nut ratio, round shape and uniform kernel calibers, pleasant kernel flavor and high incidence of pellicle removal after blanching and roasting [8, 14].

Round-shaped nut cultivars with high kernel peelability are preferred for kernel and in-shell markets (Figure 9), while low quality kernels can be used for oil extraction and cosmetics. The germplasm of European hazelnut account for more than 400 varieties worldwide, but only few tens are used in commercial hazelnut orchards [15]. Few Italian, Spanish and Turkish cultivars characterized by high kernel/nut ratio (more than 40%) and kernel size between 11 and 15 mm, matches with the high standard requests from the confectionery industries.

Figure 9 - Husks, nuts, and kernel traits of the round-shaped Italian cultivars 'Tonda Gentile Romana' (above) and 'Nocchione' (below) grown in the experimental orchard selected for the trials.

Cultivar	Area of cultivation	Main uses	Nut and kernel traits
Barcelona	Oregon, Chile, France	Kernel market and in-shell market	Large nut with round shape, 40-43% kernel, medium pellicle removal in blanched kernel
Camponica	Italy	In-shell market	Large nut with round shape, 45% kernel, medium-high pellicle removal
Tonda Gentile Romana	Italy	Kernel market	Medium nut with round shape, 45-47% kernel, medium pellicle removal
Tonda Gentile	Italy	Kernel market	Small to medium nut with triangular-round shape, 46-48% kernel, very high pellicle removal
Tonda di Giffoni	Italy	Kernel market	Medium nut with round shape, 46-48% kernel, high pellicle removal
Tombul	Turkey	Kernel market	Small nut with round shape, 54% kernel, very high pellicle removal
Pauetet	Spagna	Kernel market	Small nut with ovoid shape, 48-50% kernel, medium-high pellicle removal
Nocchione	Italy	Kernel market and in-shell market	Medium to large nut with round shape, 38-40% kernel, high pellicle removal
Negret	Spain	Kernel market	Small nut with ovoid shape, 48-50% kernel, very high pellicle removal
Palaz	Turkey	Kernel market	Small nut with round shape, 40-45% kernel, high pellicle removal, late maturity

Table 2 - Nut and kernel traits of the main commercial hazelnut cultivars grown around the world producing areas [8].

3.1 Fruit detection in commercial orchards

Sustainable agriculture is required to fulfil the pressure of growing world population, to increase production and quality without increase the environmental impact, and it could be obtained through the application of Precision Agriculture (PA), supported by advanced sensing and image processing systems [16, 17]. By combining modern machine vision with deep-learning, PA gains a revolutionary impact in various agricultural applications such as crop monitoring, disease detection and intelligent yield detection and estimation [18]. Fruit detection is essential in agronomic management, yield prediction and yield mapping since it permits the planning of other activities as automated harvesting.

Usually, fruit detection is manually performed by experts who analyze and count the fruits on the plants. However, this manual inspection is generally very time-consuming and does not allow high precision results.

Today, the interests of the fruit sectors are to apply a field estimation, particularly in orchard of medium and large size, through the combined use of image processing and deep-learning approaches. Before the use of deep-learning, automatic fruit detection was mostly done by capturing images from orchards and identifying prominent features from the images, such as size, shape, colour, and texture of the fruits, used for detection. However, this approach does not assure an accurate prediction, since it is subject to high spatial variability, due to some factors such as soil traits, fertility, water availability [19]. In contrast, deep-learning approaches bypass the manual feature extraction phase, enabling the automatic fruit detection directly from raw images. This is achieved by training the neural networks on large amounts of input data.

Although the promising results given by deep-learning approaches, e.g., [18], the major challenges limiting the automatic fruit detection are still correlated to the field variation, as well as the presence of leaves and branches which can cause occlusions and/or the presence of other fruits [17]. One of the approaches to improve the applicability of robotic harvesting is to combine human workers and robotic systems synergistically [20, 21].

4 Experimental setup for developing the protocol to detect hazelnut clusters and nuts at field-scale

4.1 Field and trial description for yield estimation at field-scale

A thorough evaluation of the plantation was conducted at the beginning of field activities to establish the trials in a portion of the hazelnut orchard suitable for the uniformity of plant growth and for giving representative to the data acquisition. The selected plants for yield estimation were geo-referenced and each plant was labeled with own assigned ID code, as reported in Table 3.

Plant ID	Latitude N	Longitude E
Yo F1	12°17'55.19''	42°16'46.99''
Yo F2	12°17'55.33''	42°16'46.99''
Yo F3	12°17'55.45''	42°16'47.01''
Yo F4	12°17'55.58''	42°16'47.02''
Yo F5	12°17'55.71''	42°16'47.03''
Yo F6	12°17'55.83''	42°16'47.04''
Yo F7	12°17'55.98''	42°16'47.04''
Yo F8	12°17'56.09''	42°16'47.05''
Yo F9	12°17'56.23''	42°16'47.06''
Yo F10	12°17'56.37''	42°16'47.07''
Ad F1	12°17'52.62''	42°16'47.14''
Ad F2	12°17'52.46''	42°16'46.84''
Ad F3	12°17'52.34''	42°16'46.53''
Ad F4	12°17'52.25''	42°16'46.20''
Ad F5	12°17'52.14''	42°16'45.85''

Table 3 - Plant ID codes and geographic coordinates (WGS84 GD) of the selected young (Yo Fn) and mature (Ad Fn) plants for fruit detection and yield estimation.

For fruit detection at field-scale, ten young trees (plants with increasing productivity) and five mature ones located in fields 16 and 18, respectively, were selected at the beginning of 2019 when the plants were still in winter dormancy (Figure 10).

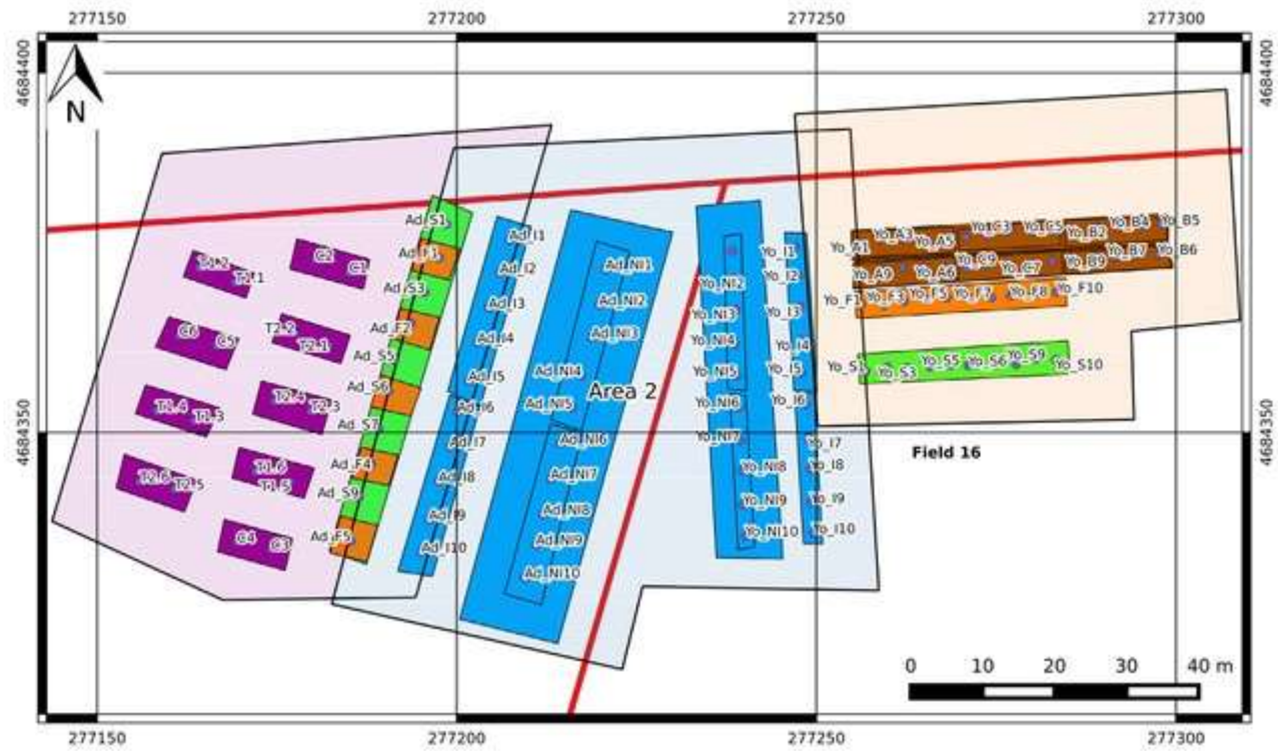

Figure 10 - Overview of the trials established in the experimental farm. The plots highlighted in orange identify the selected plants for fruit detection and estimation.

Field 16 hosted the most relevant portion of the work since the plants were not too tall, encouraging ground monitoring of clusters and nuts by hand at field-scale. This irrigated orchard was managed growing hazelnut cv 'Nocchione' planted in late 2014 (seventh leaf in the field during the growing season 2021) and spaced 4.5 x 3.0 m, giving 740 plants/ha, and trained as multi-stemmed bush (Figure 11 and Figure 12).

In the field 18, there are 30 years old plants of the cv 'Tonda Gentile Romana' (at the beginning of the trial) that grew spaced 5.0 x 5.0 m, giving 400 plants/ha, and trained as multi-stemmed bush. Some representative branches of the selected mature plants were identified in order to make a representative clusters estimation of large bushy plants, which are difficult to monitor in their entire canopy.

Figure 11 - Overview of the selected plants in the field 16 (cv 'Nocchione') monitored in late August 2019. Pictures describe the plants in ascending order from Yo F1 to Yo F10.

Figure 12 - Overview of the selected plants in the field 16 (cv 'Nocchione') monitored in early August 2020. Pictures describe the plants in ascending order from Yo F1 to Yo F10.

4.2 Clusters and nuts counting at field-scale

During the growing season 2019, an accurate manual counting of clusters and nuts was carried out in the selected plants in the field 16. This counting was performed by a two-person team of the local Unit UNITUS, as shown in Figure 13, in which one person accurately counted the clusters per branch and one person recorded the data. This prevents omitting some clusters hidden in the canopy during counting.

In the same growing season, only the manual harvest of the nuts was carried out in the field 18 on selected mature plants since the manual count in adult hazelnut shrubs is non-trivial and very time-consuming. On the contrary, manual counting of clusters compared to nuts is easier as they are larger and of light-green colour. Thus, during the following growing season (year 2020), this alternative approach was tried also for mature plants. Furthermore, since the number of nuts per cluster is a cultivar-related genetic trait with a heritability of almost 70% [7], detecting the total clusters per plant may be a satisfactory solution to determine the plant yield, also enabling the estimation of the total number of nuts per plant. The morphological traits of the clusters, such as colour and size, can also be an advantage for their detection through properly equipped UAVs and UGVs, even if clusters are often covered by leaves and are not only placed in the outer portions of the plant crown.

Figure 13 - Overview of two researchers of the local unit UNITUS during manual counting of clusters in the trial "fruit detection - young plants".

Regarding the young plants, during the growing season 2019, the manual counting of clusters and nuts per plant in field 16 took place in mid-June (Table 4) and late-August (Table 5), while the manual harvest occurred in mid-September (Table 6). Counts taken at different times during the growing season were thus compared for determining the best period to estimate yield per plant, also considering the early drop of defected nuts that do not reach the full ripeness.

Plant ID	Sampling time	Number of clusters	Number of nuts	Average nuts per cluster
Yo F1	Mid-June	8	15	1.88
Yo F2	Mid-June	25	70	2.80
Yo F3	Mid-June	21	55	2.62
Yo F4	Mid-June	47	104	2.21
Yo F5	Mid-June	34	86	2.53
Yo F6	Mid-June	76	210	2.76
Yo F7	Mid-June	26	66	2.54
Yo F8	Mid-June	67	186	2.78
Yo F9	Mid-June	57	157	2.75
Yo F10	Mid-June	66	157	2.38
Mean ± SD		**42.7 ± 23.1**	**110.6 ± 63.6**	**2.52 ± 0.3**

Table 4 - Plant ID, date of count, total number of clusters and nuts and estimated average nuts per cluster detected in mid-June 2019 on the young trees cv 'Nocchione'(mean ± standard deviation).

The count of clusters carried out in mid-June on the ten selected plants showed as the young plants of cultivar 'Nocchione' have a rather variable number of clusters, with an average of 42.7 clusters per plant. This is also reflected in the number of nuts per plant, while the average number of nuts per cluster was rather constant in all plants (mean of 2.52) with a low deviation from the mean value.

The data reported in Table 5 (related to late August) and Table 6 (related to harvesting in mid-September) largely reflect what was observed in mid-June. In particular, in late August it was observed that some clusters early fall down before the full kernel filling (dropped during in late July - early August), and the average number of clusters and total nuts showed a slight decrease with respect to the previous count (Table 4). The same trend was also observed at harvest time, as reported in Table 6.

Plant ID	Sampling time	Number of clusters	Number of nuts	Average nuts per cluster
Yo F1	Late August	8	15	1.88
Yo F2	Late August	23	65	2.80
Yo F3	Late August	21	54	2.62
Yo F4	Late August	46	103	2.21
Yo F5	Late August	34	85	2.53
Yo F6	Late August	65	180	2.76
Yo F7	Late August	25	63	2.54
Yo F8	Late August	55	153	2.78
Yo F9	Late August	54	149	2.75
Yo F10	Late August	65	155	2.38
Mean ± SD		**39.6 ± 20.1**	**102.1 ± 54.5**	**2.52 ± 0.3**

Table 5 - Plant ID, date of count, total number of clusters and nuts and estimated average nuts per cluster detected in late August 2019 on the young trees cv 'Nocchione' (mean ± standard deviation).

Plant ID	Sampling time	Number of clusters	Number of nuts	Average nuts per cluster
Yo F1	Mid-September	8	15	1.88
Yo F2	Mid-September	22	61	2.80
Yo F3	Mid-September	17	45	2.62
Yo F4	Mid-September	45	99	2.21
Yo F5	Mid-September	33	83	2.53
Yo F6	Mid-September	64	176	2.76
Yo F7	Mid-September	24	60	2.54
Yo F8	Mid-September	55	153	2.78
Yo F9	Mid-September	51	140	2.75
Yo F10	Mid-September	64	152	2.38
Mean ± SD		**38.2 ± 20.2**	**98.4 ± 54.3**	**2.52 ± 0.3**

Table 6 - Plant ID, date of count, total number of clusters and nuts and estimated average nuts per cluster harvested in mid-September 2019 on the young trees cv 'Nocchione' (mean ± standard deviation).

The approach followed during the first year of the trial, even if tested on the young plants only, was highly time-consuming. However, this allowed to verify the reliability of cluster counting only, which should be carried out after the early drop of defective nuts (late July - early August depending by the cultivar and environment). The count of the clusters only is enough to estimate the yield per plant, due to the high heritability of the genetic trait “nuts per cluster”, which makes it possible to compute the total number of hazelnuts produced by the single plant. After nut and kernel traits determination it is possible to estimate the real yield per plant with more accuracy.

Referring to the mature plants (field 18 - cv 'Tonda Gentile Romana'), two manual nuts harvests per plant were carried out, since the hazelnut shows scalarity in nuts ripeness. The yield of mature plants, expressed as sum of the two harvesting conducted in early and middle September, respectively, showed an average value of 9 Kg of in-shell nuts per plant. Furthermore, more than three quarters of the production per plant was on the ground during the first harvest, as shown in Table 7.

Plant ID	Sampling time First harvesting	Yield First harvesting (kg/plant)	Sampling time Second harvesting	Yield Second harvesting (kg/plant)	Total Yield (kg/plant)
AdF1	Early September	4.90	Mid-September	1.47	6.37
Ad F2	Early September	5.77	Mid-September	1.73	7.50
Ad F3	Early September	6.88	Mid-September	1.72	8.60
Ad F4	Early September	10.13	Mid-September	1.01	11.14
Ad F5	Early September	9.67	Mid-September	0.97	10.63
Mean ± SD		**7.47 ± 2.33**		**1.38 ± 0.37**	**8.85 ± 2.03**

Table 7 - Plant ID, time of harvesting and yield determined in 2019 on the mature plants cv 'Tonda Gentile Romana' (mean ± standard deviation).

Regarding the young plants, during the growing season 2020 (second year of the trial), the manual count of the clusters took place in early July, middle August, and early September, on the selected plants in field 16. In June and August, the total number of nuts per plant was estimated from the average number of hazelnuts per cluster, while at harvest time, total number of nuts per plant was also determined through their real counting.

Since the plants in the young orchard (field 16) grew considerably, also entering in the phase of increasing productivity, there were significantly more clusters per plant during the growing season 2020 compared to the previous year.

In early July 2020 (Table 8), the observed average number of clusters per plant was significantly higher than the one recorded the previous year in the same plants (1,127 in 2020 versus 37 in mid-June 2019). Nevertheless, as expected, the average number of nuts per cluster was very similar to that observed in 2019.

The manual counting of clusters per plant during 2020 was also supported by 3D acquisitions conducted using the UGV by the team of Local Unit UNIROMA3, during the same days that manual counts were conducted.

Plant ID	Sampling time	Number of clusters	Number of nuts	Average nuts per cluster
Yo F1	Early July	459	1,299	1.88
Yo F2	Early July	559	1,156	2.80
Yo F3	Early July	948	1,368	2.62
Yo F4	Early July	1,092	2,349	2.21
Yo F5	Early July	981	2,756	2.53
Yo F6	Early July	1,564	4,060	2.76
Yo F7	Early July	1,404	3,491	2.54
Yo F8	Early July	1,463	3,917	2.78
Yo F9	Early July	1,288	2,758	2.75
Yo F10	Early July	1,516	5,678	2.38
Mean ± SD		**1,127.4 ± 391.8**	**2,883.2 ± 1443.1**	**2.49 ± 0.6**

Table 8 - Plant ID, date of count, total number of clusters and nuts and estimated average nuts per cluster detected in early July 2020 on the trees of cv 'Nocchione' (mean ± standard deviation).

Table 9 summarizes the data acquired in the second seasonal manual count of clusters per plant, conducted in mid-August. As observed in the previous year, the number of clusters and nuts was very similar to that observed in the previous seasonal count, confirming that production losses during this period of the growing season do not highly affect the final yield.

Furthermore, as showed in Table 10, on September 10th, the number of clusters and hazelnuts per plant manually harvested (Figure 14) and counted was lower in comparison to that observed during the previous and early counts. However, it is worth highlighting that such a loss of production is expected in fruit crops since it is related to many eco-physiological and agronomical issues [20].

Figure 14 - Overview of the local unit UNITUS team during manual harvesting of nuts. On left - nuts manually grouped in front of own plant; On right - researchers harvesting nuts by hands.

Plant ID	Sampling time	Number of clusters	Number of nuts	Average nuts per cluster
Yo F1	Mid-August	420	1,189	1.88
Yo F2	Mid-August	536	1,108	2.80
Yo F3	Mid-August	932	1,345	2.62
Yo F4	Mid-August	1,014	2,181	2.21
Yo F5	Mid-August	930	2,613	2.53
Yo F6	Mid-August	1,551	4,026	2.76
Yo F7	Mid-August	1,380	3,431	2.54
Yo F8	Mid-August	1,460	3,909	2.78
Yo F9	Mid-August	1,266	2,711	2.75
Yo F10	Mid-August	1,506	5,641	2.38
Mean ± SD		**1,099.5 ± 401.2**	**2,815.2 ± 1,460.6**	**2.49 ± 0.6**

Table 9 - Plant ID, date of count, total number of clusters and nuts and estimated average nuts per cluster detected in mid-August 2020 on the trees of cv 'Nocchione' (mean ± standard deviation).

Plant ID	Sampling time	Number of clusters	Number of nuts	Average nuts per cluster
Yo F1	Early September	378	1,070	1.88
Yo F2	Early September	354	731	2.80
Yo F3	Early September	764	1,103	2.62
Yo F4	Early September	953	2,050	2.21
Yo F5	Early September	893	2508	2.53
Yo F6	Early September	1,210	3,140	2.76
Yo F7	Early September	1,076	2,676	2.54
Yo F8	Early September	1,314	3,518	2.78
Yo F9	Early September	949	2,033	2.75
Yo F10	Early September	1,145	4,287	2.38
Mean ± SD		**903.5 ± 325.4**	**2,311.6 ± 1148.0**	**2.49 ± 0.6**

Table 10 - Plant ID, date of count, total number of clusters and nuts and estimated average nuts per cluster harvested in early September 2020 on the trees of cv 'Nocchione' (mean ± standard deviation).

As far as the mature plants (field 18 - cv 'Tonda Gentile Romana') are concerned, during the growing season 2020 some representative branches of the selected plants were labelled and, similarly to the young plants, 30 clusters per branch were counted for their nuts content in early August, to determine the average number of nuts per cluster. As shown in Table 11, the average number of nuts per cluster revealed for the cultivar 'Tonda Gentile Romana' was slightly higher when compared to 'Nocchione' with average values of 2.8 nuts per cluster. This finding is in agreement with the literature [22].

Regarding 2021, similarly to the previous two years, the counting of clusters per plant in both trials is in progress (data not shown). Anyway, this growing season is not very suitable for replicating field observations, as the seasonal production of the plants was almost entirely compromised by extensive cold damages to plants caused by late spring frosts (- 8°C degrees in early April 2021). For this reason, additional five plants were added to the young trial in order to increase the chances for validation of the developed mathematical production estimation model.

Plant ID	Sampling time	Average nuts per cluster
AdF1	Early August	3.2±1.3
Ad F2	Early August	2.4 ± 0.8
Ad F3	Early August	3.0 ±1.0
Ad F4	Early August	2.8 ±1.0
Ad F5	Early August	2.6 ±1.1
Mean ± SD		**2.8±1.1**

Table 11 - Plant ID, time of clusters and nuts counting and total nuts per cluster(selected 30 clusters per branch) determined during the growing season 2020 on the mature plants cv 'Tonda Gentile Romana' (mean ± standard deviation).

4.3 Yield efficiency of the selected plants

In order to reinforce the yield determination at field-scale, the vigour of the selected plants in the field 16 (cv 'Nocchione') was measured during the growing season 2020 as vegetative trait to use for calculating the plant yield efficiency (YE).

YE is expressed as ratio between production of in-shell nuts per trunk-cross sectional area (TCSA) according to the formula YE= yield * $TCSA^{-1}$.

TCSA is commonly measured in perennial fruit crops as vigour parameter of the plant and in the hazelnut tree it is determined by measuring the trunk-cross sectional area 20 cm above the ground of each lignified stem of the shrub at the end of vegetative season (after nut harvesting). The real TCSA of the shrub is given by the sum of each measured TCSA value.

Table 12 reports the YE of the ten selected plants in field 16 which showed high uniformity of the parameter ranging from 0.012 in Yo F2 to 0.016 in Yo F8, with an average value of 0.0137, expressed as mean of YE per plant.

This parameter, which is more stable than the plant yield alone, can be particularly useful for the development of the mathematical model for hazelnut yield estimation, as described in the next section.

Plant ID	Yield (kg)	TCSA (cm^2)	YE (Kg/cm^2)
Yo F1	2.68	215.45	0.0124
Yo F2	2.52	210.39	0.0119
Yo F3	4.36	292.67	0.0148
Yo F4	4.83	343.37	0.0140
Yo F5	4.36	321.66	0.0135
Yo F6	5.34	409.44	0.0130
Yo_F7	4.55	361.18	0.0126
Yo F8	5.19	337.94	0.0153
Yo F9	5.10	396.97	0.0128
Yo F10	6.38	389.18	0.0163
Mean ± SD	**4.53± 1.17**	**327.8 ± 70.2**	**0.0137 ± 0.001**

Table 12 - Plant ID, yield, TSCA at 20 cm above the ground expressed as sum of the various shrub stems and YE calculated at the end of growing season 2020 on the plants of cv 'Nocchione' (mean ± standard deviation).

4.4 Nut and kernel traits and defected nuts: evaluation at lab scale

In order to determine the quality of the nuts and determine the net yield excluding defective hazelnuts (especially empty nuts), the nut and kernel weight, width, thickness and height, and shell weight were recorded on sub-samples of 50 nuts per plant in the laboratory. Nut and kernel shape, obtained as [(width+thickness/2)/height], and kernel/nut ratio were calculated according to [22].

The most important trait revealed during the nut and kernel traits characterization was the incidence of empty nuts at harvest time, which was slightly high, with mean values equal to 14% (Table 13).

This trait was considered during the determination of the net yield per plant, subtracting from the gross yield per plant the percentage of empty hazelnuts detected during the lab activities, similarly to what happens in the hazelnut trading sector.

Plant ID	Nut weight (g)	Nut diam > (mm)	Nut diam < (mm)	Nut height (mm)	Nut round index	Kernel weight (g)	Shell weight (g)	Kernel/nut ratio (%)	Empty nuts (%)
Yo F1	2.1±0.6	18.9±1.3	16.3±1.3	18±1.5	1.0±0.1	0.8±0.3	1.4±0.3	36.8±5.8	6
YoF2	2.1±0.8	19.2±1.1	16.9±1.4	19.1±1.4	0.9±0.1	1.2±0.2	1.4±0.2	45.6±3.1	34
Yo F3	1.5±0.5	17.7±1.7	15.3±1.6	17.0±1.6	1.1±0.1	0.6±0.2	1.1±0.2	33.7±6.3	16
Yo F4	2.2±0.5	18.7±1.1	16.8±1.4	18.6±1.3	1.0±0.1	1.1±0.2	1.2±0.2	46.6±4.1	6
Yo F5	1.8±0.5	18.5±1.3	15.9±1.2	18.0±1.1	1.0±0.1	0.7±0.2	1.2±0.3	34.7±4.1	4
Yo F6	1.3±0.4	17.5±1.3	14.8±1.2	17.1±1.3	0.9±0.1	0.5±0.2	1.0±0.2	31.1±5.1	22
Yo F7	1.4±0.4	17.4±1.5	15.0±1.2	17.4±1.3	0.9±0.1	0.5±0.2	1.0±0.5	31.7±5.4	20
Yo F8	1.5±0.3	18.1±1.2	15.5±1.0	17.9±1.2	0.9±0.1	0.4±0.1	1.0±0.2	29.0±5.3	4
Yo F9	1.4±0.5	17.9±1.5	15.2±1.1	17.4±1.1	1.9±0.1	0.4±0.2	1.0±0.2	28.7±7.8	18
YoF10	1.3±0.3	17.5±1.4	15.0±1.3	17.6±1.5	0.9±0.0	0.4±0.1	0.9±0.2	28.5±6.5	14
Mean ± S.D.	**1.7±0.4**	**18.1±0.6**	**15.7±0.8**	**17.8±0.7**	**1.1±0.3**	**0.7±0.3**	**1.1±0.2**	**34.6±6.6**	**14.4±9.7**

Table 13 - Plant ID, nut and kernel traits and incidence of blank (empty nuts) revealed in the plants of the trial Yo Fn (cv 'Nocchione'), carried out in the growing season 2020.

5 Mathematical model for early detection of clusters and nuts and orchard yield estimation

5.1 Motivation and initial analysis

Fruit monitoring and production estimation are two very important aspects of the management of hazelnut orchards. They can be used to predict the final production of the orchard and to adapt the actions and resources accordingly.

On this subject, most research in PA focuses on fruit detection [21 – 23] where the aim is at computing an accurate number of fruits per each single tree. Of course, this is a very important aspect of fruit production monitoring, as an accurate counting of the current production is essential to obtain a meaningful estimation. However, for small fruit crops like hazelnuts and in most large-scale scenarios this alone is not a sufficient solution for the production monitoring process, as it is not realistic to assume that all trees can be measured individually. Additionally one must notice that a proper production estimation requires the continuous evaluation of the plants as they are subject to climate disturbances that may damage the fruits and reduce the overall production [24, 25]. Of course, measuring continuously every single tree is even more unrealistic.

Accordingly, the development of tools for the estimation of the production in large-scale orchards based on sparse measurements in space and time is essential. A starting assumption to do so is the observation that all trees in a same sufficiently small area of a plantation are subject to similar conditions and thus have comparable losses, as seen in Section 4. This fact allows to estimate the state of unmeasured trees and to assess the total production based only on measurement of some plants. In this part of the deliverable, we will present our seminal results on how to exploit this information in order to obtain an estimation of the total yield of the orchard. The results of this work are based on experimental data obtained from some members of the PANTHEON consortium and on the data collected by orchard from the local Unit UNITUS team, as shown in the previous sections.

5.2 Production monitoring, public data available, and the importance of reliable measurements

Precise production monitoring strategies are usually implemented in controlled environment and small areas, such are greenhouses or small orchards [26, 27]. In these scenarios, all plants are frequently measured by different detection systems providing an estimation of the current production. However, the methodologies developed for these scenarios, while allowing a great accuracy in the counting of fruits, cannot be scaled up to large-scale orchards as they base their approach on the assumption that all plants can be constantly measured.

For hazelnuts, monitoring activities of the production of large-scale orchards based on partial measurements are already being implemented in several real contexts. In these cases, punctual measurements are used to extrapolate the total production estimation assuming a homogenous development.

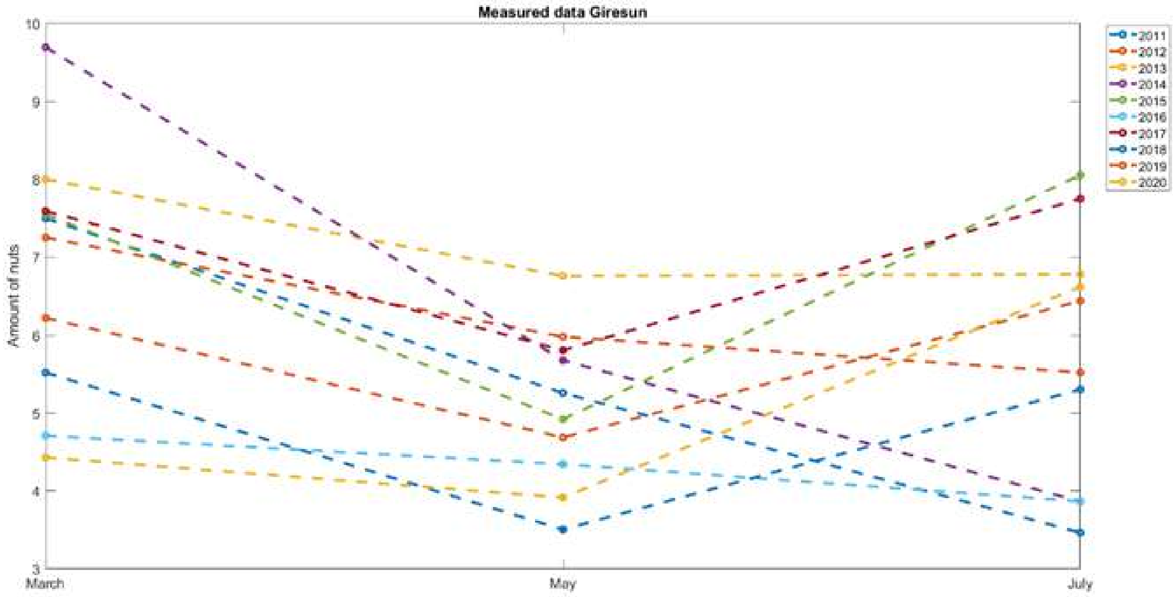

Figure 15 - Variation in the production monitoring in the province of Giresun, Turkey.

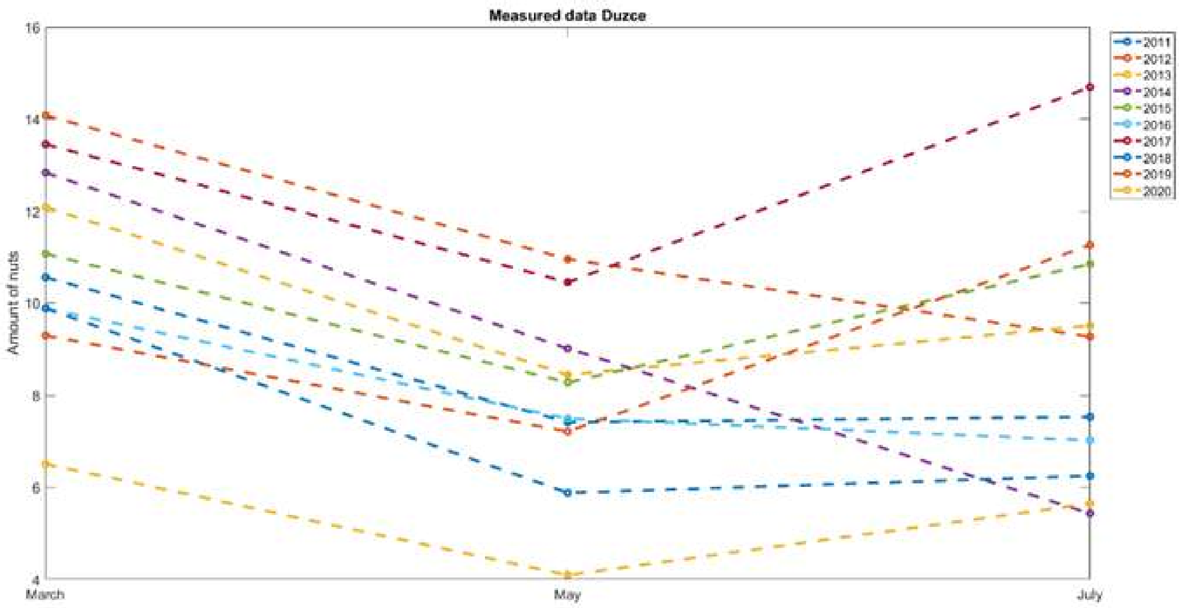

Figure 16 - Variation in the production monitoring in the province of Duzce, Turkey.

The starting point of our work on production monitoring was analyzing available data from fruit counting activities carried around the world. However, the result of this analysis was that, mostly due to the nonmechanization of this task, the data obtained in real production contexts often contains errors and large variances which impair their reliability. Indeed, the data we obtained shows a large variability and some clearly anomalous behaviors. Consider for instance Figure 15 and Figure 16, which depict the measured production of nuts during several years in two different regions of Turkey. In these plots, it can be seen that the quantity of nuts follows unexpected patterns, providing for some years a marked decreasing and increasing of the production during the same growing season, which can only be explained by errors in the way nuts were counted.

A first observation that arises from this preliminary analysis is that an insufficiently careful counting of the number of nuts may lead to results of very dubious usefulness. As shown in the previous section, performing this operation manually is extremely labor-intensive and time-consuming and cannot be realistically carried out in a real production context. This observation advocates for the importance of the automatic counting algorithms that are currently being developed within PANTHEON.

A second observation is that it is not possible to use available data collected outside PANTHEON to set and validate production models. Accordingly, we were forced to use only data coming from the PANTHEON trials to identify a preliminary model as, to the best of our knowledge, it is the only reliable dataset available for hazelnuts.

5.3 PANTHEON Data Analysis

As described in the Section 4, ten trees from the field 16 of the PANTHEON experimental orchard were chosen for the task of fruit development and production monitoring.

During the growing seasons 2019 and 2020, the selected trees were monitored in terms of number of clusters and nuts. The measurements provided the quantity of nuts and clusters at three different instants of the growing season. The collected data allowed to obtain the evolution of the number of nuts during two full growing seasons.

The data obtained during the year 2019 are summarized in Table 14 based on the tables reported in the Subsection 4.2.

Id plant	Measurement 1 (29/06/2019)	Measurement 2 (23/08/2019)	Measurement 3 (15/09/2019)
Yo F1	15	15	15
Yo F2	70	65	61
Yo F3	55	54	45
Yo F4	104	103	99
Yo F5	86	85	83
Yo F6	210	180	176
Yo F7	66	63	60
Yo F8	186	153	153
Yo F9	157	149	140
Yo F10	157	155	152

Table 14 - Number of in-shell nuts counted in 2019 in the "young trial" fruit detection.

From these measurements, it can be observed that the number of nuts per plant is very low compared to the year 2020 (reported in Table 15 and discussed below). This can be explained by the fact that during the season 2019 the plants were still young, and they had just initiated the production phase.

To provide a clearer visualization of the data, the evolution of the number of nuts in each tree is depicted in Figure 17. The figure shows that most of the plants keep the same number of nuts during the summer, with just some of them losing a small quantity of fruits in between the measurements. This behavior can be explained by the small production of nuts in each plant, which arguably made the plants less prone to lose nuts.

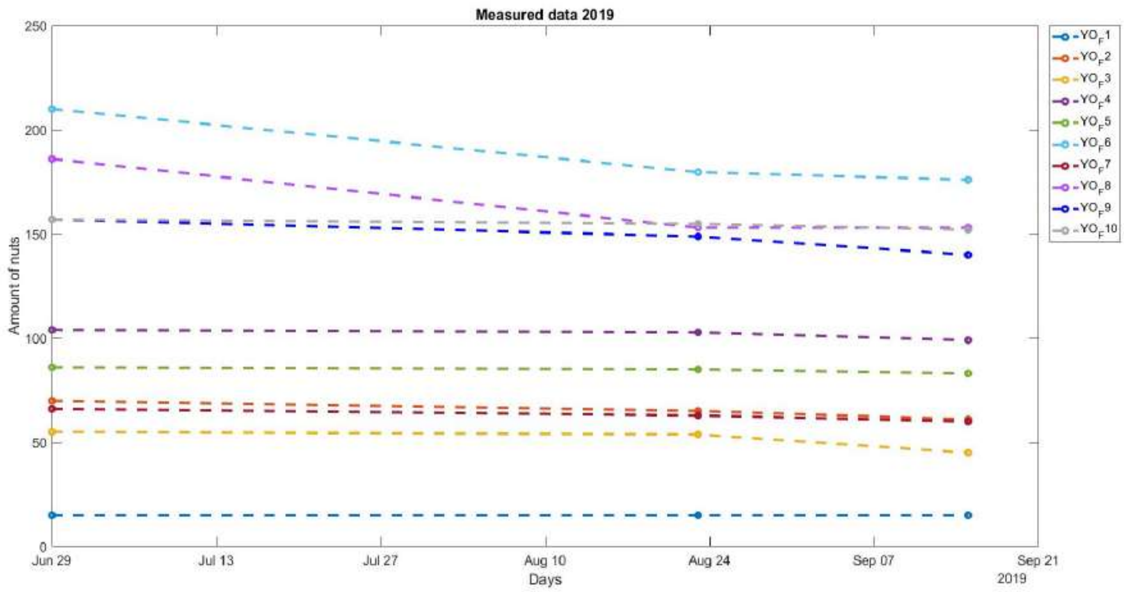

Figure 17 - Evolution of the number of nuts in the selected plants (growing season 2019).

The measurements from the season 2020 are synthetically shown in Table 15. The reported values show a clear increment in the production during the second year. The quantity of clusters and nuts lost between measurements is more evident than in the previous year, supporting the assumption that the small production of nuts made 2019 an abnormal year in terms of production evolution. The agronomists of PANTHEON also confirmed that, from a qualitative viewpoint, this behavior is more representative of a typical adult tree.

The evolution of the data collected during the second year of measurements is shown in Figure 18. In this plot, the progression of the number of nuts is clearly monotonically decreasing during the season, contrary to the data shown in Figure 15 and Figure 16. Additionally, the loss of nuts during the second part of the season (end of August and beginning of September) is visibly higher than the first part of the season in almost all monitored plants.

Id plant	Measurement 1 (08/07/2020)	Measurement 2 (12/08/2020)	Measurement 3 (10/09/2020)
Yo F1	459	420	378
Yo F2	559	536	354
Yo F3	948	932	764
Yo F4	1,092	1014	953
Yo F5	981	930	893
Yo F6	1,564	1,551	1,210
Yo F7	1,404	1,380	1,076
Yo F8	1,463	1,460	1,314
Yo F9	1,288	1,266	949
Yo F10	1,516	1,506	1,145

Table 15 - Number of clusters counted in 2020 in the "young trial" fruit detection.

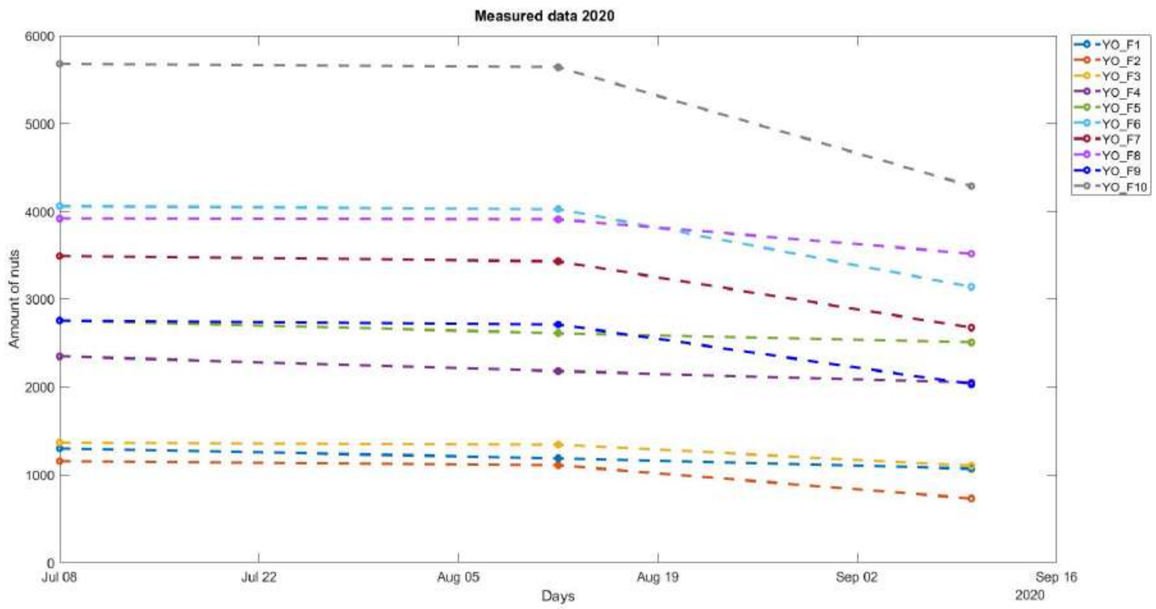

Figure 18 - Evolution of the number of nuts in the selected plants (growing season 2020).

From the data collected over the period 2019-2020, we can observe that all monitored trees followed a similar behavior regarding the evolution of the number of nuts, which is a first promising step to obtain an estimated model of the production.

Note that there are remarkable differences between the two monitored seasons: arguably the measurements of the year 2019 show a plant still young and not producing a great amount of nuts. On the contrary, data from 2020 show more mature plants. Since we believe this behavior to be more representative of what happens with adult trees in a plantation, the rest of this analysis focuses only on the data obtained during the year 2020.

5.4 Normalized data

The next step of our analysis was to normalize the measurements based on the initial quantity of fruits at the beginning of the season. This operation provides the evolution of the production in terms of percentage of initial nuts per plant, allowing us to better compare the trends in the production of the different plants despite their difference in total number of nuts. The objective of this step was to evaluate if the loss of nuts during the season could be approximated to a *rate* based on the number of nuts or just to independent disturbances related to other factors.

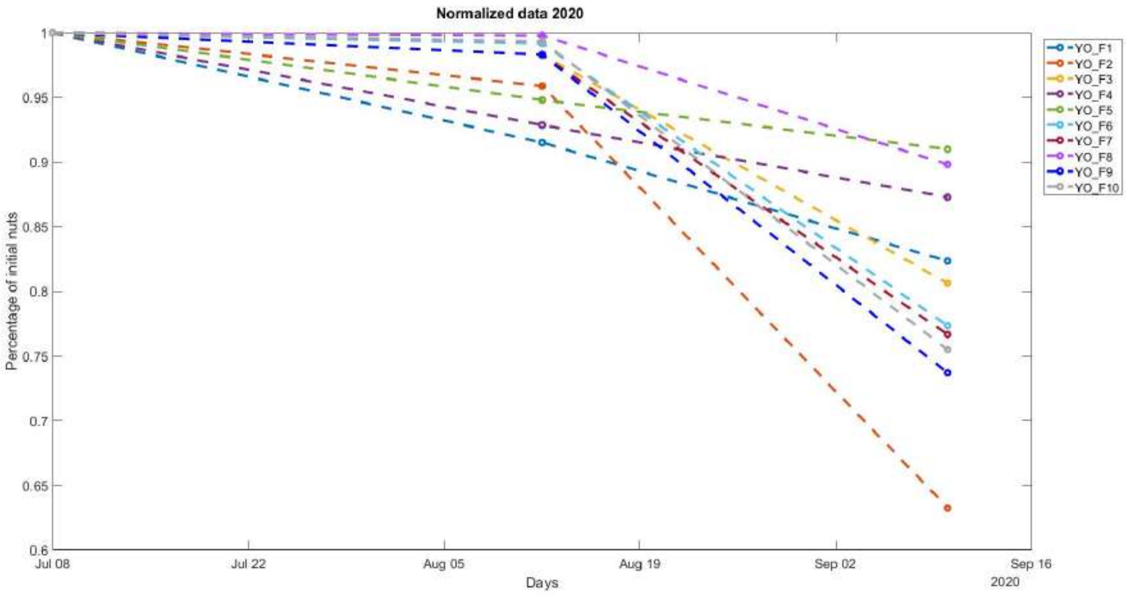

Figure 19 - Evolution of the percentage per plant with respect to the initial measurement.

In Figure 19, the ratio of nuts with respect to the initial quantity is depicted for each plant at the three different measurements. This plot shows how the behavior of the ten plants is more or less similar showing a greater percentage of lost nuts toward the end of the summer. The ratio $r_i(t)$ represented for each tree i is computed as follows:

$$r_i(t) = \frac{x_i(t)}{x_i(t_0)} \quad (1)$$

where $x_i(t)$ is the number of nuts counted at time t and t_0 represents the time of the initial measurement.

Similarly, we can plot the ratio of lost nuts in each plant with respect to the initial value, such as:

$$l_i(t) = 1 - r_i(t) \quad (2)$$

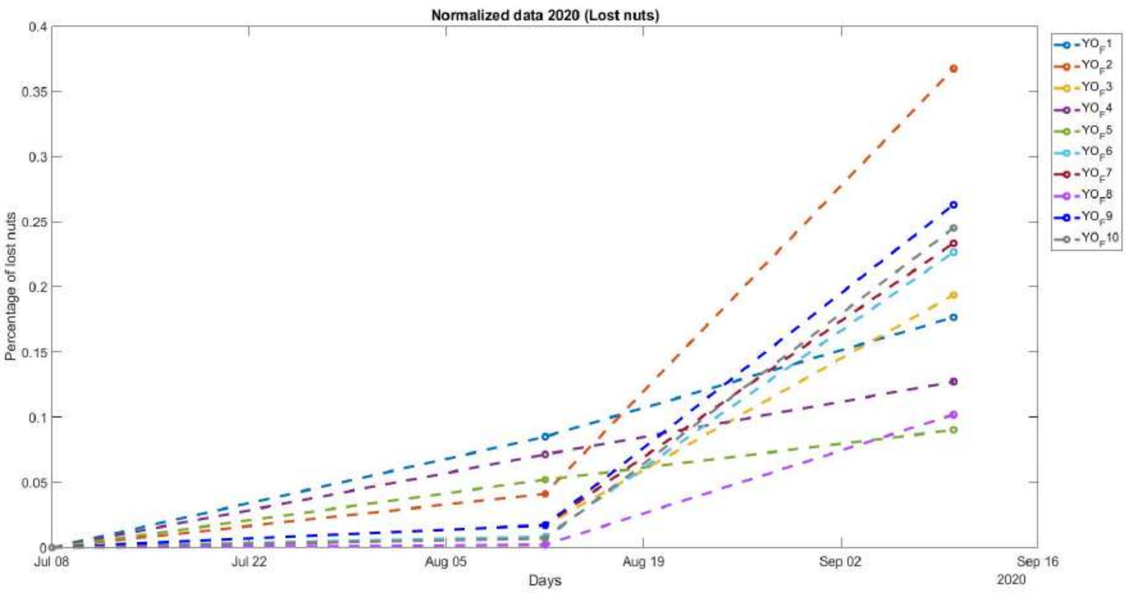

Figure 20 - Evolution of the percentage of lost nuts per tree.

Figure 20 shows the evolution in the ratio of lost nuts for each plant. This plot helps to appreciate better the similarity in the proportional response of the different plants.

In summary, the results from this subsection show that the trend in most of the monitored trees is similar, showing a clear increase in the loss of nuts at the end of the season. They are also a promising step towards the achievement of a model to estimate the production of the trees which is explored in the following sections.

5.5 Choice of a mathematical model

In order to perform an estimation and a prediction, it is fundamental to build a model of the way the number of nuts evolves over time. A starting point of the model is that the evolution of the number of nuts in the plant is monotonically decreasing (as nuts cannot be created halfway in the season) and that the decrease is time-varying. Given the limited availability of data this kind of behavior can be mathematically modeled in many different ways. We investigated two models which seemed reasonable.

Single integrator with disturbances - Let us define $x_i(t)$ as the quantity of nuts in a plant i at time t. In this case, we assume that the number of nuts tends to remain constant, and it decreases only subject to disturbances $w_i(t) < 0\ \forall t$, related to heat, rain or wind, that decrease the number of nuts.

Therefore, the time evolution of the number of nuts in the plant i can be mathematically described as:

$$\dot{x}_i = w_i(t) \tag{3}$$

where the loss disturbance $w_i(t)$ of each plant is correlated with the ones of neighboring plants.

Time-varying rate - In this case, by considering $x_i(t)$ as the quantity of nuts, we assume that the number of nuts is decreasing at a certain rate proportional to $x_i(t)$. This rate $\kappa_i(t) \geq 0$ can be time variant providing the following evolution of the system:

$$\dot{x}_i = -\kappa_i(t)x_i(t) \tag{4}$$

where the $\kappa_i(t)$ of neighboring plants is correlated. Furthermore, we assume that $\kappa_i(t) = \hat{\kappa}_i(t) + \tilde{\kappa}_i(t)$where $\hat{\kappa}_i(t)$ is the nominal "expected" loss of the ith plant and $\tilde{\kappa}_i(t)$ represents the effect of disturbances on this estimation.

Although in line of principle these two models are equivalent and thus data can be justified with both of them, from the practical viewpoint a model with a time-varying loss rate seems to allow an easier interpretation of the data and in the way $\kappa_i(t)$ are correlated. Indeed, observing Figure 19 we can see that there is a very similar rate of loss of nuts for all the monitored plants. Accordingly, an easy and promising way to estimate a sufficiently homogeneous portion of an orchard is to assume a common rate for a nominal model in the form:

$$\dot{x}_i = -\hat{\kappa}(t)x_i(t) \quad i = 1, \dots, N \tag{5}$$

and using the observations to estimate the disturbances $\tilde{\kappa}_i(t)$, which are assumed to be (almost) constant in a same homogeneous portion of an orchard.

The next subsection aims at identifying the parameter for the tentative model (4) for the experimental orchard PANTHEON based on the data collected. Of course, this model should be refined with more data, when available.

5.6 Model identification

In this subsection we show that by using the data collected in 2020, we can identify a model of the losses of the system in the form of a rate associated to the number of nuts in the plant.

Based on the data obtained during the experimental trials in PANTHEON, the identification of the specific continuous time time-varying rate is not realistic given the limited amount of measurements available. Accordingly, we made the simplifying assumption that the loss rate is constant over time, i.e., $\hat{\kappa}_i(t) = \hat{k}\ \forall i$.

In general, for an autonomous LTI system like the one we are trying to identify, the explicit solution to (5) has the following form:

$$x(t) = e^{-\hat{\kappa}t} x_0 \tag{6}$$

x_0 is the initial condition and where we assume $t_0 = 0$.

It should be noticed that Figure 19, which depicts the evolution of the number of nuts based on the initial measurement, represents precisely the following evolution:

$$r(t) = \frac{x(t)}{x_0} = \frac{e^{-\hat{k}t} x_0}{x_0} = e^{-\hat{k}t} \tag{7}$$

for the different monitored trees of the PANTHEON orchard. Thus, it can be used to estimate a tentative nominal rate $\hat{k}$.

The identified curve

$$r(t) = e^{-\hat{\kappa}t} \tag{8}$$

obtained by using the three measurements from the ten selected plants provided the following parameter (± standard error)

$$\hat{\kappa} = 0.002767\ (\pm 0.00067).$$

The different indicators describing the goodness of the obtained fit are shown in Table 16.

Indicator	Value
SSE	0.1251
R^2	0.5941
RMSE	0.05038

Table 16 - Numerical evaluation of the goodness of the fitting.

The values in Table 16 reveal that the fitting obtained based on the data collected from PANTHEON is reasonable. The indicators residual of the sum of squares (SSE) and root mean square error (RMSE) represent *errors* of the fitting, meaning the lower the values, the better the estimation. In the case of the R^2 indicator, i.e., coefficient of determination, it tends to 1 when all outputs can be perfectly explained by the inputs. Therefore, a value higher than 0.5 describes a good correlation.

It follows that we can compute the explicit solution to the system in the form:

$$x(t) = e^{-\hat{\kappa}t}x_0 \tag{9}$$

based on the identified parameter $\hat{\kappa}$.

5.7 Naïve validation of the identified model

This subsection aims at performing (in the limit of the possible) a validation of some of the hypotheses of our model. In particular, we will show that that the assumption about the structure of $\kappa_i(t)$, i.e., $\kappa_i(t) = \hat{\kappa} + \tilde{\kappa}(t)$ where $\tilde{\kappa}(t)$ is a production disturbance that is common to all the trees in a certain area, is reasonable. In line of principle, for this purpose, we would need a much larger dataset containing data from multiple years. As this is not possible within the context of PANTHEON project, we propose a simple procedure that, although cannot fully validate the assumption, can at least reassure us the fact that it looks reasonable.

In particular, we computed $\kappa_i(t) = \hat{\kappa} + \tilde{\kappa}(t),\ i = 1, 2$ from 20% of our data (two randomly selected plants) and show that, by extending the $\tilde{\kappa}(t)$ estimated on these two trees to all the trees, a reasonable estimation of the production is still produced (assuming to know the initial conditions).

The procedure was repeated on all N_c possible combinations of datasets, following the methodology represented in Figure 21.

This procedure can be described as follows:

1. We identify $\tilde{\kappa}(t)$ using the data from two plants;
2. We compute the solution of the model $\hat{x}_i(t_3) = e^{-(\hat{\kappa}+\tilde{\kappa}(t))t}x_i(0)$ with $x(0)$ the first measurement of the other eight plants and t_3 is the time of the third measurement;
3. We compare $\hat{x}_i(t_3)$ from the estimated model with the measurement $x_i(t_3)$
$$e_i = \left|\frac{x_i(t_3) - \hat{x}_i(t_3)}{x_i(t_3)}\right| \tag{10}$$
where i corresponds to the index of the selected plants;
4. We compute the average error of all plants:

$$error = \frac{\sum e_i}{N - 2} \tag{11}$$

where N is the total number of plants.

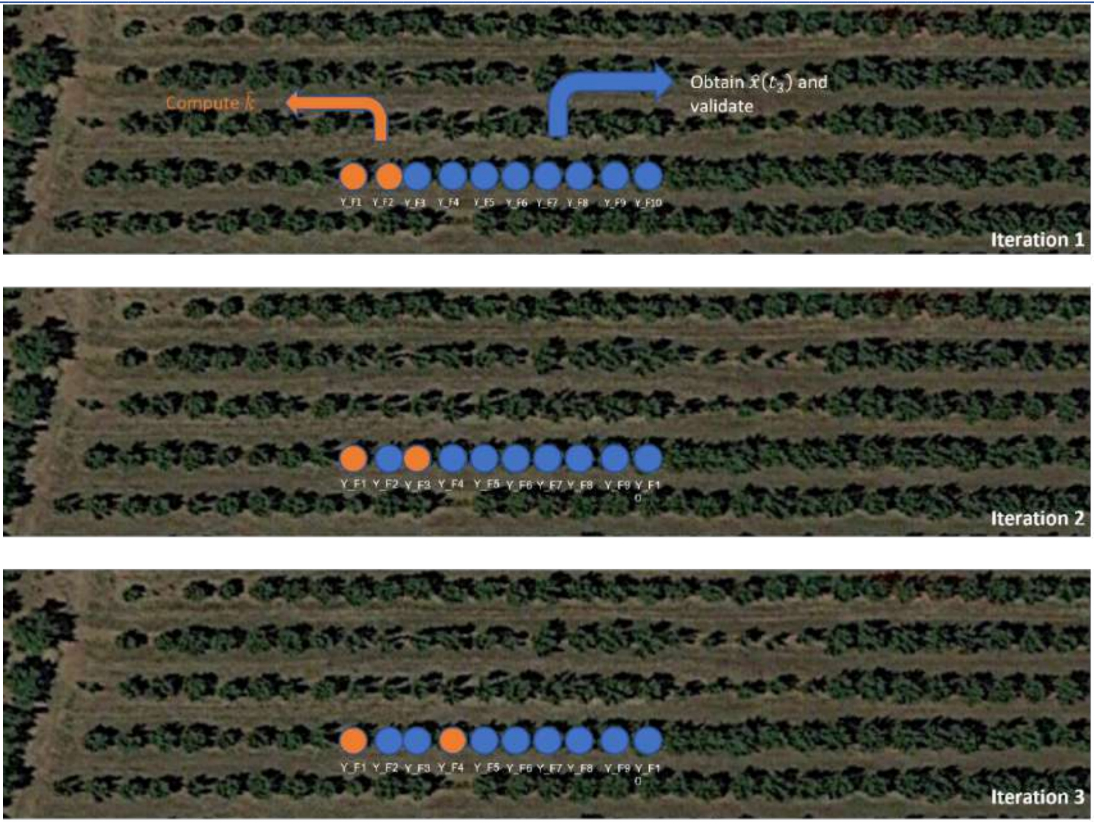

Figure 21 - Procedure of selecting the training and validation datasets.

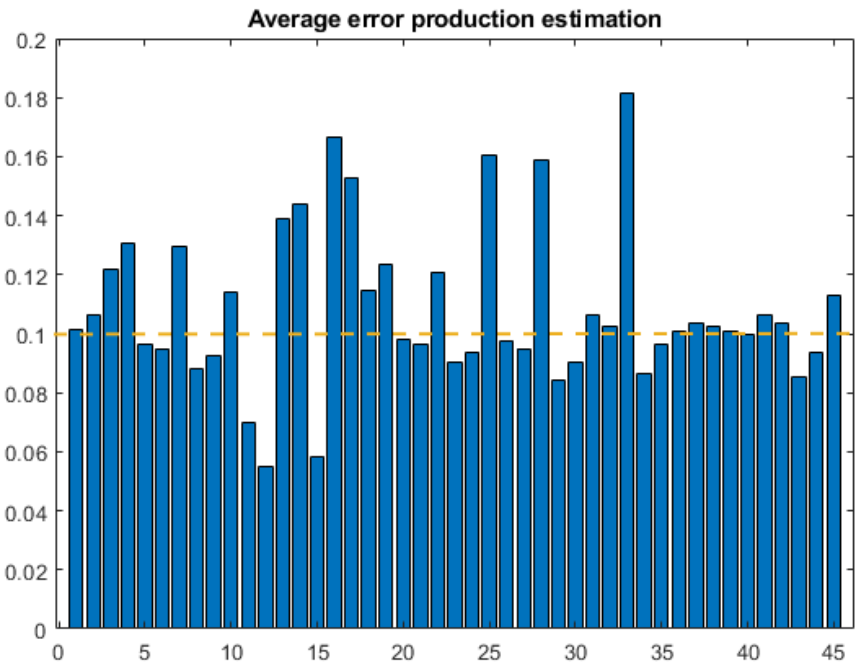

Figure 22 - Average error of the percentage of the production for all N_c combinations.

The results for all computed datasets are shown in Figure 22. The figure highlights that in most cases the estimation error is under 10% providing indication that the response of all plants is similar enough to use some measurements to estimate the others.

Of course, this is a very preliminary study that would require more data for a more robust validation and that is impaired by the lack of available data from the literature.

5.8 Estimation of the initial conditions

Although the previous subsection showed that the assumption on the model is reasonable, its prediction relies on an estimation of the initial conditions.

Therefore, as anticipated in Section 4.3, we propose that the estimation of the initial conditions can be done quite reliably by resorting to the relation $\frac{Yield}{TCSA}$ between the plants of a same plot. In this ratio $Yield$ represents the total yield produced by a plant in terms of weight and *TCSA* is the total trunk section sectional area measured combining all branches at 20 cm from the ground, as defined in Section 4.3.

The insight is that, according to empirical agronomical practice, in relatively small areas of the same plant variety and under the same conditions and treatment, this relation tends to be similar for all plants. The data collected in PANTHEON (see Table 16) shows that indeed this relation is quite reliable showing a mean of 0.0137 and a standard deviation of ± 0.001.

This fact allows us to state that, for each couple of plants i and j insisting on the same plot and subject to the same treatments, it holds

$$YE_i = \frac{Yield_i}{TCSA_i} \approx \frac{Yield_j}{TCSA_j} = YE_j \tag{12}$$

where YE_i denotes the yield efficiency of plant i. Note that the conversion between the $Yield$ (which is the most important parameter for production estimation) and the number of nuts (which is the quantity measured on the field during the growing season) is

$$Yield_i = \varpi_i n_i, \tag{13}$$

being ϖ_i the average weight of the in-shell nuts in plant i and n_i the number of nuts in this plant. Note that, for all adult plants of a same cultivar the average weigh ϖ_i is a constant that can be retrieved from the literature.

The idea of estimating the initial conditions of unmeasured trees using the $TCSA$ is very realistic in the context of PANTHEON where each plant tree geometry reconstruction is available and thus a reliable $TCSA$ for each plant can be measured.

To show how the computation of the average YE made on a a small percentage of trees can be reliably extended to all the trees, we carried out the following validation procedure using again 20% of the plants to compute YE and the rest to validate:

1. We compute the average yield efficiency YE using the data fromtwo plants;

2. We compute the initial conditions as $\hat{x}_i \quad (t_0) = YE\frac{TCSA_i}{\varpi_i}$ for the other eight plants;
3. We compare the estimated initial conditions $\hat{x}_i(t_0)$ with the measurement $x_i(t_0)$

$$e_i = \left|\frac{x_i(t_0) - \hat{x}_i(t_0)}{x_i(t_0)}\right| \tag{14}$$

   where i corresponds to the index of the selected plant;
4. We compute the average error of all plants

$$error = \frac{\sum e_i}{N-2}. \tag{15}$$

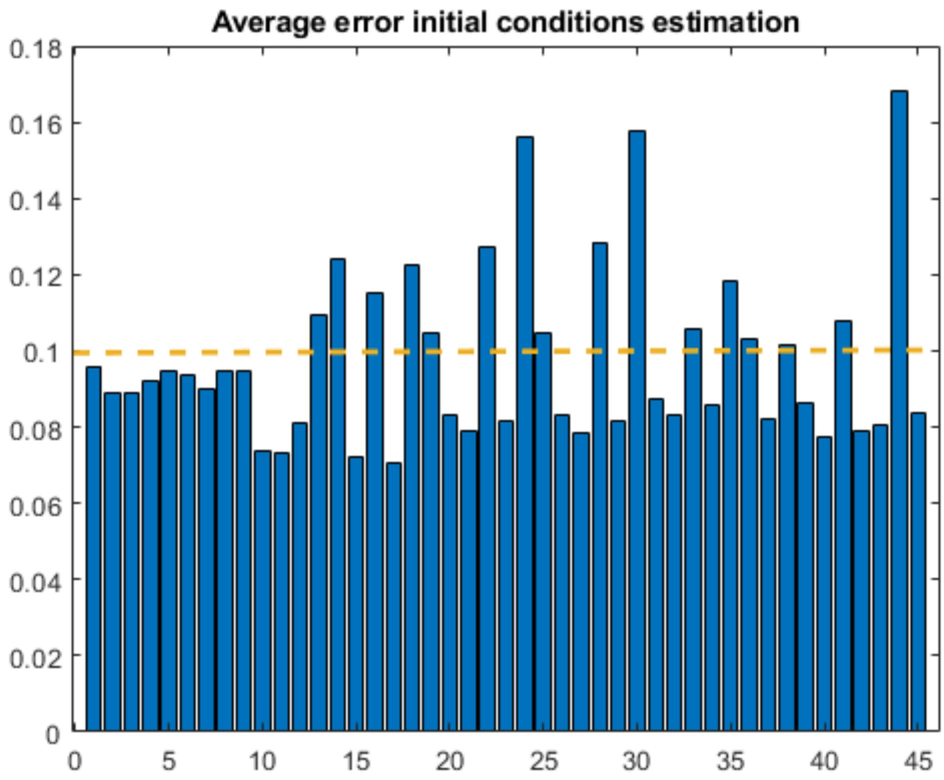

Figure 23 - Average error in the estimation of the initial conditions.

Figure 23 shows the average percentage error in the estimation of the initial conditions based on the efficiency relationship. It can be observed that this error is mostly under 10%, providing an average for all the datasets of 9.76%. These are promising results given the fact that only two plants are used to estimate this correlation factor and the initial conditions.

The advantage of this approach is that it can be applied to the manual counting of certain trees, as the one shown in previous sections and used for the validation, or to automatic fruit counting methods as the protocol currently developed in Task 4.8 of PANTHEON. It must be noted that in this subsection we assumed that only a small percentage of trees could be initially measured. However, thanks to the advances in the

Fruit detection task of PANTHEON the number of trees measured at the beginning of the growing season can be very large or even constitute the whole orchard. In that case, the measurements obtained from this automatic counting can be directly used as input for the model developed in Section 5.6 or combined with the information coming from manual measurements and Yield efficiency (Section 5.8) to improve the accuracy of this initial estimation.

6. Proposed protocol

The previous section proposed a reasonable nonlinear model to describe the loss of hazelnuts along a growing season, which thus allows to predict the evolution of the number of nuts based on the initial conditions. We also showed that the initial conditions of an entire sub-plot can be reliably estimated by using the fact that the ratio between the Yield of the plant and its total trunk section sectional area, measured combining all branches at 20 cm from the ground, is constant with a good approximation.

Based on this information, we propose a protocol which aims at exploiting the presented properties to obtain a good prediction and update of the production estimation.

The protocol can be defined as follows:

SETUP -

1. **Subdivision into areas.** The plot is divided in areas, see Figure 24, that can be considered homogeneous in terms of conditions (soil nature, exposition, type of irrigation, planting scheme) and treatment received by the plants. The more homogeneous these areas are, the better for the estimation.

13 $\kappa_{13}(t)$	14 $\kappa_{14}(t)$	15 $\kappa_{15}(t)$	16 $\kappa_{16}(t)$
9 $\kappa_{9}(t)$	10 $\kappa_{9}(t)$	11 $\kappa_{11}(t)$	12 $\kappa_{12}(t)$
5 $\kappa_{5}(t)$	6 $\kappa_{6}(t)$	7 $\kappa_{7}(t)$	8 $\kappa_{8}(t)$
1 $\kappa_{1}(t)$	2 $\kappa_{2}(t)$	3 $\kappa_{3}(t)$	4 $\kappa_{4}(t)$

Figure 24 - Example of a possible division on the orchard.

We assume that each area follows a similar evolution based on a common losses rate such as

$$\dot{x}_i(t) = -(\hat{\kappa}_j + \tilde{\kappa}_j(t))x_i(t) \quad i \in N_j, \qquad j = 1, \dots, N_a \tag{16}$$

where $x_i(t)$ is the number of nuts at the plant i at time t, $\hat{\kappa}_j$ is the losses rate of the area j, $N_j$is the set of trees belonging to the area j and N_a is the total number of areas in which the orchard has been divided.

1. **Identification of the rate $\hat{\boldsymbol{\kappa}}_j$ for each subarea** based on several measurements of different plants during at least one growing season.

ESTIMATION PROTOCOL DURING THE GROWING SEASON

1. Use of the initial measurements of some plants in the different subareas to estimate a given efficiency to **compute the initial conditions of all plants**.

2. Monitoring of the **production evolution based on the identified model and possible punctual measurements**. The use of a model to predict the evolution of the production allows us to obtain a reasonable estimation of the final production, but this estimation can be improved using punctual measurements of some plants at different instants of the growing season. To do so we propose the use of an Extended Kalman filter with intermittent observations [28, 29]. The approach proposed here is a nonlinear extension of the Kalman filter with intermittent observation that was used within PANTHEON for water stress [30].

   As a first step, we define the following discretized version of the losses model of the plant i based on a sampling period T_s

$$x_i((k+1)T_s) = (\hat{\kappa}_{T_s,j} + \tilde{\kappa}_{T_s,j}(kT_s))x_i(kT_s)\ i \in N_j, \qquad j = 1, \dots, N_a\ \ \forall k \in \mathbb{N} \tag{17}$$

   The system for the whole orchard can be rewritten in compact form as

$$x_{k+1} = f(x_k, w_k) \tag{18}$$

   where

$$x_k = x(kT_s) = \begin{bmatrix} x_1(kT_s) \\ \vdots \\ x_N(kT_s) \end{bmatrix} \tag{19}$$

   and where to use a standard notation within the Kalman Filtering community we denote the disturbance $\tilde{\kappa}_{T_s,j}(kT_s)$ as $w_{j,k}$ with

$$w_k = \begin{bmatrix} w_{1,k} \\ \vdots \\ w_{N_a,k} \end{bmatrix} \tag{20}$$

   Regarding the measurements, as only some measurements will be performed at each instant, we can define the availability of measurement of the$i$-th plant with the binary variable $\gamma_{i,k}$,i.e., $\gamma_{i,k} = 1$ if the measurements on the$i$-th plant are available at time k, and $\gamma_{i,k} = 0$ otherwise.

This provides the following expression for the measurements at time k in the orchard

$$y_k = \Gamma_k(x_k + v_k) \tag{21}$$

where Γ_k is a selection matrix defined as

$$\Gamma_k = [e_i^T]_{i|\gamma_{i,k=1}} \tag{22}$$

where e_i is the i-th vector of the canonical basis and$v_k \sim (0, R)$ denotes white noise with a variance matrix R.

Therefore, the number of nuts can be estimated by means of an intermittent Extended Kalman Filter, where the prediction step is:

$$\hat{x}_{k|k-1} = f(x_{k-1}, w_{k-1}) \tag{23}$$

$$P_{k|k-1} = F_{k-1}P_{k|k-1}F_{k-1}^T + L_{k-1}Q_{k-1}L_{k-1}^T \tag{24}$$

with

$$F_{k-1} = \frac{\partial f}{\partial x}|_{x_{k-1}} \tag{25}$$

$$L_{k-1} = \frac{\partial f}{\partial w}|_{w_{k-1}} \tag{26}$$

And the correction step:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - \Gamma_k x_k) \tag{27}$$

$$K_k = P_{k|k-1}\Gamma_k^T(P_{k|k-1} + R)\Gamma_k^T \tag{28}$$

$$P_{k|k} = P_{k|k-1} - K_k\Gamma_k P_{k|k-1} \tag{29}$$

where Q_{k-1} is a design parameter, see [30] for more details about its design.

The current limited availability of data does not allow to validate the effectiveness of the method and will be subject of future works.

References:

[1] Boccacci, M. Aramini, N. Valentini, L. Bacchetta, M. Rovira, P. Drogoudi, A.P. Silva, A. Solar, F. Calizzano, V. Erdoğan, L.F. Ciarmiello, V. Cristofori, C. Contessa, I.J. Ferreira, F.P. Marra and R. Botta, “Molecular and morphological diversity of on-farm hazelnut (Corylus avellana L.) landraces from southern Europe and their role in the origin and diffusion of cultivated germplasm”, in *Tree Genetics & Genomics*, 2013, 9: 1465-1480. DOI: 10.1007/s11295-013-0651-7.

[2] V. Cristofori, S. Speranza, and C. Silvestri, “Developing hazelnuts as a sustainable and industrialcrop,” in *BurleighDodds Series in Agricultural Science*, 2019, 465–504, DOI: 10.19103/AS.2018.0042.21.

[3] INC, 2021. https://www.nutfruit.org/industry/statistics. International Nut and Dried Fruit Council. Carrer de la FruitaSeca, 4 - PolígonTecnoparc - 43204 REUS, Spain.

[4] A. Frary, S.C. Ozturk, H.I. Balik, S.K. Balik, G. Kizilci, S. Doganlar and A. Frary, “Association mapping of agro-morphological traits in European hazelnut (Corylus avellana)”, in *Euphytica*, 215, 21. DOI: 10.1007/s10681-019-2352-2.

[5] A.N. Azarenko, R.L. McCluskey and W.C. Chambers, “Does canopy management help to alleviate biennial bearing in 'Ennis' and 'Montebello' hazelnut trees in Oregon?”, in *Acta Hortic,* 2005, 686: 237-242. DOI: 10.17660/ActaHortic.2005.686.31.

[6] C.R. Hampson, A.N. Azarenko and J.R. Potter, “Photosynthetic rate, flowering, and yield component alteration in hazelnut in response to different light environments”, in *J. Amer. Soc. Hort. Sci*, 1996, 121(6):1103-1111. DOI: 10.21273/JASHS.121.6.1103.

[7] S.A. Mehlenbacher, “Hazelnuts (Corylus)”, in *Acta Hort*, 1991, 290: 791–836. DOI: 10.17660/ActaHortic.1991.290.18.

[8] C. Silvestri, L. Bacchetta, A. Bellincontro and V. Cristofori, “Advances in cultivar choice, hazelnut orchard management and nuts storage for enhancing product quality and safety: an overview”, in *Journal of the Science of Food and Agriculture*, 2021, 101(1): 27-43. DOI: 10.1002/jsfa.10557.

[9] E. Germain, “The reproduction of hazelnut (Corylus avellana L.): a review”, in *Acta Hortic*, 1991, 351: 195-209.DOI: 10.17660/ActaHortic.1994.351.19.

[10] C. Bignami and S. Natali, “Influence of irrigation on the growth and production of young hazelnuts”, in *Acta Hortic*, 1991, 445: 247-262. DOI: 10.17660/ActaHortic.1997.445.34.

[11] M.M. Thompson, H.B. Lagerstedt and S.A. Mehlenbacher, “Hazelnuts, in Fruit Breeding: Nuts”, 1996 Vol. 3, ed. by Janick J and Moore JN. Wiley, New York, NY, 125-184.

[12] M. Fontana, M. Somenzi and A. Tesio, “Cultivation, harvest and postharvest aspects that influence quality and organoleptic properties of hazelnut production and related final products”, in *Acta Hortic.*, 2014, 1052: 311–314, DOI: 10.17660/ActaHortic.2014.1052.43.

[13] L. Bacchetta, M. Rovira, C. Tronci, M. Aramini, P. Drogoudi and A.P. Silva, “A multidisciplinary approach to enhance the conservation and use of hazelnut genetic resources”, in *Genet. Resour. Crop Evol*, 2015, 62: 649–663. DOI: 10.1007/s10722-014-0173-7.

[14] A. Gongal, S. Amatya, M. Karkee, Q. Zhang and k. Lewis, “Sensors and systems for fruit detection and localization: A review”, in *Computers and Electronics in Agriculture*, 2015,116:8-19. DOI: 10.1016/j.compag.2015.05.021.

[15] L. Fu, F. Gao, J. Wu, R. Li, M. Karkee and Q. Zhang, “Application of consumer RGB-D cameras for fruit detection and localization in field: A critical review”. In *Computers and Electronics in Agriculture*, 2015, 177. DOI:10.1016/j.compag.2020.105687.

[16] P. Maheswari, P. Raja, O.E. Apolo-Apolo and M. Perez Ruiz, “Intelligent Fruit Yield estimation for Orchards Uning Deep Learning Based Semantic Segmentation Techniques – A review”, 2021, in *Frontiers in Plant Science*, 12:64328. DOI:10.3389/fpls.2021.684328.

[17] N. Häni, P. Roy and V. Isler, “A comparative study of fruit detection and counting methods for yield mapping in apple orchards”, in *J. F. Robot*, 2020, 37: 263–282. DOI:10.1002/rob.21902.

[18] Y. Ge, Y. Xiong, G.L. Tenorio and P.J. From, “Fruit localization and environment perception for strawberry harvesting robots, in *IEEE Access,* 2019, 7: 147642–147652. DOI:10.1109/ACCESS.2019.2946369.

[19] B. Arad, J. Balendonck, R. Barth, O. Ben-Shahar, Y. Edan, T. Hellström, J. Hemming, P. Kurtser, O. Ringdahl, T. Tielen and B. Van Tuijl, “Development of a sweet pepper harvesting robot”, in *J. F. Robot*, 2020, 1–13. DOI:10.1002/rob.21937.

[20] V. Cristofori, S. Ferramondo, G. Bertazza and C. Bignami, "Nut and kernel traits and chemical composition of hazelnut (Corylus avellana L.) cultivars", in *Journal of the Science of Food and Agriculture*, 88 (6): 1091-1098. DOI: 10.1002/jsfa.3203.

[21]X. Liu *et al.*, “Robust Fruit Counting: Combining Deep Learning, Tracking, and Structure from Motion,” in*2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Madrid, Oct. 2018, pp. 1045–1052. doi: 10.1109/IROS.2018.8594239.

[22]N. Häni, P. Roy, and V. Isler, “A comparative study of fruit detection and counting methods for yield mapping in apple orchards,” *J. Field Robotics*, vol. 37, no. 2, pp. 263–282, Mar. 2020, doi: 10.1002/rob.21902.

[23]M. Rahnemoonfar and C. Sheppard, “Deep Count: Fruit Counting Based on Deep Simulated Learning,” *Sensors*, vol. 17, no. 4, p. 905, Apr. 2017, doi: 10.3390/s17040905.

[24]C. Bignami, V. Cristofori, and G. Bertazza, “Effects of water availability on hazelnut yield and seed composition during fruit growth,” *Acta Horticulturae*, no. 922, pp. 333–340, Dec. 2011, doi: 10.17660/ActaHortic.2011.922.43.

[25]V. Cristofori, R. Muleo, C. Bignami, and E. Rugini, “Long term evaluation of hazelnut response to drip irrigation,” *Acta Horticulturae*, no. 1052, pp. 179–185, Sep. 2014, doi: 10.17660/ActaHortic.2014.1052.23.

[26]Y. Song, C. A. Glasbey, G. W. Horgan, G. Polder, J. A. Dieleman, and G. W. A. M. van der Heijden, “Automatic fruit recognition and counting from multiple images,” Biosyst. Eng., vol. 118, pp. 203–215, Feb. 2014, doi: 10.1016/j.biosystemseng.2013.12.008.

[27]M. Afonso *et al.*, “Tomato Fruit Detection and Counting in Greenhouses Using Deep Learning,” *Front. Plant Sci.*, vol. 11, p. 1759, 2020, doi: 10.3389/fpls.2020.571299.

[28]B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, “Kalman Filtering With Intermittent Observations,” *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004, doi: 10.1109/TAC.2004.834121.

[29]E. Garone, B. Sinopoli, A. Goldsmith, and A. Casavola, “LQG Control for MIMO Systems Over Multiple Erasure Channels With Perfect Acknowledgment,” *IEEE Trans. Automat. Contr.*, vol. 57, no. 2, pp. 450–456, Feb. 2012, doi: 10.1109/TAC.2011.2167789.

[30]N. Bono Rossello, R. Fabrizio Carpio, A. Gasparri, and E. Garone, “A novel Observer-based Architecture for Water Management in Large-Scale (Hazelnut) Orchards,” *IFAC-PapersOnLine*, vol. 52, no. 30, pp. 62–69, 2019, doi: 10.1016/j.ifacol.2019.12.498.