Smart Agricultural Technology 13 (2026) 101777

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

CANBUS to drawbar load estimation: Mapping real-world tractor loads for mission profiling

Luca Colendi [a], Manuel Tentarelli [b], Massimiliano Varani [a,*], Michele Mattetti [a]

[a] Department of Agricultural and Food Sciences – Alma Mater Sudiorum - University of Bologna, viale G. Fanin 50, Bologna, Italy
[b] CNH Industrial Italia S.p.A, , Italy

ARTICLE INFO

Keywords:
Drawbar force mapping
Fuel efficiency
Real-world data
Transmission modelling
CANBUS

ABSTRACT

With increasingly stringent requirements for greenhouse gas reduction, assessing current tractor drivetrains and developing cleaner alternatives are of growing importance. Designing such systems requires extensive field-load data, which remain difficult to obtain due to the absence of scalable, long-term measurement methods. This study presents a non-invasive methodology for evaluating tractor field loads using only controller area network (CANBUS) messages, universally available on modern tractors. A static mathematical model of a hydromechanical continuously variable transmission (IHMCVT) is developed, validated through experimental tests, and applied to map drawbar forces under real-world conditions. The model employs convergence-based calculations with multiple feedback loops and shallow neural networks to determine hydrostatic unit efficiencies. Validation against power-based measurements from road and field tests demonstrates accuracy within ±10%, suitable for practical applications. Applied to one year of operational data, the method reveals that tractors mainly operate at low speeds (<10 km h^{-1}) under heavy soil tillage and transport conditions. Specific fuel consumption mapping further highlights inefficiencies due to engine–transmission interactions. Overall, the proposed CANBUS-based model provides a reliable, scalable, and low-complexity approach for real-world mission profiling and future drivetrain optimization.

1. Introduction

Reducing greenhouse gas emissions and fossil fuel use is essential for promoting sustainable agriculture and mitigating climate change. Tractors are one of the most popular off-highway vehicles operating worldwide; they are used in varying operating conditions, from field operations with implements to on-road driving with heavy trailers. One of the most frequent energy-demanding operations of agricultural tractors is the development of a drawbar load for towing implements or trailers, and both may use up 78 % of the time and up to 92 % of the fuel [1].

Following the lead of the automotive industry, a large part of the recent research on agricultural tractors has focused on evaluating and designing hybrid or electrified powertrains [1–3]. These would result in significant fuel, CO_2, and cost savings, but the design of such solutions requires a set of field data to size components and evaluate their potential benefits from the point of view of the farmer.

In most studies, researchers have evaluated these benefits using a limited set of field measurements performed for their projects [4,5]. Even if the approach appears reasonable, the intrinsic variability of field conditions due to heterogeneous soil properties, topography, and crop requirements greatly limits the generalisation of the results and the comparison of the proposed solutions.

To overcome such limitations, the automotive industry has based the design of novel solutions on reference driving cycles (e.g., Worldwide Harmonized Light Vehicle Test Procedure (WLTP), Real Driving Emissions (RDE) testing) [6,7]. There are no legislative driving cycles for agricultural tractors; however, a methodology for outlining a reference driving cycle for engines installed on non-road mobile machinery has been recently proposed [8]. This cycle was calculated to aid the development of engines, but it cannot be used for the development of entire drivetrains, as, in this case, field-load data are needed. The DLG-PowerMix test [9] is often used as a standardized benchmark for evaluating tractor performance and fuel consumption under representative load conditions, although it does not constitute a regulatory driving cycle and is based on stationary load points rather than continuous field-load histories.

Field-load data are the collection of loads (i.e., forces, moments, and environmental conditions) experienced by a vehicle while operating

---

* Corresponding author.
E-mail address: massimiliano.varani@unibo.it (M. Varani).

https://doi.org/10.1016/j.atech.2026.101777
Received 28 October 2025; Received in revised form 27 December 2025; Accepted 1 January 2026
Available online 3 January 2026

Nomenclature

$P_x(e_{abs})$	e_{abs} at its x^{th} percentile [N]
e_{abs}	Absolute error [N]
$toll$	Algorithm stopping tolerance [-]
μSFC	Arithmetic mean of specific fuel consumption [g kWh^{-1}]
μ	Arithmetic mean operator [-]
ω_c	Carrier speed [rad s^{-1}]
T_c	Carrier torque [Nm]
CI	Confidence interval
CANBUS	Controller area network [-]
P_c	Cooling hydraulic circuit power absorption [W]
R	Correlation coefficient [-]
P_{axle}	Driving axles power [W]
$\dot{f}_f$	Engine fuel rate [m^3 s^{-1}]
θ_f	Engine fuel temperature [K]
ω_{ICE}	Engine speed [rad s^{-1}]
η_{est}	Estimated hydraulic unit efficiency [-]
ω_{s1}	First sun speed [rad s^{-1}]
T_{s1}	First sun torque [Nm]
4WD	Four-wheel drive [-]
$F_{x_{FL}}$	Front-left longitudinal wheel force [N]
$F_{x_{FR}}$	Front-right longitudinal wheel force [N]
$\rho_f(\theta_f)$	Fuel density [kg m^{-3}]
ω_1	Gearwheel 1 speed [rad s^{-1}]
T_1	Gearwheel 1 torque [Nm]
ω_2	Gearwheel 2 speed [rad s^{-1}]
T_2	Gearwheel 2 torque [Nm]
η_{gw}	Gearwheel mechanical efficiency [-]
AC	Global cycles accumulation [-]
GNSS	Global navigation satellite system [-]
$P_{g,ICE}$	Gross engine power [W]
$T_{g,ICE}$	Gross engine torque [Nm]
p_{max}	HST maximum pressure [MPa]
p	HST pressure [MPa]
K_ω	HST speed ratio [-]
K_T	HST torque ratio [-]
HST	Hydrostatic transmission [-]
IHMCVT	Input coupled hydromechanical continuously variable transmission [-]
ICE	Internal combustion engine [-]
IQR(e_{abs})	Interquartile range of e_{abs} [N]
KDE	Kernel density estimation [-]
LS	Load spectra [-]
P_l	Lube hydraulic circuit power absorption [W]
$T_{wheels_{max}}$	Maximum traction torque to the tractor wheels [Nm]
MSE	Mean square error
$F^{meas}_{d_{jWD},\ i}$	Measured ith drawbar force level [N]
N^{meas}_i	Measured number of cycles at the ith drawbar force level [-]
Med(e_{abs})	Median of e_{abs} [N]
F^{mod}_d	Modelled drawbar force [N]
$F^{mod}_{d,\ i}$	Modelled ith drawbar force level [N]
N^{mod}_i	Modelled number of cycles at the ith drawbar force level [-]
$\overset{=}{\omega}_{pin}$	Modelled pin speed matrix [rad s^{-1}]
T^{mod}_{pin}	Modelled pinion torque [Nm]
$\overset{=}{T}_{pin}$	Modelled pinion torque matrix [Nm]
T_{ICE}	Net transmission torque [Nm]
NRMSE	Normalized root mean square error
η_{OG}	OG mechanical efficiency [-]
τ_{OG}	OG speed ratio [-]
FD	On-field drawbar test [-]
FT	On-field tillage test [-]
RT	On-road transport test [-]
OG	Ordinary gear [-]
ω_{pin}	Pinion speed [rad s^{-1}]
η^n_m	Pinion-to-driving axles mechanical efficiency [-]
U1	Piston pump unit 1 [-]
U2	Piston pump unit 2 [-]
PG	Planetary gear [-]
c	Planetary gear carrier [-]
s1	Planetary gear first sun [-]
τ_I	Planetary gear first-stage speed ratio [-]
r	Planetary gear ring [-]
s2	Planetary gear second sun [-]
τ_{II}	Planetary gear second-stage speed ratio [-]
$P(\omega_{ICE})$	Power absorption functions [W]
PTO	Power take-off [-]
P_p	Pre-charge pumps power absorption [W]
η_{real}	Real hydraulic unit efficiency [-]
$\omega_{r_{axle}}$	Rear axle angular speed [rad s^{-1}]
$\tau_{r_{axle}}$	Rear axle ratio [-]
pin	Rear pinion shaft [-]
V_{th}	Rear wheel peripheral speed [m s^{-1}]
$F_{x_{RL}}$	Rear-left longitudinal wheel force [N]
$F_{x_{RR}}$	Rear-right longitudinal wheel force [N]
$e_{\%}$	Relative error [%]
ω_r	Ring speed [rad s^{-1}]
T_r	Ring torque [Nm]
RMSE	Root mean square error [%]
SNN	Shallow neural network [-]
SFC	Specific fuel consumption [g kWh^{-1}]
α	Stage parameter [-]
λ	Swash plate position parameter [-]
$F^{meas}_{d_{jWD}}$	Total measured drawbar force [N]
δ_j	Traction condition parameter [-]
η_{TE}	Traction efficiency [-]
s	Tractor slippage [-]
δ	Transfer case switch selector signal [-]
D_{u1}	Unit 1 displacement [m^3]
$\omega_{u1\,max}$	Unit 1 maximum speed [rad s^{-1}]
D_{u2}	Unit 2 displacement [m^3]
ω_{u2max}	Unit 2 maximum speed [rad s^{-1}]
V_{grd}	Vehicle ground speed signal [m s^{-1}]

under real-world conditions. Drivetrains should be described as time histories of field grade, ground speed, and drawbar load. Field load data can be recorded on a specific vehicle through dedicated tests [10] using the correct instruments, such as load cells [11] and wheel force transducers [12]. However, these instruments cannot be installed on a vehicle operating in prolonged field usage due to their limited durability, required vehicle modifications, and high costs, so there is no reliable, scalable way to capture year-long drawbar load data from tractors in operation. Recent studies have investigated CAN-based analytics for tractor mission profiling [13] and load estimation [14], as well as hydro-mechanical CVT modelling approaches [15,16]. However, these works either rely on additional sensors, focus on short-term measurements, or address dynamic simulation rather than scalable long-term monitoring. Despite the widespread availability of CANBUS technology, no existing methodology provides a comprehensive, non-invasive solution for estimating tractor drawbar loads over extended periods using only standard signals. Current approaches either require additional hardware or lack integration with transmission models, limiting their scalability and accuracy. This gap motivates the development of a CAN-based framework capable of supporting mission profiling and drivetrain optimization at fleet level. This limits the amount of field-load data that can be collected, further limiting the coverage of the wide range of operating conditions occurring in the real world (i.e., vehicle mass, type of driveline, soil, and implement).

Load-based models can be used to estimate real-world fuel use and emissions, and they can model the real-world power demand of vehicles [17]. Nevertheless, the wide variability of the model parameters caused by the wide range of operating conditions encountered by tractors during their life cycles hinders the adoption of such an approach for field-load data estimation. However, the controller area network (CANBUS) technology has been available for years in all tractors in the marketplace. CANBUS can easily record the operating parameters of the engine and transmission; these data could be used to estimate the field-load data using a vehicle model.

The goal of this study is to propose a methodology for extensively evaluating tractor field loads only using the standard CANBUS messages available on all machines. The specific objectives are:

- To develop a static IHMCVT-based model capable of estimating drawbar force using standard CANBUS signals.
- To validate the proposed model under diverse operating conditions through experimental tests.
- To apply the methodology to a one-year dataset for real-world mission profiling.

The approach relies on a static transmission model that enables an accurate estimation of the drawbar force delivered by the tractor, The model is first validated through dedicated tests carried out both in the field and on the road. Subsequently, the methodology is applied to a real-world dataset collected over more than a year, in order to identify the dominant operational patterns. This approach offers a practical advantage for manufacturers, as it allows them to gain a detailed understanding of the actual loads experienced by tractors during operation without installing any additional sensors. In a future perspective, such knowledge could support the optimization of drivetrain design, durability assessment, and energy management strategies based solely on data already available on board.

2. Materials and methods

This study is structured into two main phases, summarised as follows:

- **data collection and model development:** data were collected through multiple experimental tests and real-world activities, and a vehicle model was developed to estimate drawbar load using CANBUS data.
- **model validation and real-world drawbar analysis:** the model was validated using CANBUS and drawbar test data, while real-world CANBUS data were employed to estimate drawbar load under actual operating conditions.

2.1. Data collection and model development

The study was based on the New Holland T7.315 HD tractor (CNH Industrial N.V., Amsterdam, The Netherlands). Its technical information is shown in Table 1.

This tractor model is equipped with a dual-stage input coupled hydromechanical continuously variable transmission (IHMCVT), a hybrid continuously variable transmission (CVT) that transmits power from the internal combustion engine (ICE) to the wheels through both mechanical and hydraulic pathways. This type of transmission controls the output power of the tractor by modulating three key inputs: engine torque, engine speed, and the continuously adjustable transmission ratio, which directly determines the speed of the vehicle.

Table 1
Declared technical specifications and equipment.

Engine maximum power [kW]	230
Maximum speed [km h^{-1} @ rpm]	50 @ 1600
Transmission	dual-stage input coupled hydromechanical continuously variable transmission (IHMCVT)
Unballasted static mass [kg]	11,570
Maximum permissible mass [kg]	16,800
Front tyre	Michelin AXIOBIB 2 VF 600/70 R30 168D/165E TL
Rear tyre	Michelin AXIOBIB 2 VF 710/70 R42 182D/179E TL

2.1.1. Data collection

For the extensive drawbar force mapping activity, a sample New Holland T7.315 HD was monitored using a CANBUS data logger (Kvaser Memorator Pro 5xHS, KVASER, Molndal, Sweden). Data were collected over one year of farming activities (from January 2023 to December 2023) at the Agricultural Farm of the University of Bologna (Cadriano, BO, Italy, 44°32′54″N 11°24′45″E), accumulating 500 h of use. The data logger was configured to record the CANBUS messages whenever the tractor engine was activated, ensuring that the recording did not interfere with farming activities. The CANBUS input signals having the following Suspect Parameter Numbers (SPNs) and Parameter Group Numbers (PGNs) (ISO, 2012; SAE, 2013) were recorded:

- SPN 190 and PGN 61,444: the engine speed (ω_{ICE}), which is the revolution speed of the engine crankshaft (sample rate 10 Hz).
- SPN 544 and PGN 65,251: the engine reference torque ($T_{r,ICE}$) reporting the maximum torque the engine can deliver (sample rate 0.2 Hz).
- SPN 513 and PGN 61,444: the actual engine - percent torque ($T_{\%,ICE}$) reporting the torque as a percent of $T_{r,ICE}$ (sample rate 50 Hz).
- SPN 513 and PGN 5398: the nominal friction-percent torque ($T_{\%,f}$) reporting the frictional and thermodynamic loss of the engine, pumping torque loss, and the losses of fuel, oil, and cooling pumps, as a percent of $T_{r,ICE}$ (sample rate 20 Hz).
- SPN 183 and PGN 65,266: the engine fuel rate ($\underset{f}{\dot{f}}$), reporting the amount of fuel consumed by the engine per unit of time (sample rate 10 Hz).
- SPN 5540 and PGN 64,740: the engine fuel temperature (θ_f) reporting the fuel temperature (sample rate 10 Hz).
- SPN 191 and PGN 61,442: the transmission output shaft speed (ω_{pin}), representing the angular velocity of the pinion shaft mechanically connected to the rear differential (sample rate 10 Hz).
- SPN 84 and PGN 65,229: the wheel-based vehicle speed (V_{th}), which is the rear wheel peripheral speed of the vehicle (sample rate 10 Hz).
- SPN 2796 and PGN 64,980: the transfer case selector switch (δ), which determines whether the tractor is in two-wheel drive (2WD) or four-wheel drive (4WD) (sample rate 10 Hz).
- SPN 3948 and PGN 64,932: the engagement of the power take-off (PTO), providing the information whether at least one PTO is currently engaged (sample rate 10 Hz).

Lastly, the tractor was equipped with a global navigation satellite system (GNSS) receiver (xProGPS_nano, Suchy Data Systems, GmbH, Erdweg, Germany) with a probable circular error of 2.5 m and a sampling frequency of 25 Hz. This allowed for the acquisition of the vehicle ground speed signal (V_{grd}).

As PTO loads cannot be directly measured with standard CAN signals, the analysis included only data collected with the PTO inactive, representing 84 % of the original dataset. This restriction implies that the operational picture does not cover scenarios where the PTO is engaged. It should be emphasized that this choice is not due to a limitation of the model, but rather reflects its design objective of operating solely using standard CAN signals, without requiring additional hardware. It is important to note that, although the CANBUS provides PTO engagement status and PTO speed, it does not provide the corresponding PTO torque or power absorption. Without this information, the net power input to the transmission cannot be accurately reconstructed when the PTO is active, making those operating points incompatible with the present model structure. This is not a limitation of the transmission model itself, but a direct consequence of the signals available through standard CANBUS logging. Furthermore, PTO-driven implements can introduce additional dynamic effects—such as small fluctuations of PTO speed around the nominal setpoint—that would require dedicated sensors and a specific modelling framework. For these reasons, the analysis was restricted to conditions with the PTO inactive, since complete characterization would require a dedicated experimental campaign, which goes beyond the scope of the current study. While PTO engagement is common in real-world farming, the current dataset is representative of a large fraction of typical transport and field operations where the PTO is inactive. Therefore, although some bias may exist for applications that heavily involve PTO-driven implements, the filtered dataset still captures the main operational regimes relevant for transmission performance estimation.

Besides, a validation activity was conducted, consisting of separate experimental tests on the tractor to gather data during specific real-world activities. Specifically, the drawbar force was recorded and compared with the model estimate. For this validation phase, additional signals were recorded (sampling frequency of 100 Hz). In particular, these additional signals were obtained using wheel force transducers (WFTs), specifically the LW-2T-100 K model from Michigan Scientific Corporation (Charlevoix, United States of America), which collected traction force signals from all four tractor wheels. In particular, the following signals were acquired: $F_{x_{RL}}$ and $F_{x_{RR}}$, which are the rear-left and the rear-right longitudinal wheel forces, respectively; and $F_{x_{FL}}$ and $F_{x_{FR}}$, which are the front-left and the front-right longitudinal wheel forces, respectively. Validation data were collected from three on-road transport tests (RT), three on-field drawbar tests (FD), and two on-field tillage tests (FT), as illustrated in Table 2, where tractor settings are shown.

Transport tests were conducted in Concordia Sulla Secchia (MO, Italy, 44°54′47″ N 10°59′11″ E), while on-field tests were carried out at the Agricultural Farm of the University of Bologna. According to the USDA soil texture classification [18], the soil at the Agricultural Farm of the University of Bologna was classified as clay loam, with sand, silt, and clay percentages of 34 %, 36 %, and 30 % in the top 300 mm layer, respectively [19]. Soil moisture content and bulk density were determined following the procedures described in ASTM standards [20]. Soil strength was assessed using a cone penetrometer (Fieldscout SC 900, Spectrum Technologies Inc., Aurora, IL, USA) equipped with a 12.7 mm diameter tip. Measurements were collected down to a depth of 300 mm, at 25 mm intervals, with a constant penetration rate of 25 mm s^{-1} [21]. In total, 120 profiles were acquired. The field, which had been cultivated with soy and harvested four months before the test session, had an environmental temperature of 8 °C, as measured by the Funo weather station (44°35′30″ N, 11°22′0″ E; elevation 24 m a.s.l.). Soil moisture content was 16.7 % (±1.7 %), and bulk density was 1318.1 kg m^{-3} (±143.9 kg m^{-3}). The average cone index (CI) values were 0.627 MPa for the 0–150 mm layer (CI_{0-150}) and 0.792 MPa for the 150–300 mm layer ($CI_{150-300}$). Based on these values, the soil was categorized as 'soft soil' (SS) according to the criteria proposed by Brixius [22].

Table 2
Validation test nomenclature and conditions.

Test ID	Terrain	Activity	Static mass [kg]	Rear mass [%]	Ballast (front-rear) [kg]	Tire pressure (front-rear) [kPa]
RT1	Road (asphalt)	Transport	11,570	60	0–0	100–100
RT2	Road (asphalt)	Transport	15,500	65	800–3130	100–100
RT3	Road (asphalt)	Transport	19,370	60	3100–4700	100–100
FD1	Field	Drawbar test	15,500	65	800–3130	100–100
FD2	Field	Drawbar test	19,370	60	3100–4700	100–100
FD3	Field	Drawbar test	19,370	50	5060–2740	100–100
FT1	Field	Tillage ploughing	16,270	60	1950–2750 (implement on rear)	110–120
FT2	Field	Tillage subsoiling	16,270	60	1950–2850 (implement on rear)	110–120

During transport tests, the transmission was set to 2WD and the gearbox set in automatic mode, allowing the control strategy of the vehicle to autonomously define engine operating points based on the required speed and traction torque. Typical town routes from the farm to the field were recorded for each transport test (see Fig. 1). These tests were conducted on a 7.1-km asphalt route with an average gradient of 0.07 %. Transport RT1 was carried out using the unballasted vehicle, while RT2 and RT3 involved the application of ballast. This setup allowed for model investigation against diverse and realistic road uses, characterised by dynamics of acceleration and deceleration.

In contrast, on-field drawbar tests were executed employing a secondary tractor as a braking unit (New Holland T8 Genesis™ 335, CNH Industrial N.V., Amsterdam, The Netherlands) that was meant to apply a braking action over the first ballasted tractor through a steady-state procedure defined by Angelucci et al. [12].

The ploughing tests were conducted, with the tractor ballasted, using a Maschio Unico XL plough (Maschio Gaspardo, S.p.A. Campodarsego,

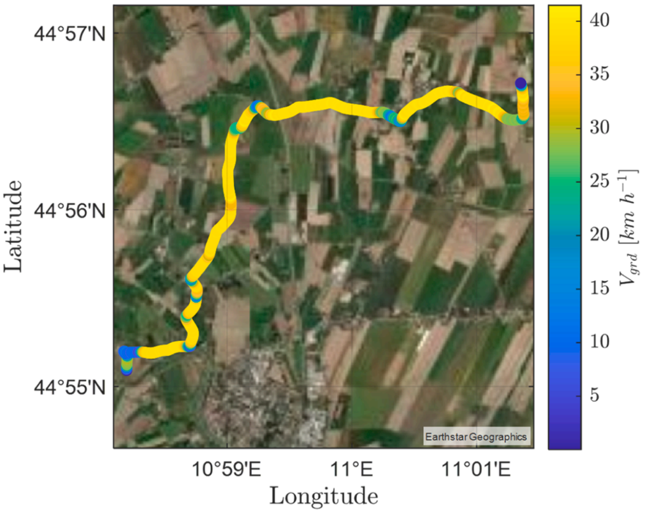


Fig. 1. The town-asphalted farm-field route in Concordia Sulla Secchia (MO, Italy) used in the on-road tests. The top-right route endpoint corresponds to the farm, and the bottom-left is the field. The marker colour represents the ground speed of the tractor (V_{grd}).

Padova, Italy), weighing approximately 2850 kg, with 4 bodies and set with a working width of 2.0 m, working depth of 0.25 m, and operating with V_{th} at 5, 6, 7, and 8 km h^{-1} for each respective pass. In contrast, the subsoiling tests were performed, again with the tractor ballasted, using a Maschio Artiglio subsoiler (Maschio Gaspardo S.p.A., Campodarsego, Padova, Italy), weighing around 2750 kg, equipped with 7 anchors and a maximum working depth of 0.55 m. For these tests, the data used for validation were collected over field passes with a working depth of 0.50 m and V_{th} set at 5, 6, and 7 km h^{-1} for each respective pass [10]. For both types of tests, the working widths or depths and speeds were chosen to reflect the typical field operation conditions in the test area. Table 3 reports the summarized soil characteristics and implements settings for the tests.

2.1.2. Data pre-processing

All signals were time-synchronized and stored in ASCII format. The CAN messages were resampled to 100 Hz using CANalyzer 10.0 SP7 (Vector Informatik GmbH, Stuttgart, Germany) to match the acquisition rate of the Wheel Force Transducers (WFTs). Discrete signals, such as the transfer case position and PTO engagement, were resampled using a zero-order hold approach to preserve their binary nature, whereas spline interpolation was applied to continuous variables. Because the model developed in this study is static, the analysis focused on the low-frequency content of the measured signals. All numerical data used for validation and drawbar-force mapping were converted into MATLAB format (The MathWorks Inc., Natick, MA, United States of America), then processed with a fourth-order Butterworth low-pass filter with a cutoff frequency of 1 Hz. The power spectral density of the WFT measurements indicated that over 70–80 % of the signal energy lies below this threshold, while higher-frequency components mainly reflect measurement noise or rapid transients not relevant to the static modelling approach.

Once filtered signals were obtained, the gross engine torque was calculated as in Eq. (1).

$$T_{g,ICE} = T_{r,ICE} * \frac{T_{\%,ICE} - T_{\%,f}}{100} \tag{1}$$

while a subsequent correction was implemented over the gross engine power signal ($P_{g,ICE}$), calculated as in Eq. (2), to obtain the net torque input to the drivetrain. This correction was made to account for power absorption upstream of the drivetrain. The obtained $T_{g,ICE}$ signal already accounts for the engine losses, but it does not reflect the power absorption due to the systems between the flywheel and the transmission.

$$P_{g,ICE} = T_{g,ICE} * \omega_{ICE} \tag{2}$$

To perform the correction, simple regression functions were introduced to consider power absorptions from the tractor hydraulic circuit pre-charge pumps (P_p), the lube hydraulic circuit (P_l), and the cooling hydraulic circuit (P_c). Each power absorption was expressed as a function of ω_{ICE} (see Eq. (3), in which P_* stands for any specific power component), where β_0 and β_1 are the regression coefficients [23]. The regression models were derived from experimental bench-data provided by the tractor manufacturer, as well as from the values reported by Migliaccio [24]. Before regression, ω_{ICE} was standardized by z-score normalization, using $\mu_{\omega_{ICE}}$ and $\sigma_{\omega_{ICE}}$, i.e., the mean and standard deviation of the regression datasets. The values of β_0, β_1, $\mu_{\omega_{ICE}}$, and $\sigma_{\omega_{ICE}}$ are available in Table 4.

$$P_*(\omega_{ICE}) = \beta_0 + \beta_1 * \frac{\omega_{ICE} - \mu_{\omega_{ICE}}}{\sigma_{\omega_{ICE}}} \tag{3}$$

$$T_{ICE} = \frac{P_{g,ICE} - P_p(\omega_{ICE}) - P_l(\omega_{ICE}) - P_c(\omega_{ICE})}{\omega_{ICE}} \tag{4}$$

2.1.3. IHMCVT model

After filtering and correcting the signals, the mathematical structure of the IHMCVT model and the procedure for calculating the drawbar force were defined, forming the foundation of the overall vehicle model. The IHMCVT architecture underpins this model, thus its configuration is first described below.

The IHMCVT is composed of three primary components: an ordinary gear (OG), a planetary gear (PG), and a hydrostatic transmission (HST), [16]. A simplified yet complete version for symbolic representation is presented in Fig. 2. For the considered IHMCVT, the ICE, the OG, and the first solar of the PG (s1) are connected to the transmission main shaft (a) that constitutes the mechanical branch of the system.

The HST comprises two hydraulic swash plate piston pump units (U1 and U2) and is connected to the drivetrain through the OG and the PG ring (r). As mentioned before, the PG is double-staged: it can extract power from the carrier (c) and the second solar (s2) of the PG. If the first stage is active, power will be transmitted from component c, through the corresponding clutch, to the rear pinion shaft (pin). Conversely, if the second stage is active, the corresponding clutch will transmit power from component s2 to the pin. Therefore, it is necessary to define the characteristic relations of the primary components of the IHMCVT.

The transmission was modelled through a set of static equations describing the macroscopic, time-independent behaviour of the system. The complete formulation of these equations is provided in Appendix A.

This modelling process aimed to calculate the speeds and torques of each rotating element of the drivetrain, given the engine and transmission operating points. Extracting these speeds and torques was essential for predicting the operational behaviour of the transmission during working conditions, enabling the following assessment of the resulting drawbar force.

2.1.4. Model output extraction

For a detailed explanation of the symbols used throughout the following sections, refer to Appendix A.

An IHMCVT can express the power exerted by the tractor by

Table 4
Regression coefficients for.

Power absorptions	β_0 [W]	β_1 [W]	$\mu_{\omega_{ICE}}$ [rad s^{-1}]	$\sigma_{\omega_{ICE}}$ [rad s^{-1}]
P_p	5063	1668	146.6	51.7
P_l	466.7	2.797×10^{-15}	157.1	51.3
P_c	1146	1.224×10^{-14}	157.2	51.2

Table 3
Validation test nomenclature and conditions.

Test ID	Soil type (sand-silt-clay) [%]	Soil moisture [%]	Bulk density [kg m^{-3}]	Cone Index (CI_{0-150} - $CI_{150-300}$) [MPa]	Working speed [km h^{-1}]	Working depth [m]
RT1	-	-	-	-	Variable	-
RT2	-	-	-	-	Variable	-
RT3	-	-	-	-	Variable	-
FD1	Clay loam (34–36–30)	16.7 ± 1.7	1318.1 ± 143.9	0.627–0.792	Steady-state	-
FD2	Clay loam (34–36–30)	16.7 ± 1.7	1318.1 ± 143.9	0.627–0.792	Steady-state	-
FD3	Clay loam (34–36–30)	16.7 ± 1.7	1318.1 ± 143.9	0.627–0.792	Steady-state	-
FT1	Clay loam (34–36–30)	16.7 ± 1.7	1318.1 ± 143.9	0.627–0.792	5–6–7–8	0.25
FT2	Clay loam (34–36–30)	16.7 ± 1.7	1318.1 ± 143.9	0.627–0.792	5–6–7	0.50

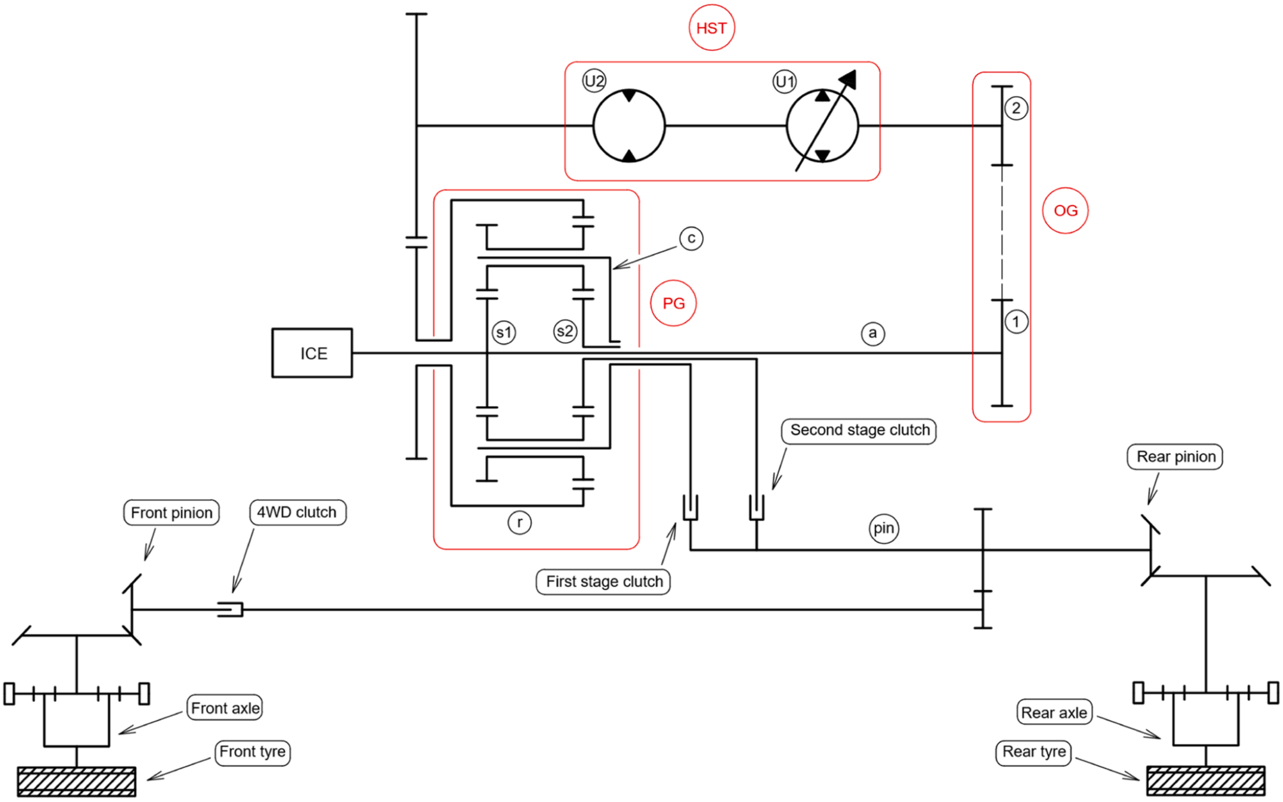


Fig. 2. Stick diagram of the dual-staged IHMCVT (input coupled hydromechanical continuously variable transmission), combined with the internal combustion engine (ICE) and the front and rear axles. Three primary components are shown: the ordinary gear (OG) with gears 1 and 2, a planetary gear (PG) composed of the carrier (c), ring (r), and two solar gears (s1 and s2), and the hydrostatic transmission (HST), with hydraulic units (U1 and U2). The connection shaft (a) and the rear pinion shaft (pin) are also shown.

modulating T_{ICE}, ω_{ICE}, and V_{th}. Indeed, considering the input signals T_{ICE} and ω_{ICE}, a pair of samples (T_{ICE}, ω_{ICE}) can identify the instantaneous operating point of the engine. The instantaneous operating point of the transmission is uniquely determined if the pin shaft speed ω_{pin} (which is proportional to V_{th}) is also known. In this way, the triplet (T_{ICE}, ω_{ICE}, ω_{pin}) can be used to identify the instantaneous transmission operating point via real-world data. For each sample of T_{ICE}, ω_{ICE}, and ω_{pin}, the results matrix $\bar{\bar{\omega}}_{pin}$ could be accessed at the transmission operative point (T_{ICE}, ω_{ICE}, ω_{pin}). The position of each working point in the $\bar{\bar{\omega}}_{pin}$ matrix was identified and then used to enter the results matrix $\bar{\bar{T}}_{pin}$ and extract the simulated pinion torque (T_{pin}^{mod}). To generalize the methodology, the drawbar force was calculated from the power flowing from the pin shaft to the driving axles. Indeed, the output power from the pinion was transferred to the tractive axles, and the modelled drawbar force (F_d^{mod}) was estimated via Eq. (5), and as a function of the traction efficiency η_{TE}.

$$F_d^{mod} = \eta_{TE} * \frac{P_{axle}}{V_{grd}} = \eta_{TE} * \frac{T_{pin}^{mod} * \omega_{pin} * \eta_m^n}{V_{grd}} \tag{5}$$

where P_{axle} is the power at the transmission outlet and available at the driving axles, η_m^n expresses the mechanical efficiency of the transmission between the pinion and the driving axles. Indeed, η_m is the mechanical efficiency of a single ordinary gear (assumed to be 0.97) accounting for friction losses, and n is the number of gearwheel ratios between the pinion and the driving axles, assumed to be 2 in 2WD mode and 3 in 4WD mode, given the transmission architecture. Thus, $\eta_m^n = 0.94$ when the transmission was in 2WD, while $\eta_m^n = 0.91$ when in 4WD. The assumption of constant gear efficiency is consistent with previous IHMCVT modelling studies [25–27], as also detailed in Appendix A. This approach addressed the problem in terms of power, avoiding the unknown percentage of power flowing to the front drive. η_{TE} is expressed following Eq. (6).

$$\eta_{TE} = \begin{cases} \eta_{TE}(s) = E * \exp(F * s) + G * \exp(H * s)\ if\ 4WD \\ 1\ if\ 2WD \end{cases} \tag{6}$$

where $\eta_{TE}(s)$ is a parametric regression function of the slippage $s \in [0 - 0.5]$ and E, G, F, and H are regression coefficients. This function, defined by Angelucci et al. [12], is applicable in cases of nonzero slippage. s was calculated as in Eq. (7).

$$s = \frac{V_{th} - V_{grd}}{V_{th}} \tag{7}$$

η_{TE} was thus assumed to be equal to 1 for 2WD operations. This choice was motivated by the very low slip typically observed during road transport (s was always below 0.02) compared to field conditions, where traction losses are negligible relative to other drivetrain losses. Under such circumstances, traction efficiency functions—which are strongly dependent on slip—become less accurate and less relevant as slip approaches zero [12]. Therefore, adopting a fixed efficiency ensures model simplicity without compromising accuracy for the intended operating scenario.

To improve clarity and to provide a compact overview of the quantities effectively entering the modelling equations, Table 5 summarizes the physical meaning, values, and sources of the parameters used in the IHMCVT model formulation. Only the parameters explicitly appearing in the analytical expressions discussed above are included.

The detailed kinematic parameters of the specific commercial IHMCVT used as a case study (individual gear ratios, hydrostatic unit displacements, maximum pressure and speeds, and maximum wheel torque) are not reported, as they pertain to the proprietary definition of the transmission. Consistent with previous IHMCVT modelling studies

Table 5
Physical meaning, values, and sources of the parameters.

Parameter	Physical meaning	Value	Source
η_m	Pinion-driving axles mechanical efficiency	0.97	Literature
η_{OG}	OG mechanical efficiency	0.97	Literature
η_{gw}	Planetary gear mechanical efficiency	0.97	Literature
n	Pinion-driving axles number of gearwheel ratios	2 (2WD), 3 (4WD)	Assumed from IHMCVT architecture
$\beta_0, \beta_1, \mu_{\omega_{ICE}}, \sigma_{\omega_{ICE}}(P_p)$	Regression coefficients for pre-charge pump power adsorption	See Table 4	Manufacturer bench-data
$\beta_0, \beta_1, \mu_{\omega_{ICE}}, \sigma_{\omega_{ICE}}(P_l)$	Regression coefficients for lube circuit power adsorption	See Table 4	Manufacturer bench-data
$\beta_0, \beta_1, \mu_{\omega_{ICE}}, \sigma_{\omega_{ICE}}(P_c)$	Regression coefficients for cooling circuit power adsorption	See Table 4	Manufacturer bench-data
$E,\ F,\ G,\ H$	Traction efficiency model coefficients	See [12]	Literature
λ	Fraction of U1 variable displacement	[+1,−1]	Physical definition

(e.g., [16,28]), the manuscript documents the overall modelling framework and all the physically relevant parameters employed in the simulations. The formulation presented in Section 2.1 and Appendix A is general and can be applied to different IHMCVT architectures by adapting the kinematic relations to the specific gear layout and using the corresponding machine-specific parameters.

2.2. Model validation and drawbar force analysis

2.2.1. Model validation assessment

The estimated drawbar force was then compared to the experimentally derived drawbar force. Measured WFT forces were considered positive when applied in the same direction as the motion of the tractor. The experimentally derived drawbar force was calculated differently for transport and pulling tests based on measured force signals. During a transport test, i.e., when the 2WD is engaged, the rear WFTs measure the net traction forces available at the rear axle. In contrast, the 4WD is engaged during pulling activities, allowing each tyre to contribute to traction. Therefore, the general equation used to express the total measured drawbar force when in 2WD ($F^{meas}_{d_{2WD}}$) and when in 4WD ($F^{meas}_{d_{4WD}}$) is defined in Eq. (8).

$$F^{meas}_{d_{jWD}} = \delta * (F_{x_{FL}} + F_{x_{FR}}) + F_{x_{RL}} + F_{x_{RR}} \tag{8}$$

with the transfer case selector signal δ being:

$$\delta = \begin{cases} 0 \ if \ 2WD \ (j = 2) \\ 1 \ if \ 4WD \ (j = 4) \end{cases}$$

The diagram in Fig. 3 identifies the model-derived physical quantities and the various combinations of measured physical quantities used to define the drawbar force when the tractor is in 2WD or 4WD mode.

The model goodness was then evaluated in two key areas: its ability to predict the drawbar force signal over time and its capability to correctly replicate load–unload cycles upon validation of the first aspect. The former details the temporal accuracy and amplitude of the signal, while the latter estimates the number of loading cycles of a certain amplitude regardless of their temporal sequence. For the first assessment, the absolute (e_{abs}) and relative error ($e_{\%}$) between the estimated and measured transmission gross drawbar forces were defined as shown by Eqs. (9) and (10).

$$e_{abs} = F^{mod}_{d} - F^{meas}_{d_{jWD}} \tag{9}$$

$$e_{\%} = \frac{e_{abs}}{F^{meas}_{d_{jWD}}} * 100 \tag{10}$$

Statistical indicators, such as interquartile range (IQR) and median (Med), were calculated from the absolute error distribution to evaluate the quality of the predictive model. These indicators were calculated as in Eq. (11) and Eq. (12).

$$IQR(e_{abs}) = P_{75}(e_{abs}) - P_{25}(e_{abs}) \tag{11}$$

$$Med(e_{abs}) = P_{50}(e_{abs}) \tag{12}$$

where $P_x(e_{abs})$ represents the value of e_{abs} at the x^{th} percentile of its distribution (i.e. 25th, 50th, and 75th). Moreover, the confidence intervals (CI) of *Med* and IQR were estimated using non-parametric bootstrap resampling. Specifically, 1000 bootstrap samples were drawn with replacement from the absolute error distribution, and the 2.5th and 97.5th percentiles of the resulting bootstrap distribution were taken as the bounds of the 95 % confidence interval.

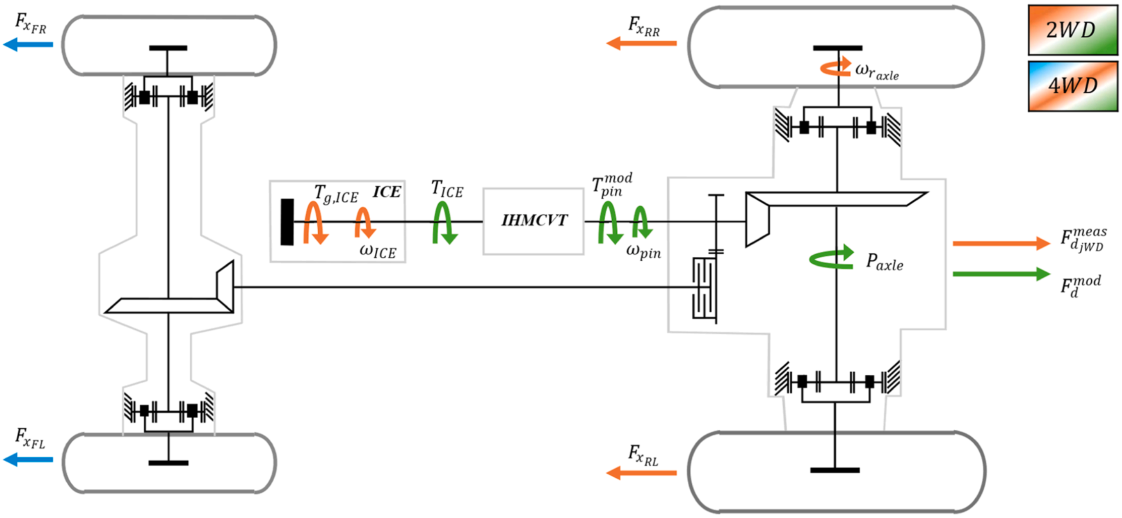


Fig. 3. Diagram of the analysed tractor showing the main measured and calculated physical quantities used in the validation and drawbar force mapping phases, as well as their positions in the kinematic diagram. Green vectors represent modelled quantities, while orange and blue denote measured or measurement-derived quantities. The colour combination of vectors and circular arrows indicates the quantities to consider, distinguishing when the tractor is in 2WD mode (orange and green vectors) and 4WD mode (orange, green, and blue vectors), as indicated in the legend at the top right of the diagram.

The load spectra (LS) of F_d^{mod} and $F_{d_{jWD}}^{meas}$ were calculated using the rainflow counting algorithm [29], which was set to discretize each load signal into 64 ranges and apply a rainflow filter to exclude cycles with amplitudes <2 % of the maximum value of the considered signal. This approach allowed for the visualisation of load counts and ensured that minimally damaging cycles were excluded. The LS were calculated by summing the number of cycles within each load range.

To statistically compare the cycle counts of the modelled signal with the measured ones, two indicators were defined (see Eq. (13) and Eq. (14)): the normalized root mean square error (NRMSE) and the percentage of total cycle accumulation (AC). Here, N_i^{mod} and N_i^{meas} denote the number of modelled and measured cycles, respectively, at the i^{th} drawbar force levels $F_{d,\ i}^{mod}$ and $F_{d_{jWD},\ i}^{meas}$, where $i = 1, \dots, 64$.

$$NRMSE = \frac{\sqrt{\mu\left(\left(N_i^{mod} - N_i^{meas}\right)^2\right)}}{\max\left(N_i^{meas}\right) - \min\left(N_i^{meas}\right)} \tag{13}$$

$$AC = \frac{\sum_i \left(N_i^{mod} * F_{d,\ i}^{mod}\right)}{\sum_i \left(N_i^{meas} * F_{d_{jWD},\ i}^{meas}\right)} \tag{14}$$

where μ is the arithmetic mean, max and min were used for normalisation over the range of the measured LS. The NRMSE quantifies the global estimation accuracy across the force spectrum, being more sensitive to large deviations and thus reflecting the relative closeness of the estimated values to the measured ones. To assess its variability, a 95 % CI was obtained using non-parametric bootstrap: 1000 bootstrap samples of the paired cycle counts were drawn with replacement, the NRMSE was recalculated for each sample using the same normalization factor ($\max\left(N_i^{meas}\right) - \min\left(N_i^{meas}\right)$), and the 2.5th and 97.5th percentiles of the resulting distribution were taken as CI bounds. In contrast, AC quantifies how closely the model reproduces the total accumulated force cycles relative to measurements. Unlike NRMSE, it does not reflect point-by-point accuracy but provides a measure of overall load accumulation fidelity. Both indicators were then computed for all RT and FT datasets to assess fatigue load accumulation under real-world operating conditions.

2.2.2. Real-world drawbar load analysis

The drivetrain model was employed to analyse an entire working season. Real-world data were first classified into operational states to determine whether the tractor was engaged in transport or field manoeuvres. Indeed, a previously developed classification algorithm was applied to the CANBUS dataset to identify typical tractor operations, specifically on-field passes, headland turns, transports, and idle activities, through the analysis of the time-based data [13]. Using the developed drivetrain model, the transmission output was then calculated, leading to the computation of the drawbar force signal and the extraction of typical drawbar force–vehicle speed plots. Furthermore, the vehicle specific fuel consumption (SFC) was calculated using Eq. (15) and Eq. (16).

$$SFC = \frac{\dot{f}_f * \rho_f\left(\theta_f\right)}{F_d^{mod} * V_{grd}} \tag{15}$$

with

$$\rho_f\left(\theta_f\right) = \sum_{k=0}^{3} d_{f_k} * \theta_f^k \tag{16}$$

where ρ_f is the fuel density calculated via a polynomial function of the fuel temperature, (θ_f) and d_{f_k} are the polynomial fitting coefficients of the diesel fuel (see Table 6). The function and coefficient considered for the calculation were used as proposed by Safarov et al. [30] and are related to the *Diesel fuel B0 2016* sample. The use of B0 properties is driven by their availability in validated analytical form and by their widespread adoption as reference fuel in engine-performance and fuel-consumption analyses (e.g., [31,32]). However, the formulation is general, and biodiesel blends can be incorporated by substituting their density–temperature relation and heating value, without any modification to the model structure.

Table 6

Polynomial fitting coefficient of Diesel *Hallen DK B0* sample.

Coefficients [unit]	Value ~
d_{f_0} [kg m^{-3}]	960.119153909765
d_{f_0} [kg m^{-3} K^{-1}]	$-$ 0.11446764457377
d_{f_0} [kg m^{-3} K^{-1}]	$-\ 0.140423325896083 * 10^{-2}$
d_{f_0} [kg m^{-3} K^{-1}]	$0.102003799041378 * 10^{-5}$

By associating each value of the SFC signal with the corresponding estimated drawbar force and vehicle speed, the arithmetic mean of the specific fuel consumption (μSFC) was calculated for each operating point on the drawbar force–vehicle speed plot. This methodology enabled the identification of the operating regions characterised by high fuel consumption values.

3. Results and discussion

3.1. Model validation

The signals acquired during the three representative test types (RT, FD and FT) exhibit markedly different dynamics, which are essential to correctly interpret the model validation results discussed in the following sections. Partial time-based traces and their detailed interpretation are provided in Appendix B (Fig. B.1), where the typical behaviours of the transmission working variables (ω_{ICE}, T_{ICE}, and V_{th}) and the corresponding ground speed V_{grd} are illustrated for RT, FD, and FT.

In brief, RT is characterised by moderate engine load variations and relatively smooth transitions between operating points, whereas FD and FT tests display stronger dynamic effects. In particular, the drawbar pulling test exhibits the highest torque amplitudes and the most pronounced variations in slippage, while the tillage test shows intermediate conditions with sustained torque levels and moderate fluctuations in ground speed. These aspects directly affect the drawbar force measured, $F_{d_{jWD}}^{meas}$, which, as discussed later, plays a central role in evaluating model performance.

To quantify these differences in force behaviour, Fig. 9 presents the kernel density estimations (KDEs) of both the measured drawbar force and its absolute temporal derivative for the representative datasets RT2, FD2, and FT2. These are given as examples since they show the widest range of measured forces, offering a thorough overview of operational conditions and their effects on the drawbar forces.

The quantitative indicators derived from these signals — namely the mean drawbar force, $\mu\left(F_{d_{jWD}}^{meas}\right)$, and the mean absolute temporal derivative, $\mu\left(\left|\dot{F}_{d_{jWD}}^{means}\right|\right)$ — were computed to summarise the differences in amplitude and dynamic behaviour among the tests. The temporal derivatives were obtained by numerical differentiation and filtered using a fourth-order Butterworth low-pass filter with a 1 Hz cut-off, applied in zero-phase mode to limit noise amplification while avoiding phase delays.

Each KDE is normalised to its maximum value, and the mean values of these indicators are reported as characteristic references. RT2 predominantly operates at low and steady drawbar loads, showing a sharp density peak around 7 kN and a narrow derivative distribution centred near zero, which indicates frequent steady-state conditions. Conversely, FD2 displays the broadest distributions of both quantities, with forces reaching up to 120 kN and high temporal variability, reflecting the strong dynamic effects induced by the tractor–tractor interaction and the increasing wheel slip. FT2 presents intermediate behaviour, with forces ranging between 50 and 100 kN and derivative values distributed more widely than in RT2 but less than in FD2.

These results confirm that both the amplitude and the rate of change of the drawbar force characterize the specific test conditions. Indeed, during high-load field operations, even small variations in speed can lead to significant changes in drawbar force due to the substantial reduction achieved by the transmission. Moreover, during fieldwork, the interaction between tractor wheels and soil generates nonlinear phenomena, including wheel slip and other dynamic responses. These are considerably more pronounced than those observed during road tests. Also, tillage operations produce variable responses through soil–implement interactions. Additionally, tests involving one tractor pulling another exhibit abrupt and dynamic variations in resistance load, attributed to internal vehicle responses and the connecting rope between the vehicles, further increasing the system dynamics.

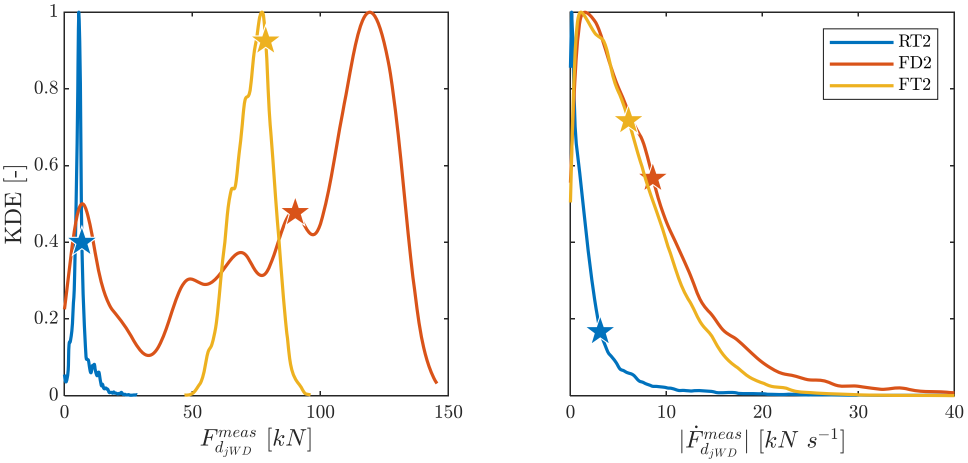
The coexistence of large force amplitudes and rapid signal variations, particularly in field tests, explains the higher complexity and potential modelling errors observed in those scenarios. Indeed, the performance of the model was assessed by defining the absolute and relative errors and using statistical indicators to examine data dispersion. The KDE of the absolute error was calculated for each dataset to show the error distribution, as presented in Fig. 5. The KDEs reveal that the error distribution of each test is typically a bell curve with a probability peak close to zero, which suggests that the model is sufficiently accurate and that errors are appropriately distributed for both positive and negative values. In RT tests, the error distribution is narrow and strongly concentrated around zero, with very short tails. This reflects limited variability in the signal and small modelling deviations, consistent with the absence of slip and the relatively smooth dynamics typical of transport conditions. In FD tests, the distribution becomes noticeably wider, and the peak shifts slightly further from zero. This behaviour is linked to the higher load variability and more pronounced transients induced by the tractor-tractor and tractor-soil interaction, which tend to amplify the dynamic effects acting on the system and consequently increase the model error. The FT tests show the broadest distribution, characterised by extended tails and a larger spread of error values, as in this case, the combination of variable slip, oscillations transmitted through the implement, and the nonlinear wheel–soil response results in more complex operating conditions. It is also worth noticing that FD2 exhibits the peak with the highest absolute deviation from zero, measuring 2.3 kN. This corresponds to a loss in the vehicle traction ratio (VTR) of only 0.012, as defined by Angelucci et al. [12].

The bisector plots in Fig. 5d–f complement the absolute-error distributions shown in Fig. 5a–c by illustrating how estimated and measured drawbar forces relate across the different operating scenarios. In all cases, the dispersion of the points remains generally centred around the bisector, confirming the absence of systematic bias. The spread patterns are consistent with the behaviour observed in the KDEs: RT tests exhibit a comparatively wider lateral spread despite their low absolute errors, an expected effect of the small force range, where even minor deviations appear as visible offsets from the diagonal. FT datasets exhibit the highest dispersion, reflecting their sustained load levels and moderate-to-high dynamical variability. By contrast, FD tests show the most compact arrangement of points, as the higher forces lead to errors that grow proportionally with the load and therefore align predominantly along the bisector rather than diverging from it. Overall, the three plots corroborate the trends highlighted by the error distributions and reveal a clear heteroscedastic structure of the estimation error across operating conditions.

To characterize the drawbar force signals $F^{meas}_{d_{jWD}}$ of each dataset, the mean values and the mean absolute temporal derivative of these forces were computed. These metrics quantify both the magnitude and change rate of the signal (e.g. in Fig. 4 for RT2, FD2, and FT2) over time via two parameters. Fig. 6 illustrates the distribution of errors relative to signal characteristics by combining the mean and mean absolute derivative of $F^{meas}_{d_{jWD}}$ with the error indicators (i.e. *Med* and *IQR*) calculated over the model outputs when predicting $F^{meas}_{d_{jWD}}$.

Field tests yielded high μ values, ranging from 74 to 94 kN (corresponding VTR from 0.46 to 0.59), and correspondingly high mean absolute temporal derivative values, with mean force variation rates spanning from 6 to 8.8 kN s^{-1}. Conversely, road tests demonstrated significantly lower μ values, consistently under 8 kN. Dynamic variations were notably reduced, with force derivative rates around 3 kN s^{-1}, approximately half to one-third of the rate of change observed in field tests.

Fig. 6 also shows a correlation between the two variables and a clear distinction between field and road test data, whereas Table 7 reports a summary of the error indicators, *Med* and *IQR* together with their confidence intervals (CI), and the global performance for each test type.



(a)

(b)

Fig. 4. Kernel density estimation (KDE) of (a) measured drawbar force ($F^{meas}_{d_{jWD}}$) and (b) absolute temporal derivative $\left(|\dot{F}^{means}_{d_{jWD}}|\right)$ of measured drawbar force for datasets RT2, FD2, and FT2. Each curve is normalised to 1 using its maximum value. For each dataset, the points corresponding to the mean (μ) of $F^{meas}_{d_{jWD}}$ and of $|\dot{F}^{means}_{d_{jWD}}|$ are indicated by stars.

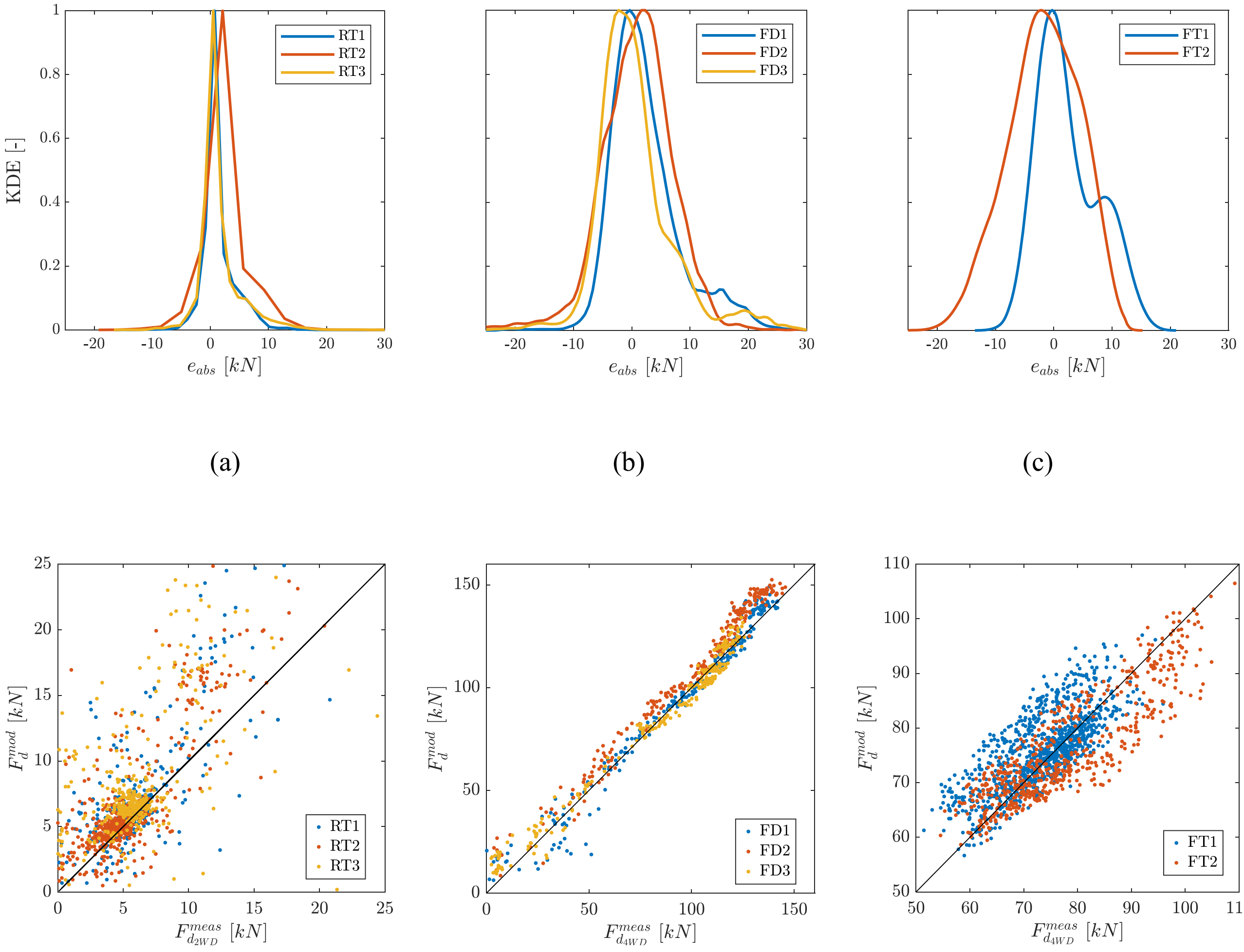
Table 7
Statistical error indicators from validation tests.

Test ID	*Med* [kN]	95 % CI of *Med* [kN]	*IQR* [kN]	95 % CI of *IQR* [kN]	mean |*Med*| [kN]	mean |*IQR*| [kN]
RT1	0.595	[0.587, 0.602]	2.109	[2.084, 2.145]	~0.67	~2.35
RT2	0.863	[0.853, 0.873]	2.413	[2.387, 2.445]		
RT3	0.538	[0.525, 0.551]	2.539	[2.491, 2.580]		
FD1	1.212	[1.158, 1.275]	7.519	[7.442, 7.605]	~0.98	~6.88
FD2	1.199	[1.147, 1.256]	6.619	[6.539, 6.696]		
FD3	−0.521	[−0.576, −0.469]	6.520	[6.432, 6.590]		
FT1	1.120	[1.094, 1.145]	7.765	[7.705, 7.832]	~1.36	~8.12
FT2	−1.599	[−1.642, −1.558]	8.480	[8.424, 8.541]		

From Fig. 6, road tests cluster in the lower left quadrant, exhibiting low signal magnitudes and variation rates, with comparably low IQR values around 2.25 kN. Moreover, the median values remain lower than 0.86 kN for each RT (this is confirmed by Fig. 5, where the KDEs show a fairly symmetric distribution with peaks close to zero error), underscoring a consistent response trend of the model across all RTs.

In contrast, field tests predominantly occupy the upper right quadrant, characterised by high signal magnitudes and high variation rates, with median values ranging between −1.60 and 2.31 kN and IQR values approximately three times higher than those of the road tests (6.50–8.50 kN). These results are consistent across field tests and reflect the real-world conditions of agricultural operations. The contrast in signal characteristics and error indicators between field and road conditions underscores the impact of dynamic loads and variable resistance encountered in practice. Such variability increases the susceptibility of the model to uncertainty, as the transmission model is static for computational reasons. Nevertheless, the analysis of 95 % confidence intervals for both median and IQR (see Table 7) confirms that the central tendency of the absolute error remains robust across all tests. While road tests exhibit narrow IQR and CI, indicating high precision, field tests show higher dispersion but with tight CI bounds, suggesting that these larger errors and occasional biases are systematic and reproducible rather than random.

The following considerations demonstrate the ability of the model to predict load accumulation. Although these insights apply to all RT and FT test datasets, the plots and results specifically refer to the RT2 and FT1 datasets as representative examples. The RT2 dataset was selected because it represents a typical mass condition frequently encountered in transport. The FT1 test was selected due to its richer dataset, as it includes an 8 km h^{-1} field pass and the speed repetitions found in the FT2



(a)

(b)

(c)

(d)

(e)

(f)

Fig. 5. Kernel density estimation (KDE) of absolute error (e_{abs}) for each dataset: (a) RT, (b) FD, and (c) FT tests. Each KDE curve is normalised to 1.

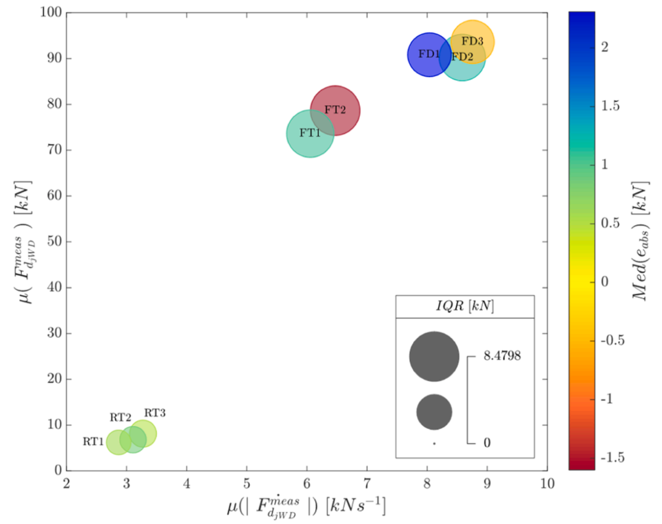


Fig. 6. Comparison of model absolute error median (*Med*) and interquartile range (*IQR*) of RT, FD, and FT tests based on the mean values of $F^{meas}_{d_j}$ and the mean absolute temporal derivative of $F^{meas}_{d_j}$. For each point, the inner circle represents *Med*, and the outer the *IQR*. Road tests cluster in the lower left quadrant, exhibiting low signal amplitudes and variation rates, while field tests show the opposite behaviour.

dataset. This resulted in a more substantial load accumulation. Fig. 7 contains two subplots, each depicting the load spectra calculated from the drawbar force signals, measured ($F^{meas}_{d_{jWD}}$) and modelled (F^{mod}_{d}), and normalized via the range of the measured LS (i.e., the normalisation done in Eq. (13) was applied to these plots).

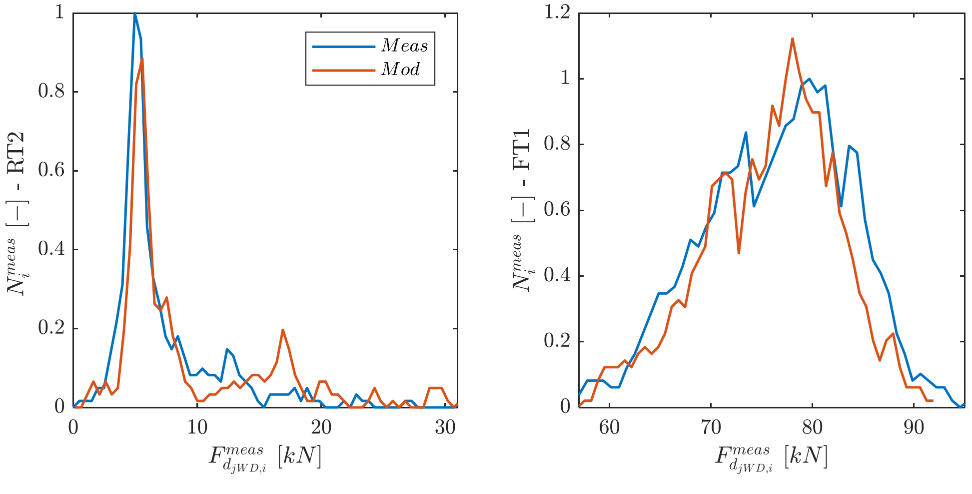
The LS in Fig. 7a, corresponding to the RT2 dataset, shows a similar distribution, highlighting the accuracy of drawbar force prediction. Both show a distinct peak around the 5 kN range, indicating that mainly low-amplitude cycles were experienced during road transport. These conditions are mechanically undemanding and contribute minimally to fatigue accumulation. In contrast, Fig. 7b, corresponding to the FT1 dataset, displays a broader and more variable distribution of load counts across a wider range of forces, from approximately 55 to 95 kN, indicative of the more dynamic and fluctuating nature of forces experienced during tillage operations. Such high-load cycles are those primarily responsible for fatigue damage and reflect the substantially more severe loading environment typical of soil-engaging operations. From an energy standpoint, these load ranges require markedly higher power, consistent with the increased traction demand typical of field work. For ranges above 83 kN, the model underestimates the cycles, which indicates the greater difficulty experienced by the static model as the cycle amplitude increases. Tillage operations are characterised not only by higher force variability compared to road transport but also by a higher frequency of load reversals. FT1 exhibited an average number of cycles (i.e., load reversals) per unit time of around 33 cycles/min, while RT2 showed a cycle count of 21 cycles/min (the corresponding modelled number of cycles (n_{cy}) were 34 and 20 cycles/min, respectively).

Considering now Table 8, the NRMSE between the estimated and measured LS was calculated (via Eq. (13)) as 0.068 for RT2 and 0.083 for FT1, suggesting that the model performs reasonably well for both scenarios, with slightly higher accuracy for road transport. The 95 % confidence intervals of the NRMSE indicate that the global estimation accuracy is robust, even under the more variable field conditions. Furthermore, AC (Eq. (14)) was 1.04 for RT2 and 0.98 for FT1. For RT2, the estimated load accumulation exceeded the measured accumulation by 4 %, whereas for FT1, the estimated load accumulation was very close to the measured value, with only a 2 % difference. This close agreement in FT1 demonstrates that the model accurately captures the overall load accumulation during tillage operations, despite the higher variability in the signal. Overall, these results highlight the effectiveness and robustness of the model in capturing the key characteristics of LS across both real-world operational scenarios.

Overall, the error metrics and the close agreement between the estimated and measured load spectra (LS) across multiple operational scenarios (road, drawbar, and tillage) indicate that the methodology provides a robust and reliable estimation of tractor drawbar loads under diverse real-world conditions.

Table 8
Statistical LS indicators for the most representative experiments.

Test ID	NRMSE [-]	95 % CI of NRMSE [-]	AC [-]	n_{cy} [cycles/min]
RT2	0.068	[0.050, 0.088]	1.04	20
FT1	0.083	[0.057, 0.111]	0.98	34



(a)

(b)

Fig. 7. Comparison between the load spectra (LS) of the measured ($F^{meas}_{d_{jWD}}$) (blue) and modelled (F^{mod}_{d}) (red) drawbar forces, for (a) RT2 dataset and (b) FT1 dataset. For each test, the number of cycles is normalised to the range of the measured LS.

3.2. Real-world drawbar load analysis

Fig. 8 reports the drawbar load data with the output of the classification algorithm. In particular, Fig. 8a reports the classification mode for each combination of F_d^{mod} and V_{grd}, highlighting three main regions. As a result of the classification, 54.6 % of the data (around 230 h) was identified as on-field passes, 29.5 % (around 124 h) as transport, and 15.9 % (around 67 h) as headland turns. On-field passes started from around 35 kN of drawbar force and speeds lower than 10 km h^{-1}. In contrast, the transport condition typically covered the entire speed span but for medium-low drawbar forces. A smaller area was identified at low speeds and low-to-medium drawbar forces corresponding to headland turns.

For each combination of F_d^{mod} and V_{grd}, the percentage of points classified as the corresponding modal operating status was calculated, and the result is shown in Fig. 8b Most of the traction domain, around 58 %, has a classification percentage exceeding 90 %. Notably, there is a strong correspondence between on-field passes and transport activity areas and these high-scoring zones.

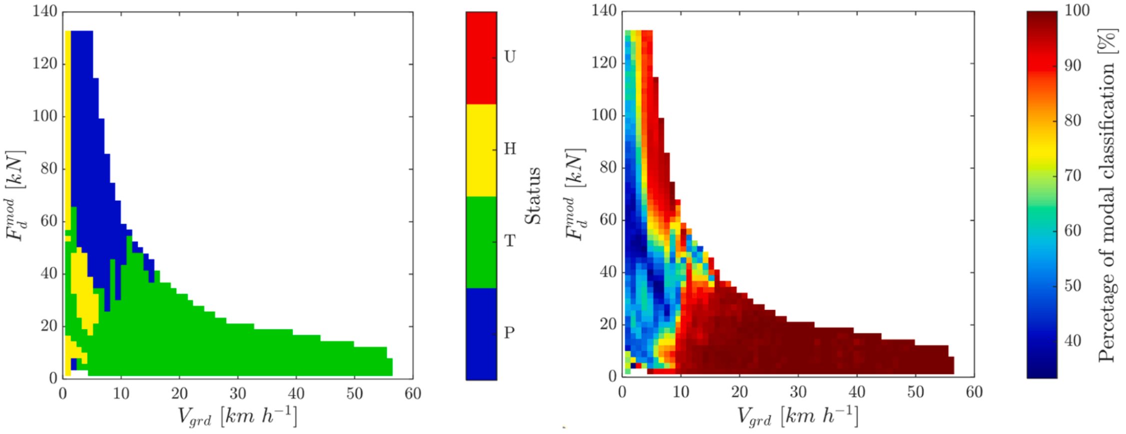
Most areas with lower classification percentages (the lowest classification percentage remains around 33 %) occur at the boundaries between different status regions, particularly at low speed and low-to-high drawbar force levels. This larger variability at boundaries was expected due to the wide diversity of agricultural tasks performed at those force–speed combinations.

To advance the analysis, the frequency histogram of traction working points was composed and represented in Fig. 9a. The analysis identified a significant area of high combined frequency for V_{grd} below 9 km h^{-1} across the entire drawbar force range. This area, encompassing both on-field drawbar activities and transport tasks, can be subdivided into two distinct high-frequency regions. The first region is characterised by speeds lower than 5 km h^{-1} and drawbar forces below 100 kN. The second, and most frequent in global terms, is located between 5 and 9 km h^{-1} with drawbar forces ranging from around 80 to 115 kN. The most frequent working point is 7 km h^{-1} with an F_d^{mod} of approximately 94 kN. These observations are reflected in the V_{grd}. The KDE plot (Fig. 9b) shows two prominent peaks at about 2 and 7 km h^{-1}. The first one corresponds to lighter activities, while the second peak (which is also the global maxima), at drawbar forces of around 80–115 kN, indicates more intensive soil activities, such as deep ploughing.

This latter peak coincides with the most frequent peak on the frequency map and is six times more prevalent than the most common transmission working points in the first peak area, indicating that the tractor was mainly used for heavy ploughing–tillage operations. A third frequent region can be recognized at approximately 6 kN stretching over the entire speed range, with a final peak at around 55 km h^{-1} (the KDE of the estimated drawbar force (F_d^{mod}) is presented in Fig. 9b). This area represents frequent transport activities.

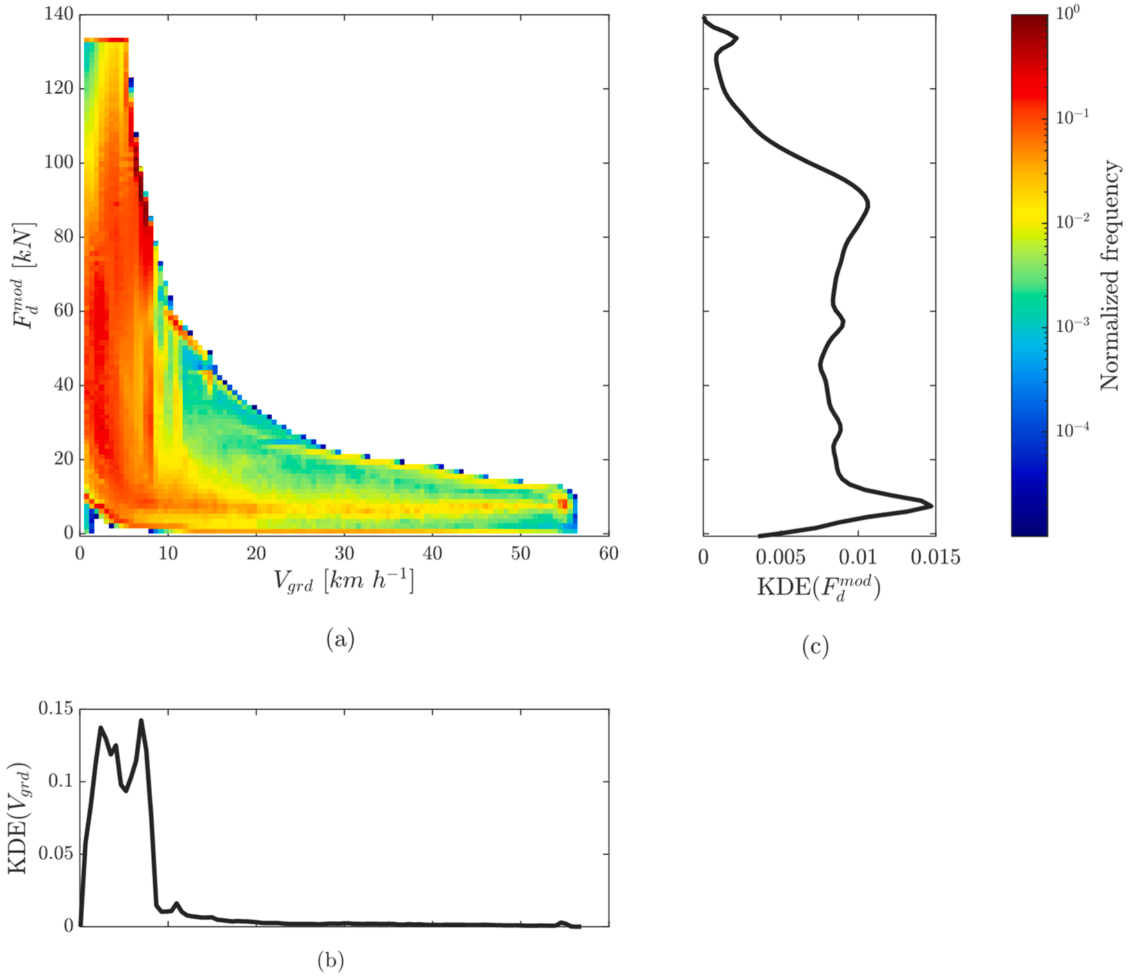
Before proceeding with the consumption map, it is useful to compare the overall ranges of *SFC* for the engine and for the vehicle. The temperature θ_f of the fuel measured during the tests ranged from 280 to 355 K. The engine *SFC* ranges approximately from 160 to 310 g kWh^{-1}, while the vehicle *SFC*, calculated from drawbar power, spans from about 190 to 380 g kWh^{-1}. The broader and higher range of the vehicle *SFC* reflects the additional losses introduced by the transmission, as well as the variability in hydraulic and mechanical efficiencies. Fig. 10 displays a map of operating points, F_d^{mod}vs V_{grd}, where each point shows the arithmetic mean specific fuel consumption of the vehicle (μSFC), highlighting the least efficient regions of the vehicle. The value for each operating point was obtained by averaging all vehicle *SFC* values belonging to that point. The plot demonstrates a significant variability in the μSFC across the traction chart. This variability can be attributed to the wide range of engine working points (T_{ICE}, ω_{ICE}) that can achieve a specific drawbar force at a given tractor speed. As the engine working point changes, so does its specific fuel consumption. Additionally, a particular engine working point and a tractor speed point uniquely determine the HST working point (i.e., the transmission working point) and its efficiencies, thereby affecting the traction force and the *SFC*.

This dual nonlinear effect highlights the necessity of defining a mean specific consumption for each point on the traction chart. The chart shows that the μSFC minima are located in the most frequently operated regions (see Fig. 9), with the global minimum at approximately 6 km h^{-1} and 98 kN, and having μSFC equal to 283 g kWh^{-1}. Conversely, three main high-consumption regions were identified. The first region is characterized by high F_d^{mod} (between 80 and 130 kN) and V_{grd} lower than 5.5 km h^{-1}, coinciding with the HST high-pressure working condition, which is marked by medium-to-low hydromechanical efficiencies of U1 and U2 and, more importantly, by low volumetric efficiencies of the two hydraulic machines. This is evident from Fig. A.1, where volumetric efficiency decreases as pressure increases. The second high-consumption area is located centrally, between 7 and 12 km h^{-1}, with drawbar forces lower than 67 kN. This region can be explained by considering the total efficiency plot of an IHMCVT [15,25]. For a particular working point (T_{ICE}, ω_{ICE}), when extracting power from the first stage, the transmission efficiency follows a typical bell shape as the V_{grd} increases, with the curve increasing, reaching a maximum, and then decreasing. As the V_{grd}



(a) (b)

Fig. 8. Modelled drawbar force (F_d^{mod}) versus ground speed (V_{grd}) charts showing (a) modal status classification and (b) percentage of points classified as the modal status. In (a), the different colours correspond to the operating status of the tractor (P: on-field passes, T: transport, H: headland turn, and U: undetected status). Each chart is discretised in a 60 × 60 grid.



(a)

(c)

(b)

Fig. 9. Frequency histograms of the transmission working point for (a) the one-year dataset. On the top and at the right of the histogram, the kernel density estimation (KDE) of (b) the estimated drawbar force (F_d^{mod}) and (c) the ground speed (V_{grd}). The frequency histogram is discretised in a 100 × 100 grid.

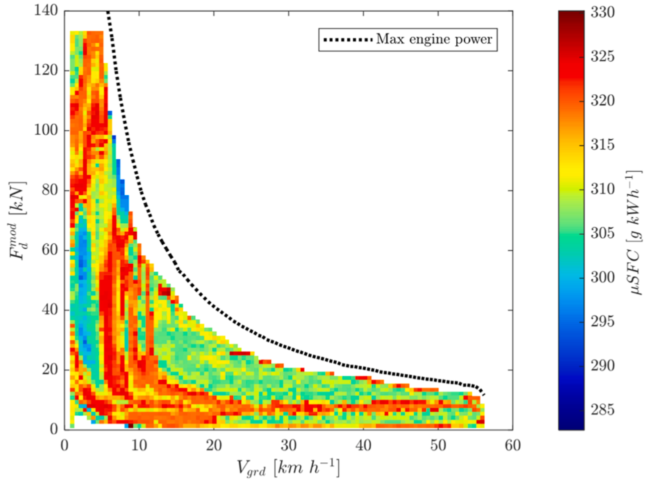


Fig. 10. Modelled drawbar force (F_d^{mod}) versus ground speed (V_{grd}), showing the arithmetic mean specific fuel consumption (μSFC). The chart is discretised in a 100 × 100 grid to accurately describe the traction graph.

increases, the transmission has to shift from the first stage to the second. Again, the efficiency follows a typical bell shape, with the curve rising, peaking, and declining. The shift from the first to the second stage occurs at the shifting speed V_{grd}, which is a function of ω_{ICE} [15].

Despite synchronization, the junction between the efficiency curves is characterised by a low-efficiency valley or, in the worst case, by an efficiency discontinuity (e.g. this valley is evident in the efficiency plots from Pettersson [15], where the transition between curves of hydro-mechanical stages always exhibits an efficiency drop). Also, under identical IHMCVT design parameters, the shape of the efficiency curves is determined by both T_{ICE} and ω_{ICE}. For a constant T_{ICE}, as ω_{ICE} increases, these curves expand horizontally without efficiency variations, thereby increasing the maximum achievable speeds of the tractor. Due to this expansion, the first efficiency curve tends to increase more slowly (at a given V_{grd}, the total efficiency will be lower), so the transmission total efficiencies are adversely affected for high ω_{ICE} and low V_{grd} (typical of high drawbar force activities). This confirms the previously discussed region of high loads and speeds below 5.5 km h^{-1} but, more importantly, as ω_{ICE} increases, the expansion of the curves moves the low-efficiency shift between stages towards higher V_{grd} (the shift point moves from a minimum of around 7 km h^{-1} to a maximum of 12 km h^{-1}), justifying the low efficiencies observed in this interval.

Continuing the analysis, the low efficiencies in the region mentioned above tend to shrink and concentrate in a narrow area, which then stretches along the entire V_{grd} range for drawbar forces around 6 kN. This identifies the third low-efficiency region and can be explained by considering the behaviour of the total efficiency curves of the transmission for constant ω_{ICE} and varying T_{ICE}. Given that this area is typical of low drawbar force activities (such as transport), when considering a constant ω_{ICE} and as the T_{ICE} decreases, the efficiency curves maintain the same vehicle speed span but tend to dilate downward. This effect creates pronounced valleys at the shift points between stages, justifying the general decrease in efficiency across the range of V_{grd}.

Finally, in Fig. 10, the theoretical traction curve at maximum power is plotted with a dashed line. This traction curve represents the drawbar force at maximum pulling power, assuming an ideal IHMCVT with 100 % efficiency from the engine output to the tractor wheels. The plot illustrates the extent of the reduction in drawbar power resulting from the inevitable losses downstream of the engine to the wheels (largely due to HST inefficiencies). The most significant reduction occurs between 6 and 9 km h^{-1}, where the actual maximum power traction curve falls short of the theoretical curve by approximately 21 kN, corresponding to an 18 % drawbar force loss. This speed range, characterised by the greatest inefficiency, also contains the most frequent operating points (see Fig. 9a). Consequently, frequently operating within this range is particularly disadvantageous, given the substantial reduction in drawbar capacity.

4. Conclusions

In this study, a model and a real-world dataset were used to determine the drawbar load of an agricultural tractor under real operating conditions. The approach relies on standard CANBUS signals, enabling the extraction of load information with simple, non-intrusive instrumentation suitable for long-term deployment across a fleet of vehicles. This overcomes the limitations of traditional sensors, such as wheel force transducers, which are often bulky, expensive, and not robust enough for extended use.

The model was validated through dedicated tests under various operating conditions, including road transport, towing, and tillage. Despite being based on steady-state equations, it achieved relative errors mostly within ±10 %, demonstrating sufficient accuracy for practical applications. Indeed, when the median error is expressed relative to the typical force level of each test category, the model shows an accuracy compatible with practical use: approximately 8 % for RT tests, about 1–1.2 % for FD tests, and roughly 1.5–2 % for FT tests. Its low computational demand allows implementation directly on embedded tractor hardware, supporting real-time traction optimization in response to changing soil conditions.

Although the long-term dataset was collected from a single machine, the model is parametric and designed to apply to any tractor with the same transmission architecture. This makes it suitable for large-scale deployment, enabling the analysis of drawbar load across a wide range of real-world conditions. Such scalability is particularly relevant for mission profiling, fatigue and damage assessment, and predictive maintenance.

Some operating conditions, such as PTO-active phases, were excluded from the dataset, accounting for roughly 17 % of total operating time. This introduces a partial bias, but future developments could address it. The increasing availability of CAN-based or dedicated sensors for PTO power measurement could be integrated into the model, allowing net power transmitted to the drivetrain to be computed by subtracting PTO power from gross input power, without altering the core structure.

The current model is static but provides a solid foundation for future developments. Incorporating dynamic behavior would enhance responsiveness and accuracy during rapid load changes, such as acceleration or gear shifting, increasing its applicability in real-time control and simulation. Extending the methodology to tractors with different transmission configurations will require standardized interfaces between existing sub-models for components such as planetary gears, conventional gears, and hydrostatic transmissions. A modular design of this type would allow adaptation to a wide range of IHMCVT architectures. Integrating a wheel–soil interaction model based on terramechanics would further improve simulation realism and accuracy. Compared with black-box approaches, this model-based strategy offers greater transparency and interpretability while avoiding costly, vehicle-specific experimental campaigns.

Looking forward, this methodology could support the development of reference duty cycles for agricultural tractors, similar to those used in the automotive sector. Such data could inform regulatory testing, drivetrain optimization, and fleet-level energy and emissions assessments. By enabling scalable, data-driven insights into tractor operation, the approach has the potential to guide future design strategies, improve fleet management, and accelerate the transition to more sustainable agricultural machinery.

Ethics statement

Not applicable: This manuscript does not include human or animal research.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used Microsoft 365 Copilot in order to assist with English grammar and language correction. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

CRediT authorship contribution statement

Luca Colendi: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Manuel Tentarelli:** Writing – review & editing. **Massimiliano Varani:** Writing – review & editing. **Michele Mattetti:** Writing – review & editing, Supervision, Resources, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Data collection was carried out within “*DATA-BUS - Digital Agriculture Technology to Achieve data to Build User-friendly Sustainability indicators*” supported by MUR (Ministry of University and Research) under the call PRIN (Research projects of significant national interest) notification 2020. Grant number: 2020SCNF4L. Data analysis was carried out within the Agritech National Research Center and received funding from the European Union Next-GenerationEU (Piano Nazionale di Ripresa e Resilienza (PNRR) – missione 4 componente 2, investimento 1.4 – D.D. 1032 June 17, 2022, CN00000022).

Appendix A

This section provides a detailed treatment of the static mathematical transmission model.

IHMCVT equations

The definition of the IHMCVT equations first focused on gathering the constructive characteristics of the drivetrain under study, determining the arrangement of the gears and subsystems within it.

The following discussion outlines the physical relationships utilised in developing the IHMCVT model. The kinematic relations between the various transmission components are presented first.

The kinematic relation of the OG of the IHMCVT is defined as the speed ratio between the two gearwheels (τ_{OG}), shown in Eq. (A.1)

$$\tau_{OG} = \frac{\omega_2}{\omega_1} \tag{A.1}$$

where ω_1 and ω_2 are the speeds of the OG gearwheels.

The kinematic relation of the PG first stage consists of the Willis formula, which is stated in Eq. (A.2)

$$\tau_I = \frac{\omega_r - \omega_c}{\omega_{s1} - \omega_c} \tag{A.2}$$

where τ_I is the speed ratio of the PG, and ω_r, ω_c, and ω_{s1} are the speeds of r, c, and s1, respectively. The transmission ratio τ_{II} of the second stage was obtained similarly.

Given an HST made of two hydraulic units, one of which has variable displacement, the speed ratio (K_ω) of the two machines was calculated through their flow rates with Eq. (A.3).

$$K_\omega = \frac{\omega_{u2}}{\omega_{u1}} = \alpha * \frac{\lambda * D_{u1}}{D_{u2}} * \left(\eta_{vol_{u1}} * \eta_{vol_{u2}}\right)^{sign(-\lambda)} \tag{A.3}$$

where ω_{u1} and ω_{u2} are the speeds of the two machines; D_{u1} and D_{u2} are their displacements; $\eta_{vol_{u1}}$ and $\eta_{vol_{u2}}$ are their volumetric efficiencies; $\lambda \in [+1, -1]$ is a parameter representing both the fraction of variable displacement selected and the hydraulic unit swash plate position; and α is a parameter equal to $+1$ when the power is extracted from the first stage and equal to -1 when the power is extracted from the second stage. Hence, λ varies from $+1$ to -1 for the first stage working condition and from -1 to $+1$ for the second stage. The *sign* function was used to identify the hydraulic circuit working mode, i.e. the power direction, for a particular swash plate position.

Having just explained the kinematic relations, the mechanical equations governing the system can now be stated. The mechanical relation of the OG of the IHMCVT can be written as in Eq. (A.4).

$$\eta_{OG}^{sign(-\lambda)} = \frac{T_2}{T_1} * \tau_{OG} \tag{A.4}$$

where η_{OG} is the mechanical efficiency of the OG, which was assumed to be constant and equal to 0.97. This choice can be considered sufficiently conservative, as typical values of OG efficiency are often higher [33], and adopting a constant value is a common practice in IHMCVT studies [25–27]. T_1 and T_2 are the torques of the OG gearwheels, and the *sign* function was once again exploited to identify the hydraulic circuit working mode because the power direction in the OG changes according to the IHMCVT working phase.

The Willis formula was used to determine the torque relations of a loss-free dual-stage PG, resulting in Eq. (A.5) Eq. (A.5) for the PG first stage and Eq. (A.6) Eq. (A.6) for the second one.

$$\tau_I = -\frac{T_{s1}}{T_r} \tag{A.5}$$

$$\frac{\tau_I * (\tau_{II} - 1)}{\tau_{II} - \tau_I} = -\frac{T_{s1}}{T_r} \tag{A.6}$$

For the HST, assuming the operating pressure of the two hydraulic machines is the same, we can define the torque ratio of the two hydraulic units (K_T) for the PG first and second stages, as expressed in Eq. (A.7).

$$K_T = \frac{T_{u1}}{T_{u2}} = \alpha * \frac{\lambda * D_{u1}}{D_{u2}} * \left(\eta_{m_{u1}} * \eta_{m_{u2}}\right)^{sign(+\lambda)} \tag{A.7}$$

where T_{u1} and T_{u2} are the torques of the machines, $\eta_{m_{u1}}$ and $\eta_{m_{u2}}$ are their hydromechanical efficiencies. It is now possible to express, for example for the first stage, the relation between the pressure difference across the hydrostatic unit U2 (p) and its output torque as in Eq. (A.8):

$$p = \frac{2\pi * T_{u2}}{D_{u2} * \eta_{mec_{u2}}^{sign(+\lambda)} * 10^6} \tag{A.8}$$

Given the above kinematic and mechanical relations, two power balances can be expressed to solve the system of steady-state equations for the unknown velocities and torques. The first power balance is calculated at transmission shaft (a) connecting ICE, OG, and PG and can be written as Eq. (A.9).

$$(T_{ICE} + T_1 + T_{s1}) * \omega_{ICE} = 0 \tag{A.9}$$

The second balance concerns the power distribution between the PG components, changing according to the transmission-selected PG stage. Below are the power balances for the first and second stages, respectively, Eq. (A.10) and Eq. (A.11).

$$T_c * \omega_c + T_{s1} * \omega_{s1} * \eta_{gw} + T_r * \omega_r * \eta_{gw}{}^{sign(-\lambda)} = 0 \tag{A.10}$$

$$T_{s2} * \omega_{s2} + T_{s1} * \omega_{s1} * \eta_{gw} + T_r * \omega_r * \eta_{gw}{}^{sign(-\lambda)} = 0 \tag{A.11}$$

where η_{gw} is the mechanical efficiency related to each gearwheel contact, which was assumed to be equal to 0.97, similarly to the assumption made for η_{OG}. Again, the last term in both equations employs the *sign* function to determine the direction of power through the ring according to the transmission operative mode. Progressing further in the driveline, if the first stage is selected, the torque assigned to the pinion shaft (pin) will be T_c. If the second stage is engaged, the torque on the shaft will be T_{s2} (see Fig. 2). Thus, the modelled pin shaft torque is referred to as T_{pin}^{mod}.

Hydraulic unit efficiencies and calculation algorithm

The hydraulic unit efficiencies are significantly influenced by their operating conditions and exhibit nonlinear dependencies on three key parameters: speed, pressure, and displacement. Fig. A1 shows the volumetric efficiency map of U1 at maximum displacement (i.e., $\lambda = 1$).

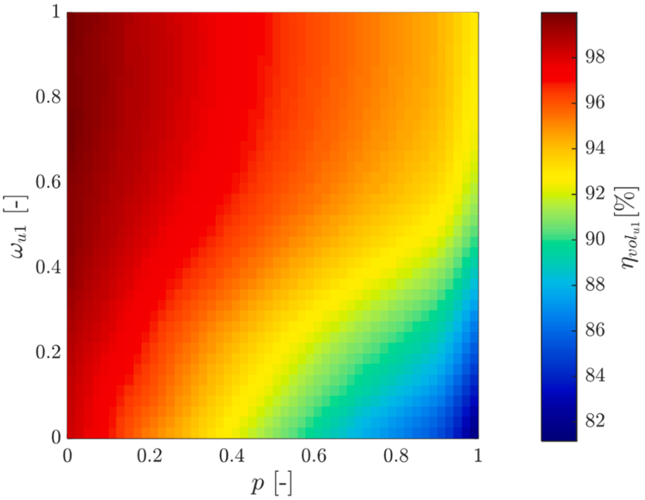


Fig. A1. Volumetric efficiency map of unit 1 ($\eta_{vol_{u1}}$) at maximum displacement (i.e., parameter λ=1), and as a function of normalised unit 1 speed (ω_{u1}) and pressure (p).

It was thus necessary to define functions (Eq. (A.12)) in the form of four shallow neural networks (SNN) to express these volumetric and mechanical efficiencies and integrate them into the model calculation.

$$\begin{aligned} \eta_{vol_{u1}} &= f(p, D_{u1}, \omega_{u1}) \\ \eta_{mec_{u1}} &= f(p, D_{u1}, \omega_{u1}) \\ \eta_{vol_{u2}} &= f(p, \omega_{u2}) \\ \eta_{mec_{u2}} &= f(p, \omega_{u2}) \end{aligned} \tag{A.12}$$

The original efficiency maps were provided by the manufacturer and defined on a three-dimensional grid (pressure p, displacement D_{u1}, speed ω_{u1}) for U1 and on a two-dimensional grid (pressure p and speed ω_{u2}) for U2, which had no variable displacement. Each variable was discretized by the manufacturer into 100 points per dimension, resulting in approximately 10^6 operative points for U1 and 10^4 for U2. These experimental maps covered the full operating domain of the hydraulic units. Each independent variable x^* was then normalized into $\hat{x}^*$ via its maximum x_{max} and minimum x_{min} values (using the formula $\hat{x}^* = (x^* - x_{min})/(x_{max} - x_{min})$) and rescaled to the [0, 1] interval before being fed into the neural networks. For confidentiality reasons, the absolute values of maximum pressure (p_{max}), maximum speeds (ω_{u1max} and ω_{u2max}) and displacements (D_{u1} and D_{u2}) cannot be disclosed. The normalization of all input variables to the [0,1] interval preserves the full shape and nonlinear gradients of the original efficiency maps, allowing the SNN-based regression model to be reproduced without requiring access to the undisclosed dimensional limits. This modelling strategy is consistent with the approach commonly adopted in previous studies (e.g., [16,28], in which the same IHMCVT was investigated), in which manufacturer-supplied efficiency maps are employed in normalized form. Moreover, technical specifications of comparable hydrostatic units (e.g., maximum pressure, speed limits, and displacements) are available in the open literature (e.g., [27]) and may be used as reference input data when evaluating similar transmission models.

Each network was fully connected and composed of a single hidden layer. For each SNN, the dataset was randomly divided into 70 % for training and 30 % for validation. The random split was performed using a fixed seed (rng(42) in MATLAB) to ensure reproducibility of the training and validation sets. The training function used was the mean squared error (MSE), with optimization performed via the Levenberg–Marquardt algorithm, a hyperbolic tangent sigmoid activation function for the hidden layer, and a linear activation function for the output layer. The maximum number of epochs was set to 1000, with early stopping implemented to avoid overfitting: training was automatically interrupted if no further improvement in the validation error was observed for six consecutive iterations. Training typically converged within 600–650 epochs for all networks.

For each SNN, the optimal number of neurons was then determined by iteratively varying the hidden layer size and selecting the configuration that minimized the RMSE on both training and validation sets without signs of overfitting. For example, for U1 volumetric efficiency, 19 neurons were chosen because the RMSE reached its minimum and remained stable below 1 % across further increases (Fig. A2 reports the regression plot of the U1 volumetric efficiency ($\eta_{vol_{u1}}$)). In Fig. A2, the regression plot shows that the predicted values are well distributed around the bisecting line, with a global correlation coefficient $R = 0.99925$ between estimated (η_{est}) and target (η_{real}) values.

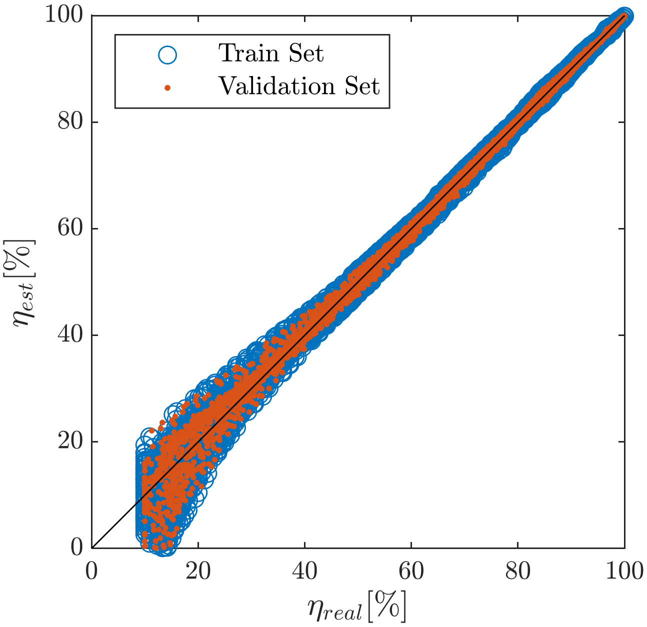


Fig. A2. Performance of U1 volumetric efficiency shallow neural network (SNN): comparison between estimated and real efficiency points for the 19-neuron SNN.

This indicates that the residuals are balanced and unbiased for both training and validation datasets, confirming the suitability of the chosen SNN architecture. All networks achieved correlation coefficients $R > 0.999$, confirming excellent agreement between predicted and target values. Table A1 shows the architecture and performance parameters of the four SNNs.

Table A1
Statistical LS indicators for the most representative experiments.

Efficiency function	Architecture (input-hidden-output)	Neurons [-]	Epochs [-]	RMSE [%]
$\eta_{vol_{u1}}$	3–19–1	19	631	0.31
$\eta_{mec_{u1}}$	3–18–1	18	640	0.33
$\eta_{vol_{u2}}$	3–15–1	15	618	0.25
$\eta_{mec_{u2}}$	3–14–1	14	622	0.24

The efficiency of the hydraulic unit is strongly related to its operative conditions, which means a direct solution of the previously presented implicit IHMCVT model is not possible. Instead, a recursive calculation was structured with the goal of numeric convergence for every input data vector fed to the model (Fig. A3). In particular, this algorithm requires the engine inputs in the form of T_{ICE} and ω_{ICE}. Then, given that the IHMCVT can be piloted by changing the selected displacement for each input pair of engine torque T_{ICE} and speed ω_{ICE}. For each swash plate position λ, the computation is initialized and then ended when a stop condition is satisfied. The main loop is responsible for evaluating the IHMCVT kinematic and mechanical outputs, calculating hydraulic pressure p, and then determining the efficiencies of the hydraulic unit through neural networks to update the model outputs at the step $(i + 1)$. The recursive process includes this main loop and three secondary feedback loops activated when specific conditions are met.

The secondary feedback loops were designed to sequentially act and limit the model outputs in terms of the maximum speed (ω_{u1max} and ω_{u2max}), pressure (p_{max}) of the hydraulic machines, and maximum traction torque to the tractor rear wheels ($T_{wheels_{max}}$). The first two loops were calibrated according to the design limits of the hydraulic units, while the last one was set to reflect the maximum admissible driveline torque, beyond which a safety mechanism disengages the transmission clutches to prevent damage. The numerical thresholds adopted in the model are consistent with those used in previous studies on the same transmission architecture, as documented by Casoli et al. [28] and Martelli et al. [16], and have been applied as fixed parameters throughout the simulation framework. If the first two loops are triggered, they update the variable efficiencies by re-evaluating the neural networks. In contrast, if the loop of maximum traction torque is activated, it recalculates the mechanical outputs and then limits the pressure of the hydrostatic circuit. The algorithm block corresponding to the limiting loop of the tractor torque was designed to serve as a potential point for integrating a realistic wheel–ground interaction model. Finally, the stopping condition is activated whenever the sum of the relative differences between the outputs of an iteration and the subsequent one is lower than a specified tolerance value (*toll*).

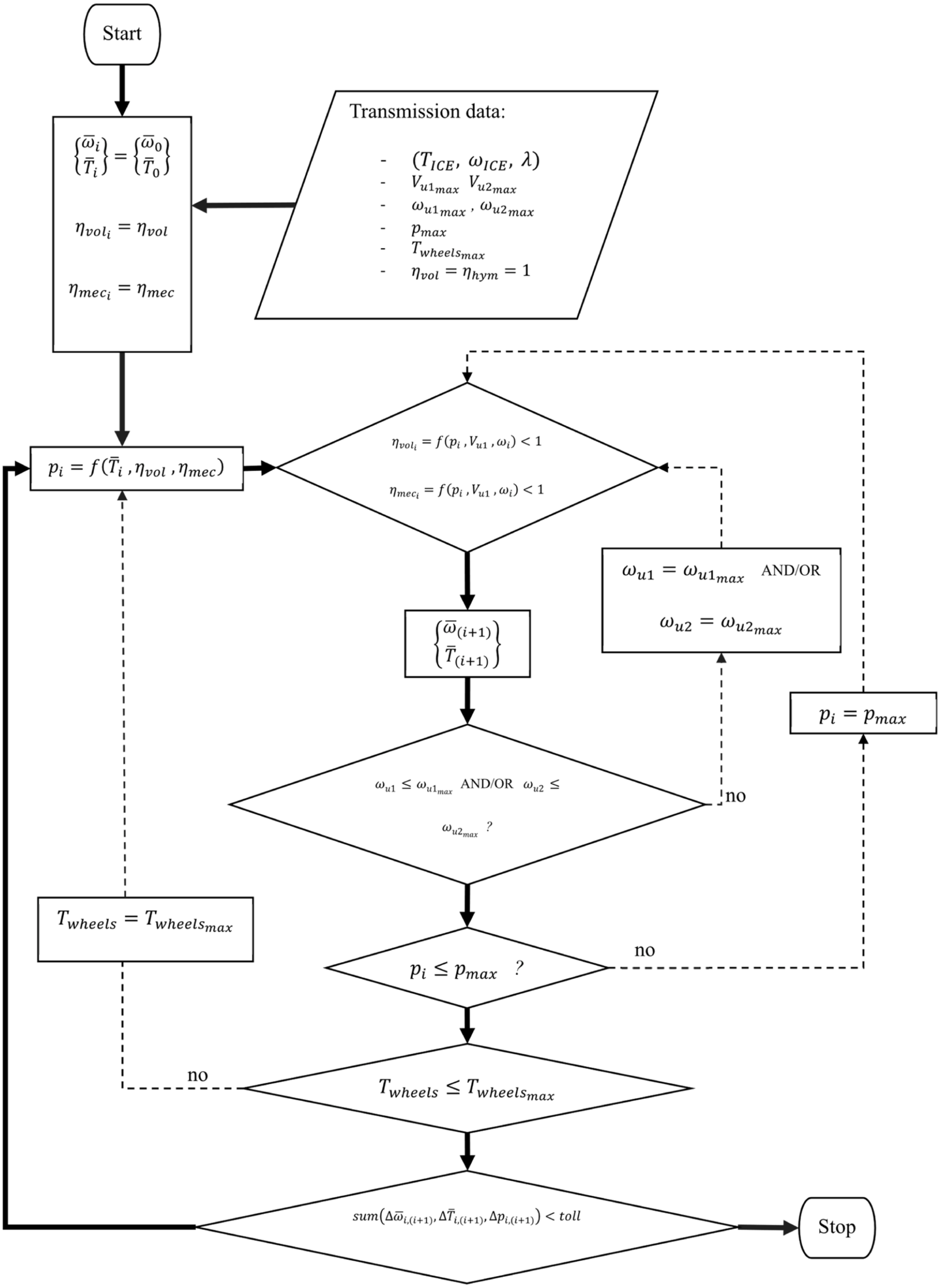


Fig. A3. Iterative algorithm for transmission calculation, according to ISO [34]. The stadium represents a scheme termination, the parallelogram indicates data sent to the initialisation block, the rectangle represents a process function, and the diamond denotes a multiple-exit block. In the diagram, the branches and the main feedback loop are represented by solid arrows, and the secondary feedback branches are indicated with dashed lines.

The procedure was designed to perform iterative calculation loops. Given the wide number of possible transmission operative points, the direct use of the above-displayed algorithm in real-world applications is not practical. For instance, if the output calculation is required from a given real-world time trace input, evaluating the model for each sample point of the trace would be extremely time-consuming. Therefore, instead of directly supplying the algorithm with real-world signals, an initial three-dimensional matrix comprising all possible combinations of T_{ICE}, ω_{ICE}, and λ (representing every possible operating point of the transmission) was generated. This matrix was then fed into the algorithm, and all possible outputs were calculated and subsequently collected into matrices (see Fig. A4). Each matrix is three-dimensionally structured so that, when the transmission works at the bin coordinates (T_{ICE}, ω_{ICE}, λ), the corresponding matrix cell contains the output of interest. Once the matrices are stored, they can be utilised to analyse any possible dataset of the tractor under consideration by rapidly extracting the output at the corresponding coordinates via real-world CANBUS signals.

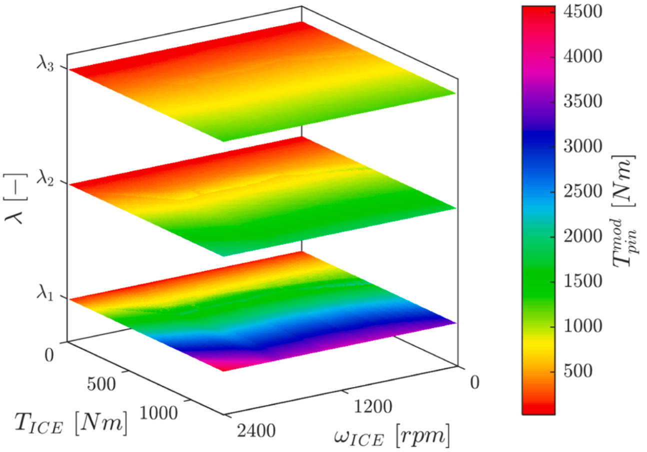
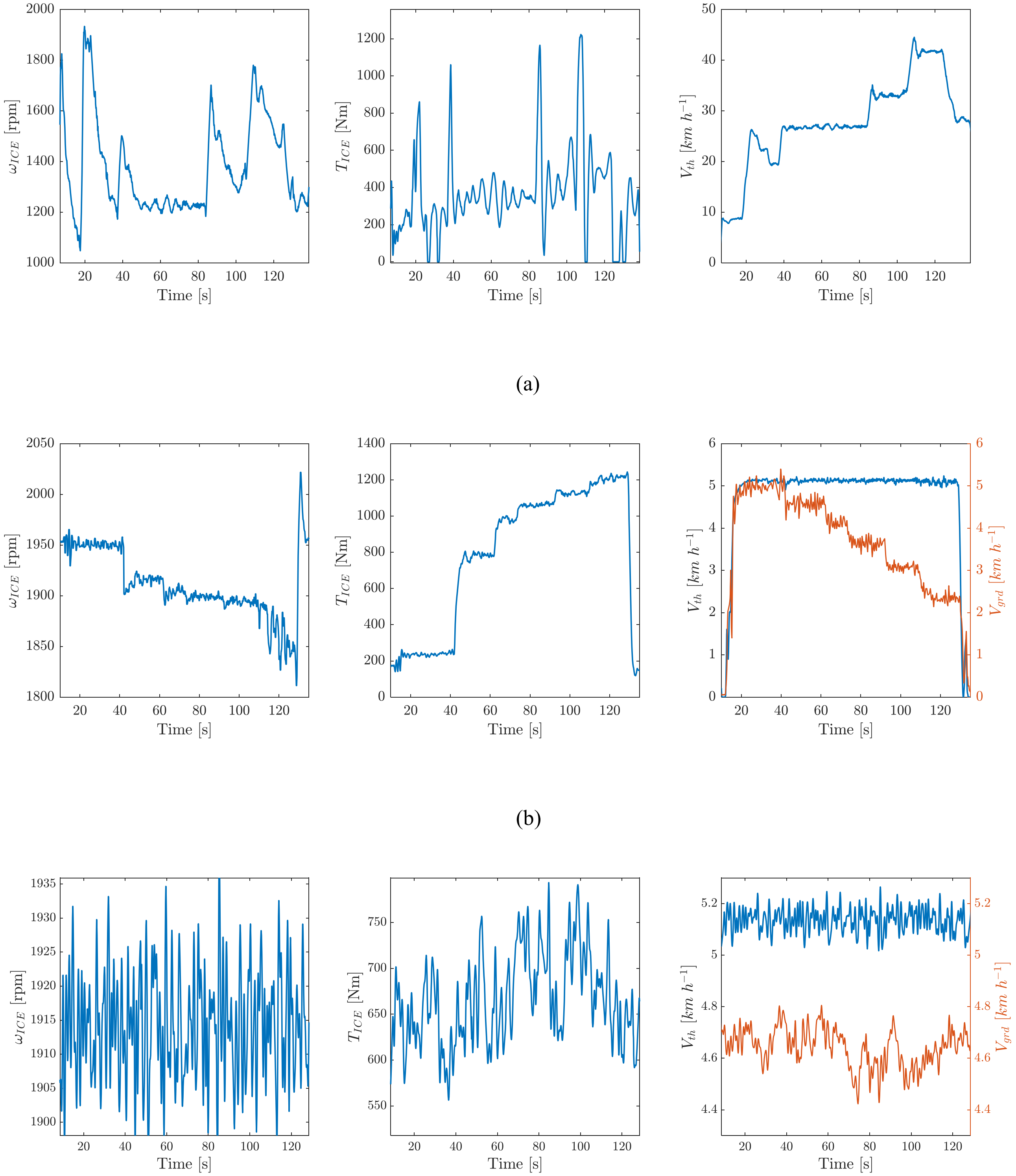


Fig. A4. Example of the pinion torque matrix ($\bar{\bar{T}}_{pin}$), for three generic λ positions as a function of T_{ICE} and ω_{ICE} obtained from the model calculation. The triplet (T_{ICE}, ω_{ICE}, λ) uniquely determines a transmission operating point of the matrix $\bar{\bar{T}}_{pin}$, characterized by a modelled pinion torque value, T_{pin}^{mod}.

For the data extraction process, two matrices were used, $\bar{\bar{\omega}}_{pin}$ and $\bar{\bar{T}}_{pin}$, corresponding to the modelled pinion speed and torque (at pin shaft), respectively.

Appendix B

This section presents partial time-based traces of the transmission variables ω_{ICE}, T_{ICE}, and V_{th} for the three characteristic test types—road transport (RT), drawbar pulling (FD), and field tillage (FT). Fig. B1 displays these temporal traces, while Fig. B1b and Fig. B1c also include the ground speed V_{grd} to highlight the discrepancies between theoretical and actual speeds caused by wheel slippage. Each plot captures a representative segment of data, illustrating the typical dynamic behavior distinctive of each test condition.



(a)

(b)

(c)

Fig. B1. Extracted time-based CAN signals of interest for (a) transport, (b) drawbar, and (c) ploughing tests. Each subfigure, from left to right, represents the transmission working variables engine speed (ω_{ICE}), engine torque (T_{ICE}), and tractor theoretical speed (V_{th}), with the latter two plots also displaying the tractor ground speed (V_{grd}) in red.

In the road transport test (Fig. B1a), ω_{ICE} oscillated between 1000 and 2000 rpm, with a few peaks approaching 1800 rpm, while T_{ICE} generally remained below 500 Nm, reaching up to 1000 Nm during acceleration phases. These acceleration phases are clearly identifiable in the plot by the rapid increases in V_{th}, which correspond to simultaneous spikes in ω_{ICE} and T_{ICE}.

During the drawbar pulling test (Fig. B1b), ω_{ICE} progressively decreased from about 1950 to 1800 rpm, whereas T_{ICE} increased steadily from roughly 200 to 1200 Nm as the load on the tractor intensified under steady-state conditions. Throughout the test, V_{th} remained almost constant at around 5 km h^{-1}, while V_{grd} declined from 5 to approximately 2.2 km h^{-1} due to increasing slippage. At the end of the test repetition, both speeds dropped to zero, T_{ICE} decreased, and ω_{ICE} exhibited a brief spike as a result of the sudden reduction in engine load.

In the ploughing (field tillage) test (Fig. B1c), ω_{ICE} remained nearly constant, between 1900 and 1930 rpm, while T_{ICE} varied moderately between 550 and 750 Nm. V_{th} was maintained at around 5.1 km h^{-1}, whereas V_{grd} fluctuated between 4.5 and 4.8 km h^{-1}, indicating a slip ratio between 5 % and 11 %.

Comparing the three tests, the torque amplitude clearly differentiates the operating conditions. Road transport required the lowest torque, field tillage demanded intermediate-to-high torque levels, and drawbar pulling imposed the highest torque loads. These variations directly influence the magnitude of the drawbar force developed in each test.

Moreover, the rate of change in the measured signals reveals several dynamic phenomena intrinsic to such systems, particularly evident in field conditions. The wheel–soil interaction, the mechanical coupling between the two tractors in the pulling test, and the implement–soil interaction during tillage all contribute to transient effects that affect the measured drawbar force. The distinct signal patterns observed in transport and field tests—especially in terms of T_{ICE} behavior and the corresponding transmission gear ratios shown in Fig. B1—are consequently reflected in the measured drawbar force $F_{d_{jWD}}^{meas}$.

Data availability

Data will be made available on request.

References

[1] M. Mattetti, G. Annesi, F.P. Intrevado, L. Alberti, Investigating the efficiency of hybrid architectures for agricultural tractors using real-world farming data, Appl. Energy 377 (2025) 124499, https://doi.org/10.1016/j.apenergy.2024.124499.

[2] D. Ali, R. De Castro, R. Ehsani, S. Vougioukas, P. Wei, Unlocking the potential of electric and hybrid tractors via sensitivity and techno-economic analysis, Appl. Energy 377 (2025) 124545, https://doi.org/10.1016/j.apenergy.2024.124545.

[3] K. Götz, P. Rosner, I. Froissart, M. Lienkamp, The potential of electric agriculture and mobility for the least-cost rural electrification strategy in Sub-Saharan Africa, Energy Sustainable Dev. 83 (2024) 101581, https://doi.org/10.1016/j.esd.2024.101581.

[4] F. Mocera, A model-based design approach for a parallel hybrid electric tractor energy management strategy using hardware in the loop technique, Vehicles (2021) 1–19, https://doi.org/10.3390/vehicles3010001.

[5] M. Varani, M. Mattetti, G. Molari, Performance evaluation of electrically driven agricultural implements powered by an external generator, Agronomy 11 (2021) 8, https://doi.org/10.3390/agronomy1108144. Article.

[6] E.G. Giakoumis, Driving and Engine Cycles, Springer International Publishing, 2017.

[7] Yimin Gao, M. Ehsani, Design and control methodology of plug-in hybrid electric vehicles, IEEe Trans. Ind. Electron. 57 (2010) 633–640, https://doi.org/10.1109/TIE.2009.2027918.

[8] L. Angelucci, M. Mattetti, The development of reference working cycles for agricultural tractors, Biosyst. Eng. 242 (2024) 29–37, https://doi.org/10.1016/j.biosystemseng.2024.04.004.

[9] e.V. DLG, DLG-PowerMix: tractor output, Efficiency Fuel Consump. (2025).

[10] M. Varani, M. Mattetti, G. Molari, A. Biglia, L. Comba, Correlation between power harrow energy demand and tilled soil aggregate dimensions, Biosyst. Eng. 225 (2023) 54–68, https://doi.org/10.1016/j.biosystemseng.2022.11.008.

[11] J. Wieckhorst, T. Fedde, L. Frerichs, G. Fiedler, A tractive sensor-integrated measurement of tire soil parameters for tractors, VDI-Berichte 2251 (2015) 219–226.

[12] L. Angelucci, F. Pinet, A. Vertua, M. Mattetti, Towards more efficient tractors: assessing and refining traction test procedures for agricultural tractors, J. Terramech. 117 (2025) 101018, https://doi.org/10.1016/j.jterra.2024.101018.

[13] M. Mattetti, M. Maraldi, N. Lenzini, S. Fiorati, E. Sereni, G. Molari, Outlining the mission profile of agricultural tractors through CAN-BUS data analytics, Comput. Electron. Agric. 184 (2021) 106078, https://doi.org/10.1016/j.compag.2021.106078.

[14] K. Götz, A. Kusuma, A. Dörfler, M. Lienkamp, Agricultural load cycles: tractor mission profiles from recorded GNSS and CAN bus data, Data Brief. 60 (2025) 111494, https://doi.org/10.1016/j.dib.2025.111494.

[15] K. Pettersson, Design Automation of Complex Hydromechanical Transmissions, Ph. D. Linköping University - Department of Management and Engineering, 2013.

[16] M. Martelli, D. Chiarabelli, S. Gessi, P. Marani, E. Mucchi, M. Polastri, Comprehensive lumped parameter and multibody approach for the dynamic simulation of agricultural tractors with tyre–soil interaction, IET Cyber-Syst. Robot. 5 (2023), https://doi.org/10.1049/csy2.12092 e12092.

[17] M. Treiber, A. Kesting, Traffic Flow Dynamics: Data, Models and Simulation, Springer, 2013.

[18] USDA, Soil Survey Manual. vol. Handbook 18, Chapter 3, U.S. Department of Agriculture, 2017.

[19] Catalog of Soils - Emilia-Romagna Region. Catalog of soils—Emilia-Romagna region. 2015.

[20] ASTM. Test methods for laboratory determination of density (Unit Weight) of soil specimens 2021. https://doi.org/10.1520/D7263-21.

[21] ASAE. Procedures for using and reporting data obtained with the soil cone penetrometer 2019.

[22] W.W. Brixius, Traction prediction equations for bias ply tires. Transactions of the ASAE, ASAE - American Society of Agricultural Engineers, St. Joseph, MI, 1987.

[23] M. Saetti, M. Mattetti, M. Varani, N. Lenzini, G. Molari, On the power demands of accessories on an agricultural tractor, Biosyst. Eng. 206 (2021) 109–122, https://doi.org/10.1016/j.biosystemseng.2021.03.015.

[24] F.G. Migliaccio, Sviluppo di una trasmissione a variazione continua per l'azionamento delle pompe idrauliche di una trattrice agricola. https://morethesis.unimore.it/theses/available/etd-03182023-195216/, 2023 (accessed February 1, 2025).

[25] A. Rossetti, A. Macor, Multi-objective optimization of hydro-mechanical power split transmissions, Mech. Mach. Theory. 62 (2013) 112–128, https://doi.org/10.1016/j.mechmachtheory.2012.11.009.

[26] Macor A., Rossetti A. Optimization of hydro-mechanical power split transmissions | Elsevier Enhanced Reader 2011. https://doi.org/10.1016/j.mechmachtheory.2011.07.007.

[27] M. Scamperle, Advanced Design Methods For Hydromechanical Transmissions in Road and Agricultural Vehicles. PhD Thesis, University of Padua - Deparment of Engineering and Management, 2017.

[28] Casoli P., Vacca A., Berta G.L., Meleti S., Vescovini M. A numerical model for the simulation of Diesel/CVT power split transmission, 2007, p. 2007-24–0137. https://doi.org/10.4271/2007-24-0137.

[29] ASTM. Practices for cycle counting in fatigue analysis 2017. https://doi.org/10.1520/E1049-85R17.

[30] J. Safarov, U. Ashurova, B. Ahmadov, E. Abdullayev, A. Shahverdiyev, E. Hassel, Thermophysical properties of Diesel fuel over a wide range of temperatures and pressures, Fuel 216 (2018) 870–889, https://doi.org/10.1016/j.fuel.2017.11.125.

[31] European Committee for Standardization (CEN). EN 590: automotive fuels — diesel — requirements and test methods 2022.

[32] International Organization for Standardization (ISO). ISO 3046-1: reciprocating internal combustion engines — Performance — Part 1: standard reference conditions and declarations of power, fuel and lubricating oil consumption 2020.

[33] T.H. Chen, H.C. Tang, Meshing efficiency analysis of Straight spur gear pair, AMM 764–765 (2015) 314–318, https://doi.org/10.4028/www.scientific.net/AMM.764-765.314.

[34] ISO, Information processing—Documentation symbols and conventions for data, program and system flowcharts, Program Netw. Charts Syst. Resour. Charts (1985).