Outlining the mission profile of agricultural tractors through CAN-BUS data analytics

(Article begins on next page)

03 May 2024

OUTLINING THE MISSION PROFILE OF AGRICULTURAL TRACTORS THROUGH CAN-BUS DATA ANALYTICS

1 Michele Mattetti[a*], Mirko Maraldi[a], Nicola Lenzini[b], Stefano Fiorati[b], Eugenio Sereni[c],
2 Giovanni Molari[a]
3 [a] Department of Agricultural and Food Sciences, Alma Mater Studiorum – University of Bologna, viale G.
4 Fanin, 50, 40127, Bologna, Italy
5 [b] CNH Industrial – Tractor Innovation Engineering, viale delle Nazioni 55, 41100, Modena, Italy
6 [c], viale delle Nazioni 55, 41100, Modena, Italy
7 [*] tel. +39 051 2096174, fax +39 051 2096178, email: _michele.mattetti@unibo.it_
8

9 ## Abstract

10 Tractor manufacturers need to know how farmers use their agricultural tractors for an

11 optimal machine design. Tractor usage is not easy to assess due to the large variability of field

12 operations. However, modern tractors embed sensors integrated into the CAN-BUS network

13 and their data is accessible through the ISO 11783 protocol. Even though this technology has

14 been available for a long time, the use of CAN-BUS data for outlining the tractor usage is still

15 limited, because a proper post-processing method is lacking. This study aimed to present a novel

16 classification scheme of CAN-BUS data which permits to outline the tractor usage. On a tractor,

17 a CAN-BUS data logger and a GNSS receiver were installed, and real-world data were recorded

18 for 579 hours. Thus, data was obtained in the most realistic condition. Tractor positions were

19 classified using GIS layers while operating conditions were classified depending on the usage

20 of the tractor's subsystems. The method highlights that showed to be able to detect the 97% of

21 the logged data and that the tractor operated on the field in working, on idle, and moving duties

22 for 65%, 18% and 16% of the time, respectively. The method allows a far more precise outline

23 of tractor usage opening opportunities to obtain large benefits from massively collected CAN-

24 BUS data.

25

Table 1: Nomenclature		
D_{OFF}	Distance travelled in out-of-work state	[m]
D_{ON}	Distance travelled in in-work state	[m]
$GNSS$	Global navigation satellite system	[$-$]
n_e	Engine speed	[rpm]
$n_{PTO,f}$	front PTO speed	[rpm]
$n_{PTO,r}$	rear PTO speed	[rpm]
P_e	Engine power	[kW]
RWI	Rear hitch in work indication	[$-$]
T_e	Actual engine-percent torque	[%]
T_f	Nominal friction-percent torque	[%]
T_H	Headland turns duration	[s]
T_r	Engine reference torque	[Nm]
V_t	Tractor ground speed	[$km\ h^{-1}$]

27

Introduction

28

29   The typical usage of a tractor model is described through its mission profile, which is a

30 synthetic description of a tractor use. Mission profiles report the factors that influence the

31 operational durability of tractor components (Johannesson & Speckert, 2013). A mission profile

32 may report:

33   • the typical tractor service life;

34   • the typical contribution of each operating modes on the service life (e.g. ploughing, on-

35   road, and off-road transportations, etc);

36   • how each component is typically used (e.g. the input power and speed on gearboxes,

37   vehicle ground speed, etc).

38

39   Mission profiles are essential for a proper design/selection of tractor components (Sehab et

40 al., 2011) or for designing durable and reliable machines with an optimal balance between

41 under-designs and over-designs (Plaskitt & Musiol, 2002).

42   Mission profiles of agricultural tractors include several factors, and these factors make their

43   mission profile estimations much more challenging than that of road vehicles. Indeed, a road

44   vehicle may travel on only three different types of roads (e.g. highway, city, country road) and

45   two types of load levels (the driver alone and with 4 passengers and luggage) (Marchesani et

46   al., 1992); on the other hand, row crop tractors may be used for a larger variety of uses (e.g.

47   road transportation, soil preparation, sowing, haying, etc.), and each can be accomplished at

48   different load levels due to the different ground conditions.

49   To estimate a mission profile, tractor usage from a sample of farmers is necessary. This is

50   typically carried out through surveys aimed at obtaining information about the farm size, yearly

51   usage of tractors, list of farming operations carried out in the farm, and how each operation is

52   performed (Mattetti et al., 2012). This approach is usually adopted for its easiness in obtaining

53   data from large samples, but the obtained information is biased toward subjective judgements,

54   which could lead to unreliable mission profiles. A different approach consists in installing

55   switchboards inside tractor cabs (Paraforos et al., 2017) or through specific smartphone apps

56   (e.g. 365Farmnet) which allow farmers to assign the task they are accomplishing with the

57   tractor. However, these approaches require a manual effort of farmers, who may forget the task

58   assignment.

59   In modern tractors, the operating parameters of all the tractor subsystems can be monitored

60   using CAN-BUS technologies together with SAE J1939, and ISO 11783 protocols (ISO, 2012;

61   Molari et al., 2013; SAE, 2006). In previous studies, CAN-BUS messages were successfully

62   used to outline the usage of specific tractor components (Mattetti et al., 2019), to determine

63   field efficiencies of agricultural machinery (Pitla et al., 2014, 2016) or to monitor specific

64   tractor operating modes (Molari et al., 2019). The best approach for a proper mission profiling

65   would be recording and analysing real-world CAN-BUS data of a large fleet of tractors. In this

66   way, the recording process would not interfere with farming activities and data would be

67 recorded in the most realistic conditions. But then, the environment where vehicles operate is

68 unknown, and advanced classification approaches are essential for a reliable estimation of the

69 mission profile (Fugiglando et al., 2019). Data classification is the process of grouping together

70 portions of signals related to the same work state (Zhang et al., 2017). A sort of data

71 classification is already provided by telemetric data service supplied by each tractor

72 manufacturer (New Holland MyPLM Connect, John Deere JDLink). In these tools, the work

73 states are defined on the basis of simple threshold-rules, in other words a work state is defined

74 when any signal exceeds a threshold. For example, in telemetric data services, tractors are:

75 • on fieldwork state when the three-point linkage is down, but farmers may drive bare

76 tractors on the field with the three-point linkage in the down position.

77 • on moving state when its speed exceeds a threshold specified by the driver (i.e.

78 25 km h^{-1}), but the proper value may change in function of the road state (i.e. presence

79 of speed humps, road damages, etc).

80 For these reasons, this approach is far too simplistic and data misclassifications are not

81 infrequent. More advanced rule-based algorithms were proposed for specific tractor operations

82 (Ettl et al., 2018), and forage harvesters (Harmon et al., 2018; Zhang et al., 2017). Rule-based

83 algorithms require knowledge from the experts and a reasonable amount of effort to design and

84 implement effective rules for real-world data. Indeed, in real-world conditions, tractor

85 operativity may change according to a variety of operating conditions (in terms of soil,

86 implement type, driving style). Thus, for outlining the tractor usage, a robust and flexible

87 method must be developed, which can deal with the variability induced by the variability of

88 tractor manoeuvres (Mattetti, Molari, et al., 2017).

89   This article aimed to develop a robust and automatic classification scheme able to identify

90 the tractor mission profile using real-world CAN-BUS, and trajectory data.

Materials and methods

Data acquisition

The analysis was applied to a New Holland T7 tractor (CNH Industrial N.V., Amsterdam, NL) whose specifications are reported in Table 2. This was chosen because tractors of this class are rich in terms of embedded sensors allowing for comprehensive monitoring of the activity of the different embedded subsystems.

Table 2 – Specifications of the tractor used in this study.		
Maximum engine power	(kW)	198
Engine displacement	(m^3)	6.728
Number of cylinders	$(-)$	6
Engine tier	$(-)$	4B
Transmission	$(-)$	Continuously variable transmission
Number of auxiliary hydraulic valves	$(-)$	4
Three-point linkage	$(-)$	Rear
PTO	$(-)$	Front and rear

The tractor was in use between the June 2018 and October 2019 by 5 professional drivers with more than twenty years of experience. The tractor was used in the Agricultural Farm of the University of Bologna. The size of the farm is 500 ha, where 67%, 10% and 23% of the land are devoted to cereals, orchards, and haying, respectively. The farm is distributed in three different units (i.e. areas where tractors are stored overnight) located in three different towns; the farther locations are 35 km apart. In this farm, this tractor is mostly used for transportation tasks and primary and secondary tillage tasks.

A stand-alone CAN-BUS data-logger optimised by CNH Industrial was installed on the tractor. The data-logger was set up to automatically record all the CAN-BUS messages anytime the tractor engine was turned on so that the recording process did not interfere with farming activities. In particular, the CAN-BUS data logger is equipped of two separated CAN-BUS channels compatible with the standards: SAE J1939-14 (SAE, 2016a) and SAE J1939-15 (SAE,

110   2018b). The data-logger embeds a BLE (Bluetooth low energy) scanner which scans the BLE

111   beacons in its surroundings (up to 10 meters) every second. Commercial BLE beacons

112   (Mokosmart M1 Beacon, Shenzen, China) were attached to the implements available at the

113   farm as suggested in other studies (Calcante & Mazzetto, 2014) (Fig. 1).

114
115   *Fig. 1: A trailer used during the project. Enclosed in the yellow circle, a BLE beacon that permitted the*
116   *identification of the trailer.*

117   The BLE scanner records the identifiers of the detected BLE beacons to record implement

118   connected to the tractor. Moreover, a Garmin Dash Cam 55 (Garmin Ltd., Olathe, KS, USA)

119   was installed on the windshield of the tractor to document the tractor activity in order to ensure

120   the reliability of the classification scheme.

121   For the purpose of this study, only signals with the following Suspect Parameter Numbers

122   (SPNs) and Parameter Group Numbers (PGNs) (ISO, 2012; SAE, 2013) were used for the

123   analysis:

124   • SPN 544 and PGN 65251: "*Engine Reference Torque*" that reports the torque as a

125   percent of Engine Reference Torque (SPN 544 and PGN 65251) and it is denoted as T_r

126   in the following.

127   •   SPN 513 and PGN 61444: "*Actual Engine - Percent Torque*" that reports the torque as a

128   percent of T_r , and it is denoted as T_e in the following.

129   •   SPN 513 and PGN 5398: "*Nominal friction-percent torque*" that reports the frictional

130   and thermodynamic loss of the engine itself, pumping torque loss and the losses of fuel,

131   oil and cooling pumps as a percent of T_r, and it is denoted as T_f in the following.

132   •   SPN 190 and PGN 61444: "*Engine Speed*" that reports the revolution speed of the engine

133   crankshaft, and it is denoted as n_e in the following.

134   •   SPN 1883 and PGN 65090: "*Rear PTO output shaft speed*", that reports the speed of the

135   rear PTO.

136   •   SPN 1882 and PGN 65090: "*Front PTO output shaft speed*" that reports the speed of the

137   front PTO.

138   •   SPN 1877 and PGN 65093: "Rear hitch in-work indication " that reports the rear hitch

139   is positioned below (in-work) or above (out-of-work) 85% of the position of the rear

140   three-point linkage (SPN 1873 and PGN 65093). This signal is denoted as RWI in the

141   following.

142   Moreover, a GNSS (global navigation satellite system) receiver with an update rate of 10

143 Hz, with no differential correction, and with a claimed accuracy of 2.5 m (in terms of circular

144 error probable) (IPESpeed, IPETronik GmbH, Baden Baden, Germany) was installed in the

145 tractor to monitor its position and its tractor ground speed (V_t).

146 *Data analysis*

147   All the signals were interpolated at 10 Hz using a cubic spline so that the sampling rate of

148 all the signals was the same. From the recorded data, the delivered engine power (P_e) was

149 calculated as follows:

150
$$P_e = T_r \cdot \frac{T_e - T_f}{100} \cdot n_e \frac{2\pi}{60}$$

151   All the portions of the recorded signals acquired when the tractor position was not logged

152 (because the GNSS receiver did not obtain a strong enough satellite signal, e.g. when the tractor

153 was moved out from an indoor environment) were excluded from the analysis. Tractor positions

154 were classified into three categories:

155   • *road*, anytime the position was closer than 3 m to any road stretch. The 3 m

156   threshold was chosen based on the circular error probability of the GNSS receiver

157   used in this study. For a full automation of the process, this was carried out by

158   checking if there is any intersection point between any road stretch and a circle, with

159   a radius of 3m, centred in the tractor position.

160   • *field*, anytime the position was inside the boundary of any field plot.

161   • *farm*, anytime the position was inside the boundary of any farm unit.

162   For the classification of the tractor position, a shapefile containing the road network, the

163 boundaries of field plots, and the boundaries of the farm units were created. The creation of the

164 shapefile started by downloading the soil use and the road network from the geoportal of the

165 Emilia Romagna region (*Dati preconfezionati — GeoER*, 2019). To this shapefile, the

166 boundaries of three farm units were added.

167   The tractor operating conditions were classified into three categories: idling, moving, and

168 three-point linkage use. Idling condition was defined as the state where the tractor was standing

169 with no use of any PTO for more than 5 s; the duration threshold was added in order to not

170 include portions where the tractor was temporary still during manoeuvring, like reversing the

171 tractor direction at the headlands. *Moving* condition was defined as the state where tractor

172 ground speed was greater than 0 km h^{-1} with no use of the three-point linkage or both PTOs.

173 *Three-point linkage use* was defined as the state where the three-point linkage was used for

174 field operations. This occurs anytime a sequence of a pass, headland, and pass was repeated.

175 *RWI* signal shows a rectangular waveform, but when the tractor operates on field operations,

176 repetitive pulses could be observed (Fig. 2 - right). A method for discerning field operations

177 from implement hitching or machine moving activities was developed and it is described in the

178 following.

179
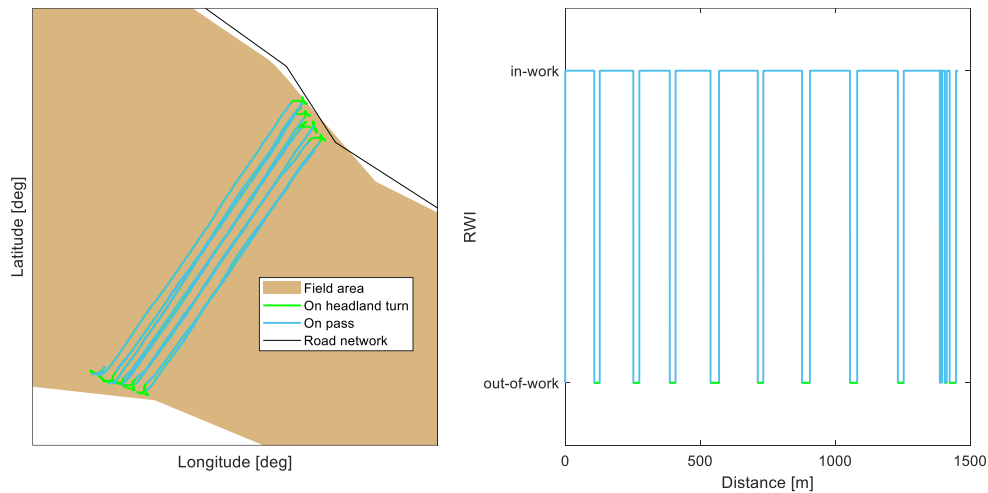
180
181 *Fig. 2: Tractor trajectory (on left), and RWI signal (on right) during a field operation with a plough.*

182 First, the distance travelled by the tractor on each in-work (D_{ON}) and on the out-of-work (D_{OFF})

183 states of *RWI* signal was calculated. For field operations, D_{ON} and D_{OFF} represent the length of

184 passes and headlands, respectively. Both depend on several operating parameters like length of

185 fields, and headland strategy (Paraforos et al., 2018). For field operations, D_{OFF} was usually

186 included in a range between 1 and 70 m. Values of D_{OFF} below the lower bound occurred on

187 implement hitching and values of D_{OFF} above the upper bound occur when the tractor switched

188 from or to a moving operation. The above range was determined using the following approach:

189   • Identification of the portions where the tractor was operating in the field from video

190   data collected with the camera.

191   • Extraction of all the logged signals in this portion.

192   • Calculation of histograms of D_{on} data of 1200 headlands and from it the threshold

193   was set.

194   The algorithm started by calculating the series of D_{ON} and D_{OFF}, and *three-point linkage use*

195   classification occurred in the timespan where D_{OFF} are included in the beforementioned range.

196   When this classification occurs, the high levels of RWI described the passes, while the low

197   levels of RWI described the headlands. Then, the headland duration was calculated as the time

198   elapsed between the falling and rising edge of RWI signal when the tractor was in *three-point*

199   *linkage usage*.

200   The work states were defined based on a combination of the classification of the tractor

201   position and the tractor operating activity (Table 3).

| Table 3 – Rule used for the work states which characterise the tractor use.			
Work states	*Classified tractor position*	*Boolean operation*	*Classified tractor operating activity*
On-road moving	*Road*	*AND*	*Moving*
Off-road moving	*NOT(Road)*	*AND*	*Moving*
Field work	*Field*	*AND*	*three-point linkage use*
Idle@field	*NOT(Road)*	*AND*	*Idle*
Idle@farm	*Farm*	*AND*	*Idle*
Marginal	*Otherwise*		

202

203   The idling on road was included into *marginal* because its contribution is of minor

204   importance to the entire idling (Molari et al., 2019).

205   An example of the classification of the work states is reported in Fig. 3. One can note that

206   in the first portion (in the first 6 min of the time histories) the tractor was classified as on-road

207   moving task; indeed, the tractor was running at around 40 km h^{-1} and P_e was on average low

208   with peaks when high tractor accelerations occurred. On the other hand, the last portion (from

209   13 min) was classified as field work, indeed V_t was lower than 10 km h^{-1} and P_e is close to

210   engine limit.
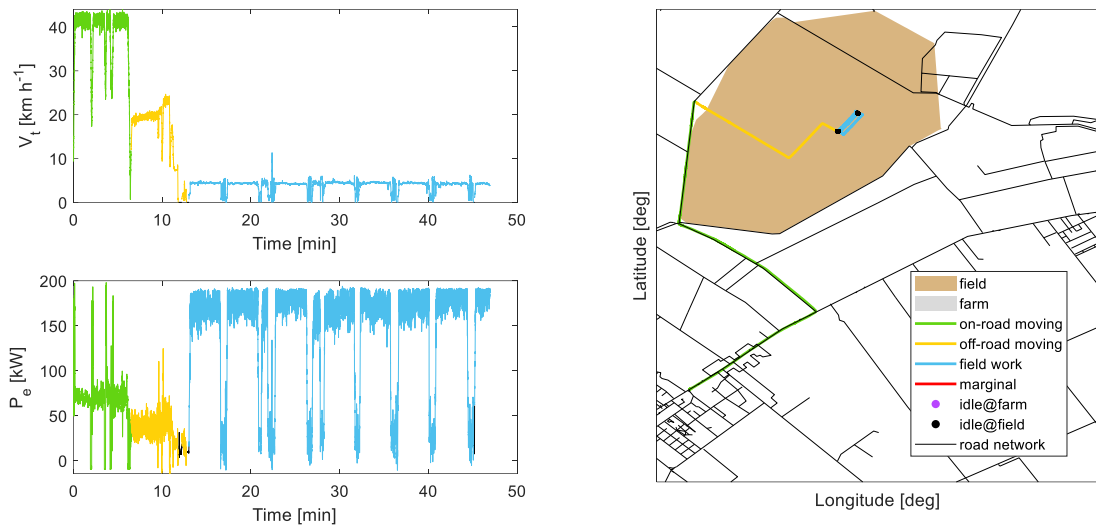
Fig. 3: Classification example of tractor ground speed (on the top left) and engine power (on the bottom left) compared with the tractor trajectory data (on right).

In this article, a task was defined as the portions where neither work states nor hitched implement was changed. For each task, the average values of all signals were calculated; for the subsequent dataset, outliers (i.e. misclassifications) were identified through the confidence ellipse method. This method consists of computing the confidence ellipse between two variables and considering outliers data points falling outside the confidence ellipse (Hodge & Austin, 2004). For drawn implements, the two variables were V_T and P_e (Fig. 4); as the power demand of this type of implements is mostly dependent on the working speed (Mattetti, Varani, et al., 2017), while for PTO-driven implements, the two variables were $n_{PTO,*}$ (* stands for f for front mounted implements and r for rear mounted implements) and P_e, as the power demand of this type of implements is dependent on the speed of the PTO (Balsari et al., 2020). The confidence level was set at 90%. A multivariate approach was necessary, since a low demanding ploughing may not be an outlier if the ground speed is low as well.
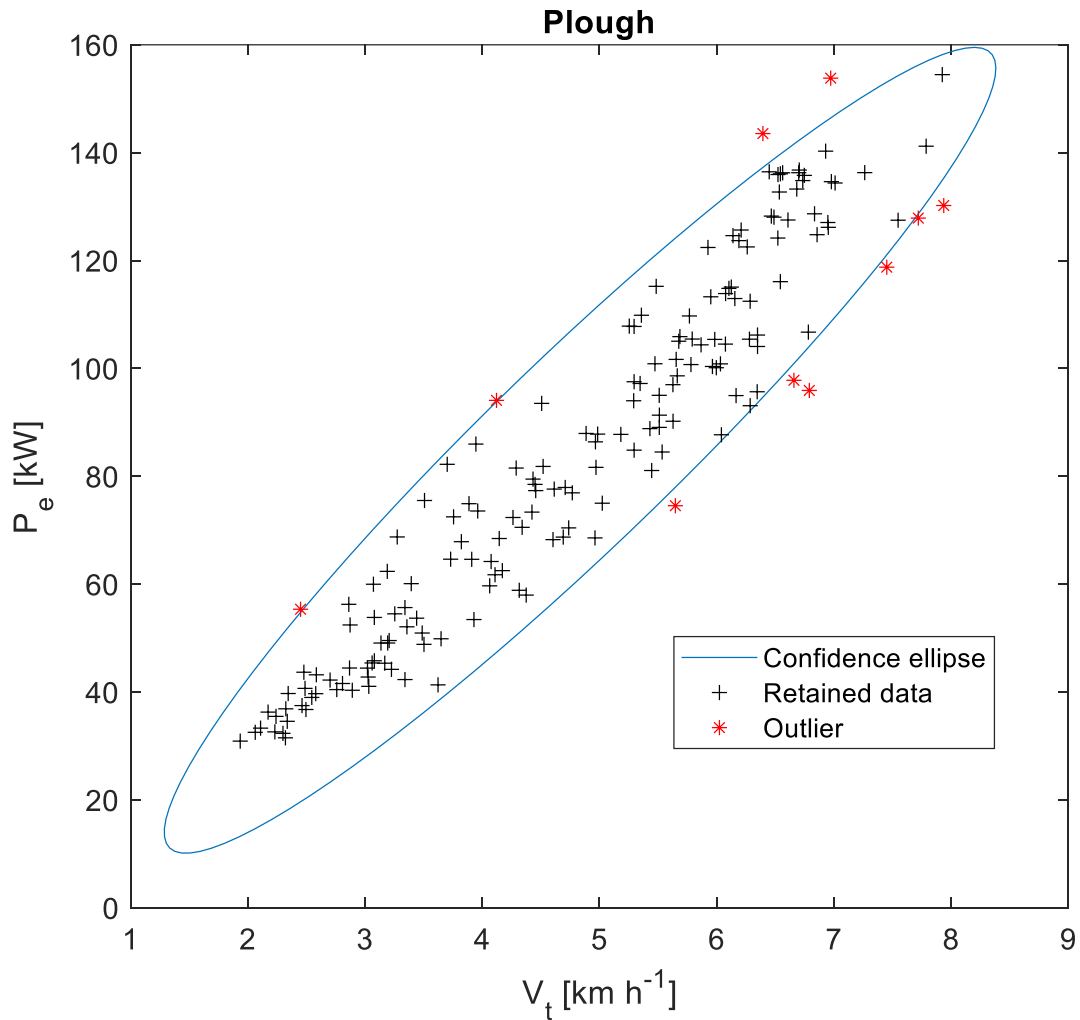
Results and Discussion

The tractor was used for 107 days amounting to 579 hours overall. The tractor was used with 11 implements, but 5 of them were used for 84% of the time (Fig. 5). For 78% of the time, the tractor was used for plough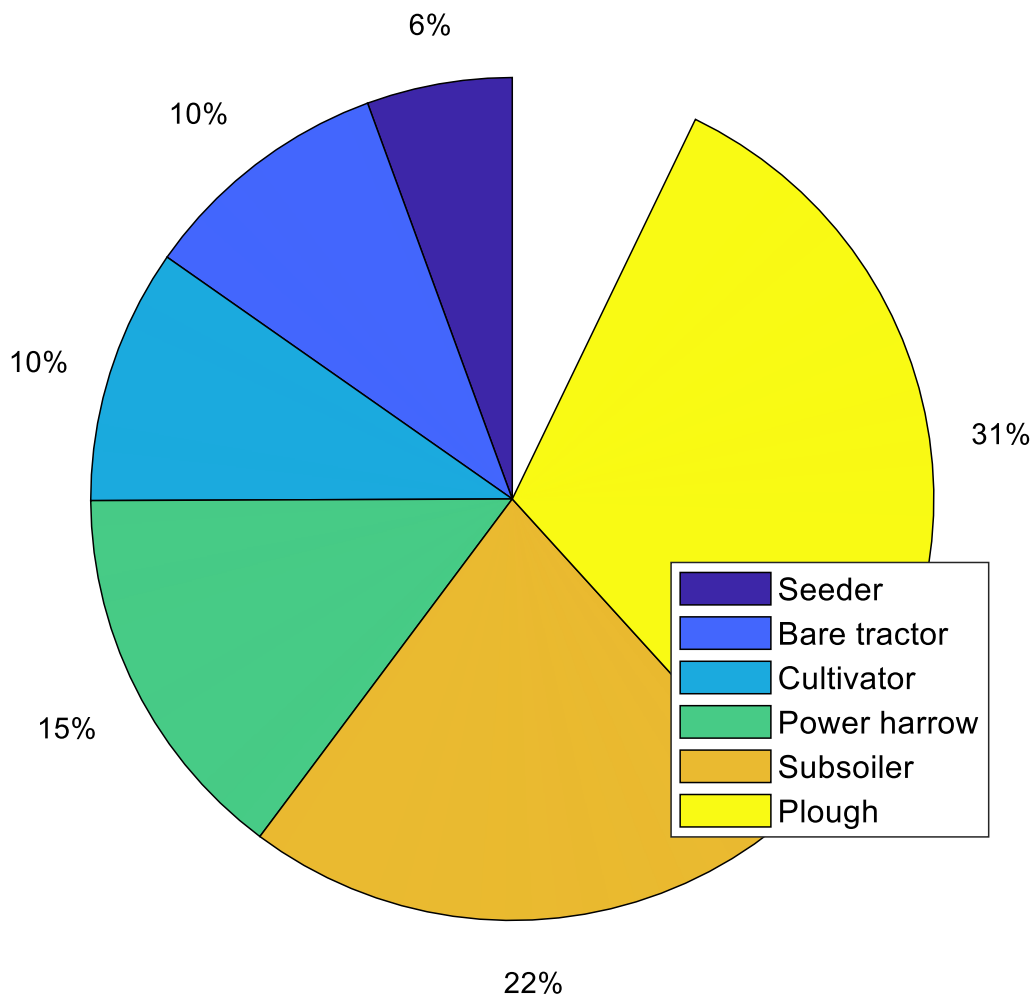ing, subsoiling, harrowing, and cultivating. Thus, the analysis was focused on the data related to those operations for the larger amount of available data. The tractor was used with no implement for 10% of the time, and in this configuration, the tractor was mostly moved from a farm unit to another.

Fig. 5: Pie chart reporting the contribution of each implement on the operating time. In the chart, the implements used for less than 20 hours were not plotted for sake of clarity.

The tractor was used for field work tasks for 65% of the time and 18% of the time for idling activities (Fig. 6). The amount of idling is below the average value reported in the study by Perozzi et al. (2016) where the idling of a large sample of tractors was analysed.
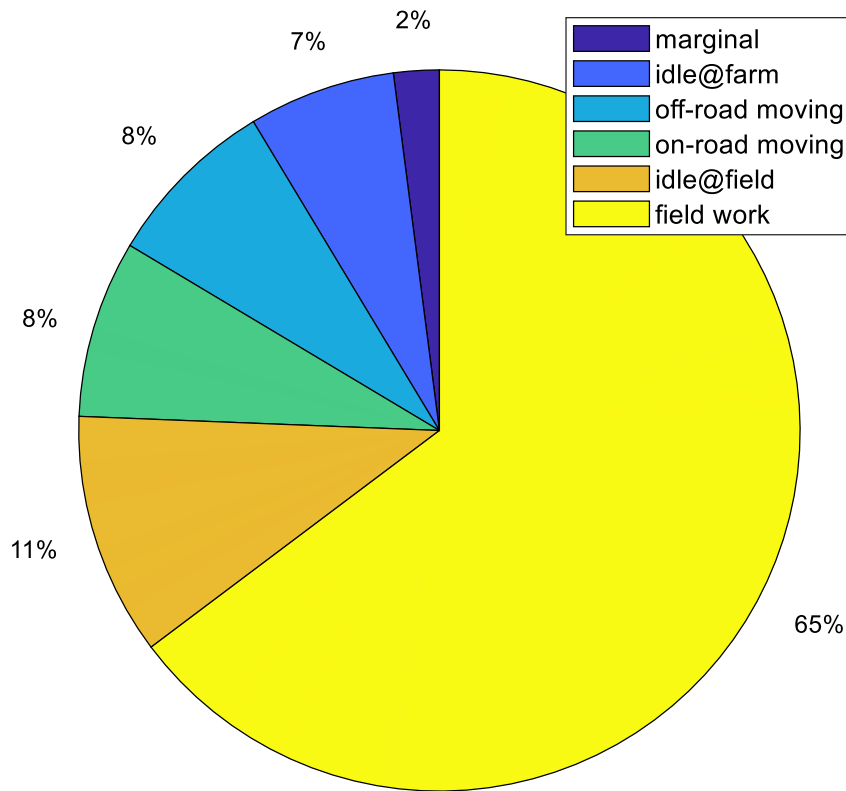
Fig. 6: Time contribution of each work state on the entire tractor activity.

243   The time contribution of *off-road moving* state is lower than that of the *on-roading moving*

244   state. This is because the farm where the tractor was used is spread over a large area, which

245   leads to infrequent but prolonged *on-road moving* states. Indeed, for *off-road moving* work

246   state, the number of tasks are 38% of all the identified tasks and their average duration is 172

247   s. On the other hand, for *on-road moving* work states, the number of tasks are and 11% of all

248   the identified tasks and their average duration is 478 s, respectively. The time contribution of

249   *idle@field* work state is larger than that of *idle@farm* because the amount of idling stops on the

250   field is much more frequent than those at the farm due to the varied source of stops, such as rest

251   stops, driver turnover, checking the performance, removal of crop residual on implements (Hunt

252   & Wilson, 2015). On the other hand, *idle@farm* state mostly occurs at the beginning and the

253   end of the workday, and mostly for machine servicing or adjustment, implement hitching and

254   machine parking (Molari et al., 2019). The sum of the time contributions for the *idle@field* and

field work states provide an insight of the portion of time where the tractor operated for field related activities, and it includes the time for actual work, headlands, field setting, and maintenance at field (Lovarelli et al., 2017). Headlands contribute to 24% of the entire *field work* state and this figure is aligned to that of Ettl et al. (2018). Moreover, the time contribution of the *marginal* is less than 3%, which means that the defined work states can describe most of the tractor operations. In Fig. 7, the relative frequency of headland durations (T_H) for 10065 headlands is reported. T_H range from 3 s up to 230 s and it is strongly dependent on the headland patterns. 50% of the headlands ranged from 20 and 40 s and this result is aligned with that reported in other studies (Ettl et al., 2018; Paraforos et al., 2018).
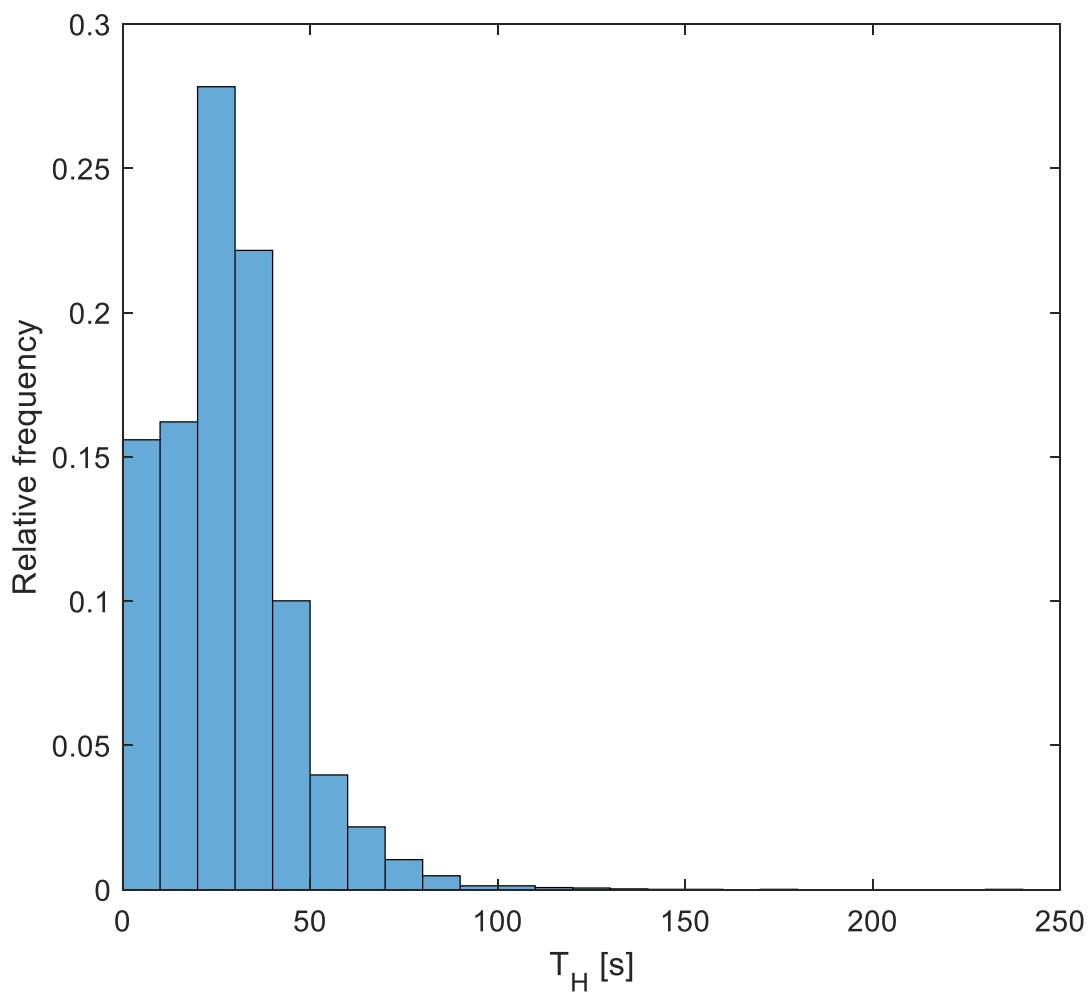
Fig. 7: Relative frequency distribution of the duration of the identified headlands.

266   Headlands shorter than 20 s are not infrequent, and they account for 32% of the headlands.

267   These occur when the tractor worked around the field border where the circuitous turn strips at

268   corner diagonals working pattern is adopted (Hunt & Wilson, 2015). In the most extreme cases,

269   T_H is lower than 10 s, this occurs for an unconventional type of headland pattern. In particular,

270   the farmer tilled two different fields separated by a country road and no turns can be observed

271   in the tractor trajectory (Fig. 8). In the same plot, also headlands longer than 100 s can be

272   observed, which occurred because the overlapping alternation pattern was adopted.
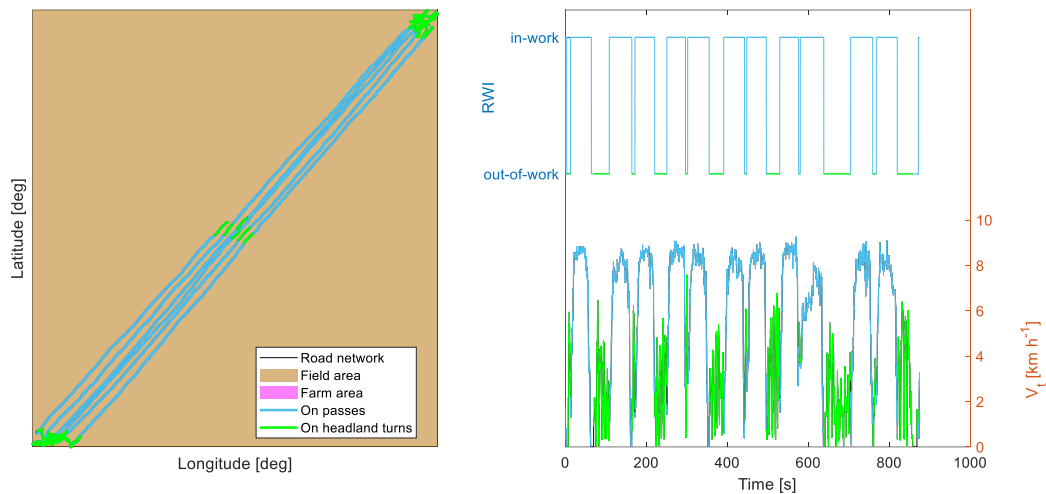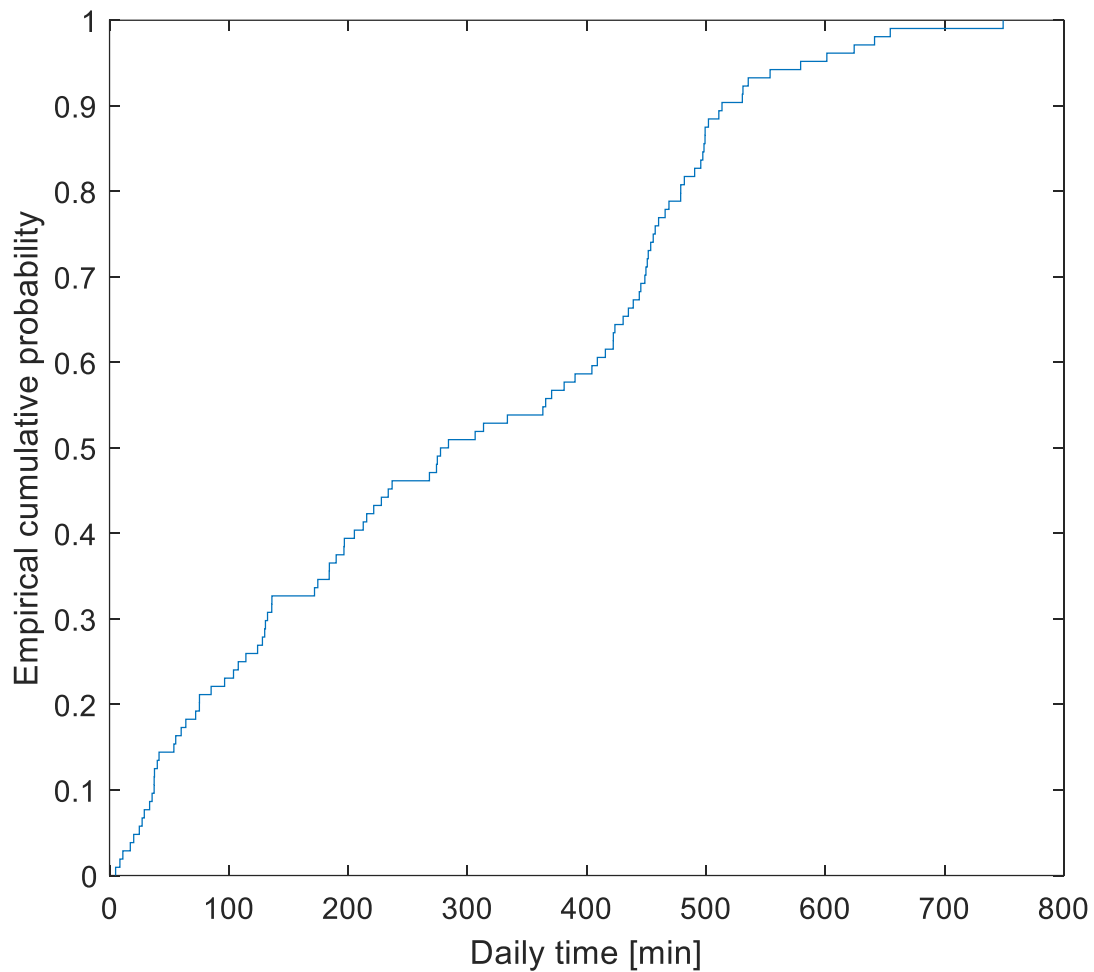

273
274   *Fig. 8: Headland pattern where T_H was lower than 10 s and longer than 100 s.*

275   The daily usage of the tractor ranges from 20 min to up 750 min; and the 50% of the days

276   the tractor was used for more than 280 min (Fig. 9).

277

Fig. 9: Daily time of the tractor during the period of analysis.

280   In Fig. 10, a typical daily operating cycle of the tractor is reported. This is composed of the

281   following activities:

282   1.   tractor idling at the farm at the beginning of the day for implement hitching or,

283   machine servicing and then moving the tractor from any farm unit to the field;

284   2.   field work where idling stops may occur for field machine maintenance;

285   3.   moving the tractor from the field to any farm unit.

286   When the driver turnover did not occur on field, or when changes in the tractor field

287   operation occurred during the day. The afore-described cycle was repeated twice in a day since

288   the driver went back to any farm unit for lunch or implement swapping.
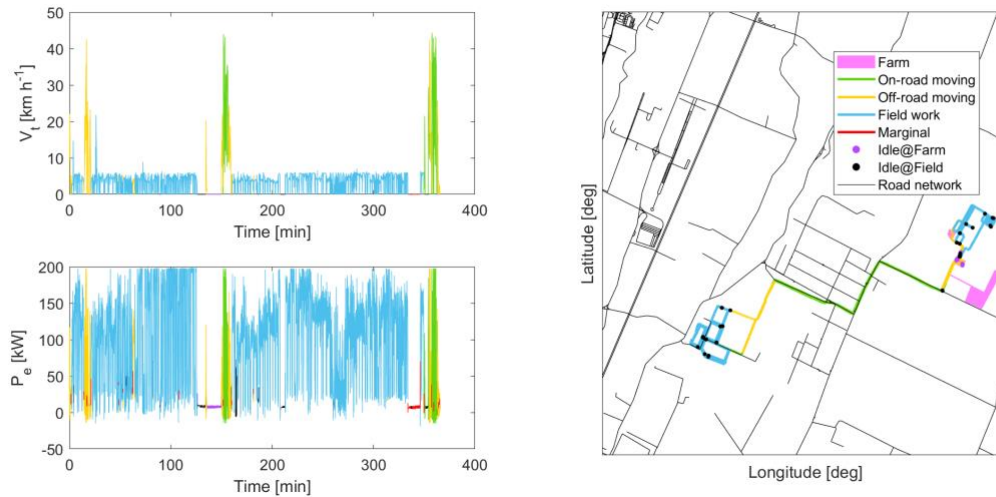
Fig. 10: Typical daily operating cycle of the tractor

In Fig. 11, the daily time contribution of the different tractor work states is reported. The largest contribution is provided by the *field work* state for 50% of the days, and it contributed to 73% of the entire daily activity. The other work states contribute less than 30% (without considering the outliers). The tractor was not used for field activities for 4 days since the daily time contribution of the *field work* state is 0%. In those days, larger contributions of the idling and moving states can be observed and the tractor was used for off-field activities because the weather conditions did not permit any field activities. Those activities consisted of machine servicing or moving implements from a farm unit to another. The results reported in other studies are aligned to the median values of the daily time contribution calculated in this study (Ettl et al., 2018; Kortenbruck et al., 2017).
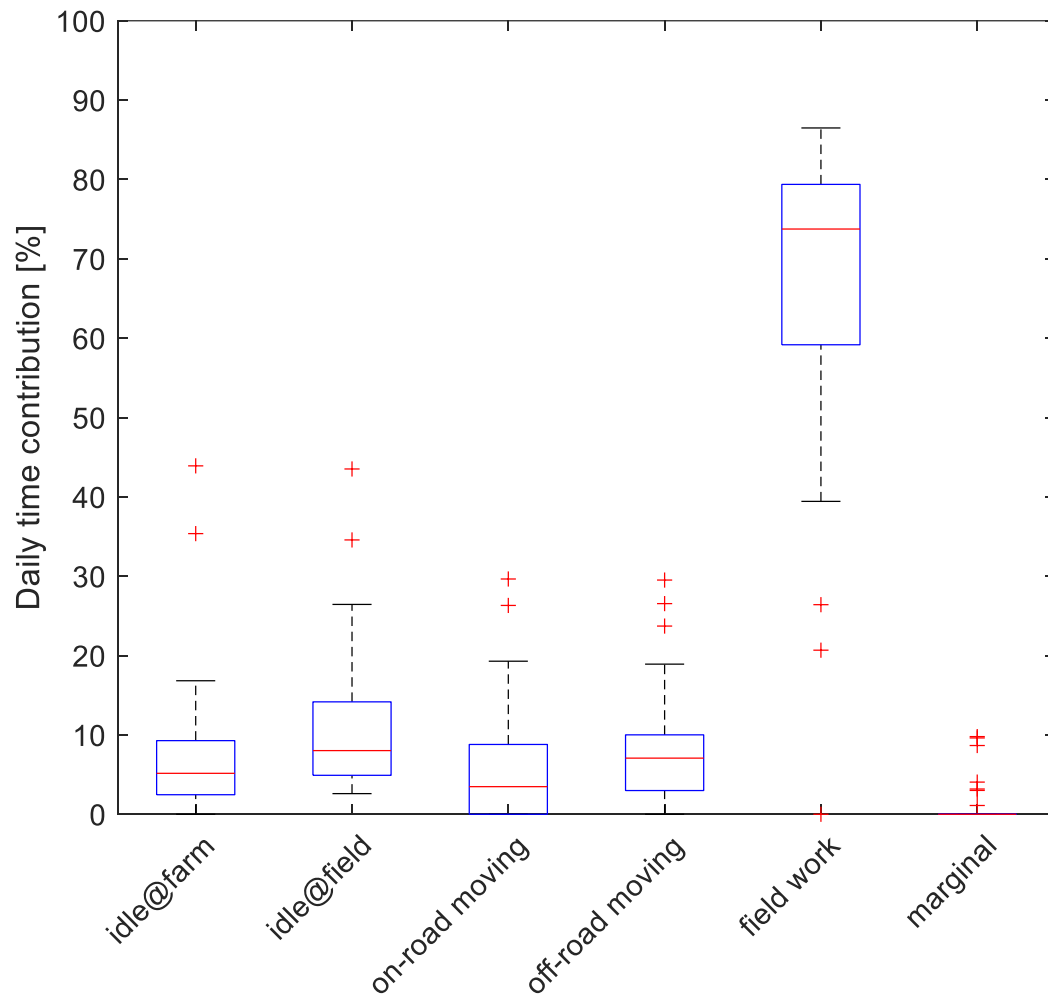
*Fig. 11: Box-plot reporting the daily time contribution of the tractor work states. Red crosses correspond to the outliers.*

305   The kernel smoothed probability distributions of V_t of the two moving states, and field work

306   states with the four most frequent implements are reported in Fig. 12. All the distributions have

307   unique modes except for that of the plough.

308

309
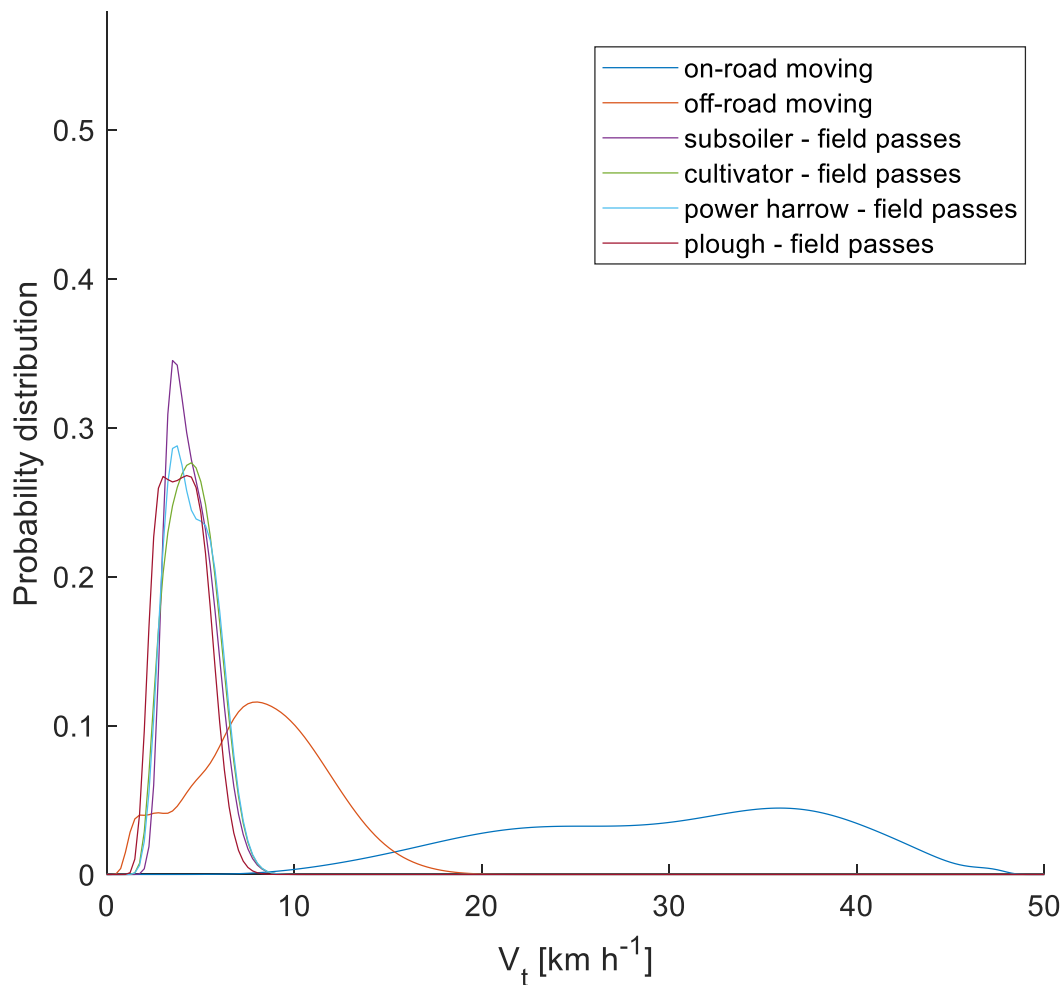
Fig. 12: Kernel smoothing distribution of tractor operating speed at various work states. Headlands were not considered on field operations.

In order to compare the distributions obtained in this study with those reported in the ASAE D497.7 (2011), the 10th and 90th percentiles, and the modes are reported in Table 4.

| Table 4 – Main statistics of the speed distributions			
Tractor states	**10th percentile [km h-1]**	**Mode [km h-1]**	**90th percentile [km h-1]**
on-road moving	20.0	36.0	40.0
off-road moving	3.5	8.0	12.4
subsoiler - field passes	3.3	3.5	5.8
cultivator - field passes	3.0	4.2	6.0
power harrow - field passes	3.2	3.7	5.9
plough - field passes	2.6	3.0 / 4.2	5.5

316   The speed values observed in this study are slightly lower than those reported in the ASAE

317   standard. For example, according to the ASAE standard, the speed range of a mouldboard

318   plough is between 5 and 10 km h^{-1}, whereas it was found to be between 2.6 km h^{-1} and 5.8 km

319   h^{-1} in this study. This difference may be because data in the ASAE standard is based on data

320   collected in the US, where tractors and fields larger than those available in the farm used in this

321   study.

322   The *off-road moving* distribution is overlapped with the distributions of the *field work* states

323   for V_t lower than 8 km h^{-1} (which corresponds to the 48th percentile of its cumulative

324   distribution), and it is overlapped with the distribution of *on-road moving* state for V_t higher

325   than 12 km h^{-1} (which corresponds to the 85th percentile of its cumulative distribution). This

326   highlights that discerning the moving states only using a threshold-rule for the speed, as it is

327   usually done with many commercial telemetric data services, may lead to misclassifications.

328   In Fig. 13, the confidence level ellipses reported for three work states demonstrate that to

329   fully discern the work states, a multivariate approach is necessary. The ellipses are clearly

330   separated with only minor overlaps occurring between the ellipses of both moving work states,

331   and between the ellipses of *field work* and *off-road moving* work states. V_t is strictly related to

332   P_e depending on the type of the work state. Indeed, on *field work* tasks, the tractor is usually

333   used with high engine loads, low speed and low gear ratios; while for moving tasks, lower

334   engine loads and longer gear ratios are typically used with, high engine loads are limited to

335   acceleration events only. For both moving tasks, P_e increases with V_t, due to the fact that the

336   main resistance forces are the motion resistance which increases with the ground speed (Wong,

337   2001). The variability of V_t and P_e inside each operating condition depends on several factors,

338   including driving style, operating, and environmental conditions. Indeed, P_e ranges from 23 kW

339   up to 143 kW for the field work state on passes. Through a visual inspection of the recorded

340   video, it was observed that the tillage operations carried out at low ground speeds (below 3 km

341 h^{-1}) occurred because the soil was severely covered by crop residuals. In these conditions,

342 farmers preferred to work slowly to avoid that the implement could get clogged with crop

343 residuals, which could force the farmer to stop at headlands for implement clearing.

344
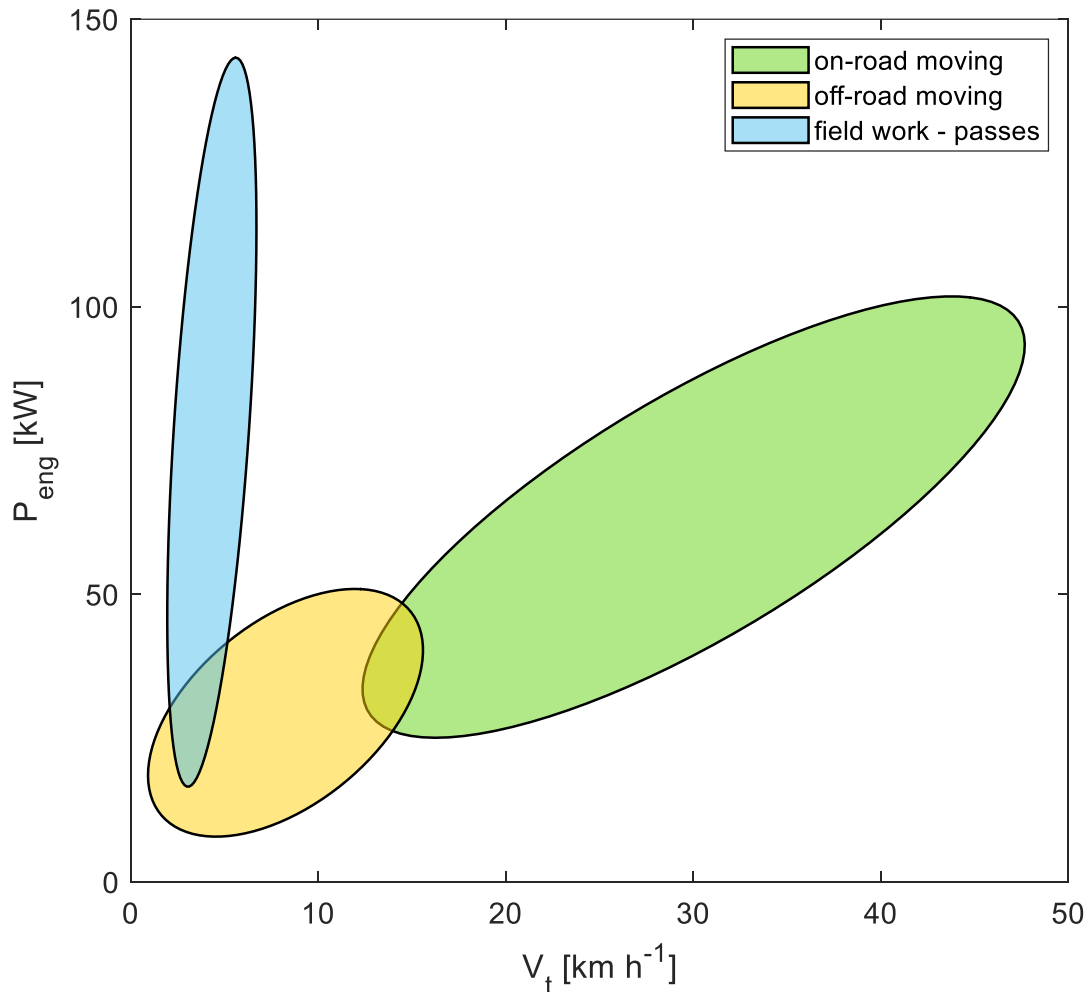
345
346 *Fig. 13: Confidence level ellipses for three non-idling operating states.*

Discussion

348   Few studies reported algorithms for classifying CAN-BUS data of agricultural tractors, and

349 in all of them, results are based on limited datasets, not collected in real-world conditions. The

350 strength of the classification scheme introduced in this study is the combined use of CAN-BUS,

351 trajectory, and geographical data. This allowed to obtain a finer description of the tractor usage.

352   Indeed, tractor movements were classified based on the road type (i.e. on-road and off-road).

353   On other similar studies, these two operating activities could not be distinguished because the

354   method relies only on CAN-BUS, and trajectory data. Indeed, Kortenbruck et al. (2017)

355   discriminated field from moving activities by evaluating the pattern of the tractor trajectory,

356   which was that of a field operation if parallel traces not further apart than the implement width

357   could be observed. The algorithm is not fully automatic because farmers have to input

358   implement widths, but CEPs of non-RTK- GNSS receivers are of the same order of magnitude

359   of typical implement widths. Thus, field operations cannot be reliably detected only with tractor

360   trajectories. Paraforos et al. (2018) detected headlands and field passes from trajectory data, in

361   particular, headlands were recognized when 180° overturns of tractor heading angle were

362   identified. However, tractors do not always overturns of 180° on headlands especially when

363   tractors work along the field contours (Fig. 8 – left). In Ettl et al. (2018), headland turns were

364   recognised by setting the threshold of the duration with the three-point linkage is fully up

365   position; however, the duration limit is dependent by the headland patterns (Paraforos et al.,

366   2018), and idling stops can be frequent during headlands (Molari et al., 2019) (Fig. 10 – right).

367   The approach for recognising the field operations adopted in this study is based on the

368   repetitive pattern of the *RWI* signal when tractors operate on field. This approach does not rely

369   on any threshold values of any CAN-BUS parameter, this makes it more effective in dealing

370   with real-world data. Setting the proper thresholds is a critical task because operating

371   parameters may change depending on the type of implement, soil conditions, ground slope, and

372   driving style (i.e. toward productivity or efficiency). The approach used in this paper works

373   only with mounted and semi-mounted implements where a variability on the *RWI* signal can

374   be observed, but the same principle could be used with other signals (e.g. steering angle) where

375   a repetitive pattern could be observed also with trailed implements.

376

Conclusions

Information on the usage of agricultural tractors is not well-documented and very often farmers, scientists, and engineers rely on handwritten logbook data. In this paper, a data acquisition system which facilitates the data collection on agricultural tractors, and a novel classification scheme were presented. The novelty of the data acquisition system is that it combines a CAN-BUS logger, a GNSS receiver and a BLE beacon scanner, thus the hitched implement could be recorded even if they are not ISOBUS compliant. Moreover, the novelty of the data analysis method is on the combined use of CAN-BUS, trajectory and geographical data which allowed to introduce a classification scheme more refined than those proposed in similar studies. Indeed, the kinds of tractor activities were classified into 5 states depending on the tractor operating condition and its position. Thanks to the proposed classification scheme, engineers may benefit from massively recorded real-world data in uncontrolled conditions, which may leverage their design method. Indeed, engineers may focus most of their efforts on optimising the most frequently used components; or they may extract the most relevant duty-cycle from a large dataset of real-world data (Bishop et al., 2012). The dynamic characteristics of front axle and cabin suspensions could be optimised in order to achieve the best balance between on and off-road performance based on the frequency of each moving state and of the most frequent operating speeds. Future work should focus on defining work states in greater details and adapting the classification scheme in order to analyse real-time CAN-BUS data. In this way, the algorithm could be embedded in in-vehicle computer systems and thus vehicle sub-systems could be controlled based on the actual work state. For example, parameters of tractor subsystems could be preventively set when the tractor is approaching a specific road type (i.e. on-road or off-road), and that could be especially useful for setting the damping coefficient of semi-active suspensions or the tyre pressure if the tractor embeds central tyre inflating system.

Acknowledgements

References

ASAE. (2011). *ASAE D497.7—Agricultural Machinery Management Data* (D497.7; ASAE Standard, pagg. 1–15). https://doi.org/10.13031/2013.36431

Balsari, P., Biglia, A., Comba, L., Alcatrão, L., Varani, M., Mattetti, M., Barge, P., Tortia, C., Manzone, M., Gay, P., & Ricauda Aimonino, D. (2020). Performance analysis of a tractor—Power harrow system under different working conditions. *Soil and Tillage Research*. https://doi.org/Submitted

Bishop, J. D. K., Axon, C. J., & McCulloch, M. D. (2012). A robust, data-driven methodology for real-world driving cycle development. *Transportation Research Part D: Transport and Environment*, *17*(5), 389–397. https://doi.org/10.1016/j.trd.2012.03.003

Calcante, A., & Mazzetto, F. (2014). Design, development and evaluation of a wireless system for the automatic identification of implements. *Computers and Electronics in Agriculture*, *101*, 118–127. https://doi.org/10.1016/j.compag.2013.12.010

Dati preconfezionati—GeoER. (2019). http://geoportale.regione.emilia-romagna.it/it/download/dati-e-prodotti-cartografici-preconfezionati/pianificazione-e-catasto/uso-del-suolo-1/2014-coperture-vettoriali-uso-del-suolo-di-dettaglio-edizione-2018/dati-preconfezionati

Ettl, J., Bernhardt, H., Pickel, P., Remmele, E., Thuneke, K., & Emberger, P. (2018). Transfer of agricultural work operation profiles to a tractor test stand for exhaust emission evaluation. *Biosystems Engineering*, *176*, 185–197. https://doi.org/10.1016/j.biosystemseng.2018.10.016

Fugiglando, U., Massaro, E., Santi, P., Milardo, S., Abida, K., Stahlmann, R., Netter, F., & Ratti, C. (2019). Driving Behavior Analysis through CAN Bus Data in an Uncontrolled Environment. *IEEE Transactions on Intelligent Transportation Systems*, *20*(2), 737–748. https://doi.org/10.1109/TITS.2018.2836308

Harmon, J. D., Luck, B. D., Shinners, K. J., Anex, R. P., & Drewry, J. L. (2018). Time-Motion Analysis of Forage Harvest: A Case Study. *Transactions of the ASABE*, *61*(2), 483–491. https://doi.org/10.13031/trans.12484

Hodge, V. J., & Austin, J. (2004). A Survey of Outlier Detection Methodologies. *Artificial Intelligence Review*, *22*(2), 85–126. https://doi.org/10.1007/s10462-004-4304-y

Hunt, D., & Wilson, D. (2015). *Farm Power and Machinery Management: Eleventh Edition*. Waveland Press.

438　ISO. (2012). *ISO 11783-7:2012—Tractors and machinery for agriculture and forestry—Serial*
439　　*control and communications data network—Part7: Implement messages application layer—*
440　　*Implement messages application layer*.

441　Johannesson, P., & Speckert, M. (2013). *Guide to Load Analysis for Durability in Vehicle*
442　　*Engineering*. John Wiley & Sons.

443　Kortenbruck, D., Griepentrog, H. W., & Paraforos, D. S. (2017). Machine operation profiles
444　　generated from ISO 11783 communication data. *Computers and Electronics in Agriculture*,
445　　*140*, 227–236. https://doi.org/10.1016/j.compag.2017.05.039

446　Lovarelli, D., Bacenetti, J., & Fiala, M. (2017). Effect of local conditions and machinery
447　　characteristics on the environmental impacts of primary soil tillage. *Journal of Cleaner*
448　　*Production*, *140*, 479–491. https://doi.org/10.1016/j.jclepro.2016.02.011

449　Marchesani, C., Parmigiani, F., & Vianello, M. (1992). *Integrated method to define the mission*
450　　*profile of a passenger car*. Innovation and reliability in automotive design and testing,
451　　Florence (I).

452　Mattetti, M., Maraldi, M., Sedoni, E., & Molari, G. (2019). Optimal criteria for durability test
453　　of stepped transmissions of agricultural tractors. *Biosystems Engineering*, *178*, 145–155.
454　　https://doi.org/10.1016/j.biosystemseng.2018.11.014

455　Mattetti, M., Molari, G., & Sedoni, E. (2012). Methodology for the realisation of accelerated
456　　structural tests on tractors. *Biosystems Engineering*, *113*(3), 266–271.
457　　https://doi.org/10.1016/j.biosystemseng.2012.08.008

458　Mattetti, M., Molari, G., & Sereni, E. (2017). Damage evaluation of driving events for
459　　agricultural tractors. *Computers and Electronics in Agriculture*, *135*, 328–337.
460　　https://doi.org/10.1016/j.compag.2017.01.018

461　Mattetti, M., Varani, M., Molari, G., & Morelli, F. (2017). Influence of the speed on soil-
462　　pressure over a plough. *Biosystems Engineering*, *156*, 136–147.
463　　https://doi.org/10.1016/j.biosystemseng.2017.01.009

464　Molari, G., Mattetti, M., Lenzini, N., & Fiorati, S. (2019). An updated methodology to analyse
465　　the idling of agricultural tractors. *Biosystems Engineering*, *187*, 160–170.
466　　https://doi.org/10.1016/j.biosystemseng.2019.09.001

467　Molari, G., Mattetti, M., Perozzi, D., & Sereni, E. (2013). Monitoring of the tractor working
468　　parameters from the CAN-Bus. *AIIA 13*. Horizons in agricultural, forestry and biosystems
469　　engineering, Viterbo.

470　Paraforos, D. S., Hübner, R., & Griepentrog, H. W. (2018). Automatic determination of
471　　headland turning from auto-steering position data for minimising the infield non-working
472　　time. *Computers and Electronics in Agriculture*, *152*, 393–400.
473　　https://doi.org/10.1016/j.compag.2018.07.035

474　Paraforos, D. S., Vassiliadis, V., Kortenbruck, D., Stamkopoulos, K., Ziogas, V., Sapounas, A.,
475　　& Griepentrog, H. W. (2017). Automating the process of importing data into an FMIS using
476　　information from tractor's CAN-Bus communication. *Advances in Animal Biosciences*, *8*,
477　　650–655. https://doi.org/10.1017/S2040470017000395

478　Perozzi, D., Mattetti, M., Molari, G., & Sereni, E. (2016). Methodology to analyse farm tractor
479　　idling time. *Biosystems Engineering*, *148*, 81–89.
480　　https://doi.org/10.1016/j.biosystemseng.2016.05.007

481 Pitla, S. K., Lin, N., Shearer, S. A., & Luck, J. D. (2014). Use of Controller Area Network
482   (CAN)   Data To Determine Field Efficiencies   of Agricultural Machinery. *Applied*
483   *Engineering in Agriculture*, *30*(6), 829–839. https://doi.org/10.13031/aea.30.10618

484 Pitla, S. K., Luck, J. D., Werner, J., Lin, N., & Shearer, S. A. (2016). In-field fuel use and load
485   states of agricultural field machinery. *Computers and Electronics in Agriculture*, *121*, 290–
486   300. https://doi.org/10.1016/j.compag.2015.12.023

487 Plaskitt, R. J., & Musiol, C. J. M. (2002). *Developing a Durable Product*. 1–20.

488 SAE. (2006). *Agricultural and Forestry Off-Road Machinery Control and Communication*
489   *Network* (N. j1939-2). https://saemobilus.sae.org/content/j1939/2_200608

490 SAE. (2013). *Vehicle Application Layer* (N. j1939-71; pagg. 1–1255).

491 SAE. (2016a). *SAE J1939-14—Physical Layer, 500 Kbps* (SAE J1939-14; pagg. 1–13).

492 SAE. (2018b). *SAE J1939-15—Physical Layer, 250 Kbps* (SAE J1939-15; pagg. 1–20).

493 Sehab, R., Barbedette, B., & Chauvin, M. (2011). Electric vehicle drivetrain: Sizing and
494   validation using general and particular mission profiles. *2011 IEEE International*
495   *Conference on Mechatronics*, 77–83. https://doi.org/10.1109/ICMECH.2011.5971228

496 Wong, J. Y. (2001). *Theory of Ground Vehicles*. John Wiley & Sons.

497 Zhang, Y., Ault, A., Krogmeier, J. V., & Buckmaster, D. (2017). Activity Recognition for
498   Harvesting   via   GPS   Tracks.   *2017   ASABE   Annual   International   Meeting*.
499   https://doi.org/10.13031/aim.201700813

500