European Journal of Forest Research
https://doi.org/10.1007/s10342-024-01717-6

RESEARCH

High-resolution harvester data for estimating rolling resistance and forest trafficability

Aura Salmivaara[1] · Eero Holmström[1] · Sampo Kulju[1] · Jari Ala-Ilomäki[1] · Petra Virjonen[2] · Paavo Nevalainen[2] · Jukka Heikkonen[2] · Samuli Launiainen[1]

Received: 21 March 2024 / Revised: 27 May 2024 / Accepted: 30 June 2024

Abstract
Information on terrain conditions is a prerequisite for planning environmentally and economically sustainable forest harvesting operations that avoid negative impact on soils. Current soil data are coarse, and collecting such data with traditional methods is expensive. Forest harvesters can be harnessed to estimate the rolling resistance coefficient (μ_{RR}), which is a proxy for forest trafficability. Using spatio-temporal data on engine power used, speed travelled, and machine inclination, μ_{RR} can be computed for harvest areas. This study describes an extensive, high-resolution data on μ_{RR} collected in a boreal forest landscape in Southern Finland during the non-frost period of 2021, covering roughly 50 km of harvester routes. We report improvements in removing some of the previous restrictions on calculating μ_{RR} on steeper slopes, enabling the calculation within a $-10°$ to $+10°$ slope range with a speed range of 0.6–1.2 ms^{-1}. We characterise the variation in μ_{RR} both between and within 11 test sites harvested during the April-August period. The site mean μ_{RR} varies from ~ 0.14 to 0.19 and shows significant differences between the sites. Using simulations of the hydrological state of the soil and open spatial data on forest and topography, we identify features that best explain the extremes of μ_{RR} within the sites. Several wetness-related indices, such as the depth-to-water index with varying thresholds, explain the μ_{RR} extremes, while biomass-related stand attributes indirectly explain these through their linkage to site and soil characteristics. Obtaining μ_{RR} from actual operational data extends the capabilities of large-scale harvester-based data collection and paves the way for building data-driven models for trafficability prediction.

Keywords Trafficability · Rolling resistance · Soil properties · ARCNET data · Spatial analysis · Forest operations

Introduction

Mechanised forest operations exert pressure on forest soils, which compromises the acceptability of forest management. Forest vehicles can cause major soil deformations and/or

Communicated by Eric R. Labelle.

✉ Aura Salmivaara
aura.salmivaara@luke.fi

Eero Holmström
eero.holmstrom@luke.fi

Sampo Kulju
sampo.kulju@luke.fi

Jari Ala-Ilomäki
jari.ala-ilomaki@luke.fi

Petra Virjonen
pekavir@utu.fi

Paavo Nevalainen
paavo.nevalainen@utu.fi

Jukka Heikkonen
jukhei@utu.fi

Samuli Launiainen
samuli.launiainen@luke.fi

[1] Natural Resources Institute Finland, Latokartanonkaari 9, 00790 Helsinki, Finland

[2] Department of Computing, University of Turku, Vesilinnantie 3, 20500 Turku, Finland

Published online: 09 July 2024

soil compaction, which have cascading impacts on soil hydraulic properties, water flow, soil biogeochemistry and plant growth (Hoffmann et al. 2022; Arvidsson et al. 2003; Hansson et al. 2019). The degree of damage is related to the machinery applied and to the local soil strength, which is influenced by soil texture, stoniness, organic matter content, root systems, soil moisture and frost/thaw dynamics (IPCC 2021; Akumu et al. 2019; Schönauer et al. 2022). To minimise negative impacts, operations need to be planned so that vehicle routing is optimised considering site conditions, and past and current weather conditions are accounted for. Accurate data on forest soil bearing capacity is a prerequisite for planning efficient and environmentally acceptable forest operations (Hoffmann et al. 2022; Schönauer et al. 2021).

Despite the availability of several mapping approaches serving forest trafficability assessment (Latterini et al. 2022a; Kankare et al. 2019; Shabani et al. 2019; Jones and Arp 2019; Mohtashami et al. 2012; Reeves et al. 2012; Murphy et al. 2011; Suvinen 2006), current trafficability maps do not adequately reflect the impact of terrain properties and varying weather conditions (Mohtashami et al. 2022; Schönauer et al. 2021; Mohtashami et al. 2017, 2023; Salmivaara et al. 2020b). This is a major knowledge gap as unfavourable conditions, including shorter periods of frozen soils and heavy rainfall events in spring and autumn, are becoming more common with climate change (Venäläinen et al. 2020; IPCC 2021), making it more difficult to conduct low-impact operations. Thus, trafficability assessments must be improved both in their spatial and temporal resolution (Labelle et al. 2022; Hoffmann et al. 2022; Salmivaara et al. 2020b).

Spatial patterns of forest soil moisture can, to some extent, be depicted using topography-based depth-to-water (DTW) maps (Lidberg et al. 2020; Kemppinen et al. 2023; Oltean et al. 2016; Ågren et al. 2014; Mohtashami et al. 2022). Some studies have also reported success in predicting the most severe rut damages using DTW (e.g. Campbell et al. 2013), but commonly the performance of DTW in predicting rutting (Mohtashami et al. 2017; Ågren et al. 2015) or soil compaction (Latterini et al. 2022b) has been very modest. This is likely because only a minor fraction of soil strength variations are related to topography (Oltean et al. 2016) and associated moisture variability (Mohtashami et al. 2023; Schönauer et al. 2022), and the large between-site variability in, e.g., soil texture compromises the usefulness of topographic wetness indices (Heppelmann et al. 2022; Vega-Nieva et al. 2009; Niemi et al. 2017). This is in line with the findings of Kemppinen et al. (2023), who found that the relationship between volumetric soil moisture and topographic indices is very site-specific.

The use of semi-empirical soil strength and rut depth models (e.g. Saarilahti 2002; Vega-Nieva et al. 2009) requires both knowing the soil moisture status and having accurate information on soil type at spatial scales relevant for forest operations (Salmivaara et al. 2020b). Currently, the resolution of such soil data is too coarse (e.g., 1:20 000 to 1:200 000 in Finland; GSF (2015)), and high uncertainties prohibit predicting rut depths using these traditional models. Collecting soil data manually is expensive and slow (Lagacherie and McBratney 2006), and therefore, machine learning methods have been used to predict soil attributes relevant for trafficability, such as organic layer depth (Ågren et al. 2022) and soil moisture (Lidberg et al. 2020) using open spatial data. However, these approaches cannot directly predict temporal variability of soil strength.

Recently, the harvester rolling resistance coefficient (μ_{RR}) was shown to correlate well with the depth of ruts formed in harvesting (Ala-Ilomäki et al. 2020; Salmivaara et al. 2020b). The more the machine movement is resisted by the wheel-terrain interaction, the higher the μ_{RR}, and higher μ_{RR} is often associated with deeper ruts (Ala-Ilomäki et al. 2020; Salmivaara et al. 2020b; Kurjenluoma et al. 2009; Bygdén et al. 2003). Consequently, a higher μ_{RR} represents less trafficable soils with higher risks of soil damage (Kurjenluoma et al. 2009; Bygdén et al. 2003).

The Controller Area Network (CAN-bus) data of a harvester offers high-resolution data on, e.g., fuel consumption, engine power usage and speed of the machine (Ala-Ilomäki et al. 2020; Suvinen and Saarilahti 2006; Melander et al. 2020). Supplementing such routinely collected machine data with inclination, mass and accurate position of the machine, enables calculating μ_{RR} for the areas where the harvester or forwarder has operated (Ala-Ilomäki et al. 2020). The use of CAN-bus data for estimating μ_{RR} and forest soil trafficability has so far been used only at specific experimental setups (Ala-Ilomäki et al. 2020). Lately, Salmivaara et al. (2020b) envisaged how operational collection of rolling resistance data, when merged with open spatial data and hydrological forecasts, could enable nation-wide trafficability prediction.

In this study, we present the first steps in that direction. We collect spatially and temporally extensive harvester machine data, covering a range of actual operational harvest sites and diverse hydrometeorological conditions during the spring-autumn period of 2021. We provide a refined model to compute transmission power from engine power to estimate μ_{RR} for a cut-to-length harvester (Ala-Ilomäki et al. 2020). With the extensive data representing roughly 50 km of harvester routes, we characterise and interpret the within and between-site variability of μ_{RR}. We identify the terrain attributes that are the most influential for μ_{RR} extremes. The collected database, including the relevant spatial data on terrain and stand characteristics, μ_{RR}, and the harvester data used for computing μ_{RR} is published as open access to catalyse future research.

Methods

Study area, harvester data and available spatial data

The harvester machine data and Global Navigation Satellite System (GNSS) data were collected from a Ponsse Ergo harvester from January 2021 to March 2022 in the Länsi-Uusimaa region in Southern Finland. In this harvester, data was obtained through the ARCNET communications protocol, which is basically similar to CAN but can transmit data at a higher rate. The harvester, instrumented with an inclinometer to measure vehicle longitudinal inclination, i.e., pitch, was operating in normal working conditions, where harvesting routes are usually covered with a brush mat. The data from the non-frost period (mean monthly temperature above 0° C, i.e., April 2021 to November 2021) covers roughly an area of 40 ha and 50 km of harvesting routes. A total of 11 test sites, harvested during April to August 2021 (Fig. 1), were selected for more detailed studies. The site attributes (mean/majority terrain feature values) are shown in Supplementary Table S2.

For computing μ_{RR}, it is necessary to know the transmission hydraulic motor power. Instead of equipping the harvester with a special set of sensors for measuring the hydraulic transmission power throughout the data collection campaign, we created a regression model for predicting hydraulic power from diesel engine power, the latter which can be directly computed from diesel engine torque and rotational speed in the ARCNET data. This approach was similar to that in Ala-Ilomäki et al. (2020) based on a Ponsse Scorpion harvester but covered a wider range of machine pitch and power. The data needed for creating the regression model was collected in a separate test drive in December 2022.

The spatial forest hydrology model SpaFHy (Launiainen et al. 2019) was used with daily gridded 10 km × 10 km weather data, covering the dates of operation, to estimate spatial variation in saturation deficit (`sat_deficit_doy`) and volumetric soil moisture content at a 16 m × 16 m resolution. The spatial data (Table 1 in Launiainen et al. 2019) on forest characteristics, soil and topography was used to set up the model for the test sites. The `sat_deficit_doy` can be interpreted as a weather-dependent topographic wetness index (TWI) (Beven and Kirkby 1979). It is a robust metric of local soil moisture regime at a given time, with the advantage of being easy to compute but still sensitive to weather variability (Beven and Kirkby 1979; Launiainen et al. 2019). It depends on local slope, the area contributing on water flow to that particular grid-cell, and the mean ground water storage of the landscape. For instance, after snowmelt and heavy rainfalls, landscape-average `sat_deficit` is small, and there are more fully saturated grid-cells (`sat_deficit` = 0) than in dry summer conditions. Volumetric soil moisture gives the percentage of water from unit volume of soil.

In addition, road data (National Land Survey of Finland 2023) were extracted for the non-frost period dataset, and other available open spatial data were compiled for the 11 test sites (See Supplementary material Tables S1 and S2).

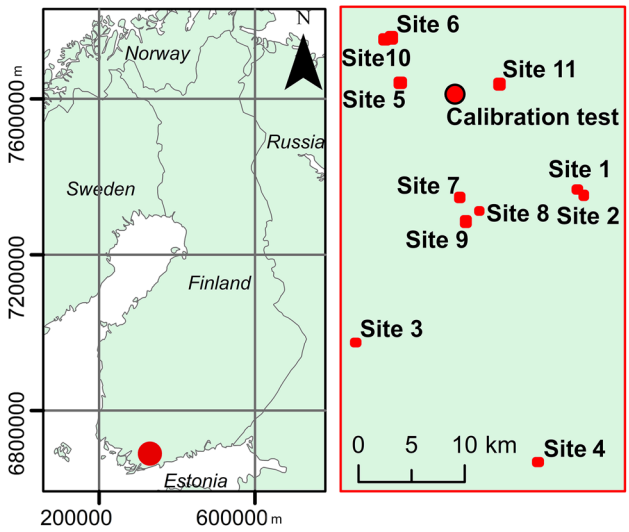

Fig. 1 Location of the test sites in Länsi-Uusimaa, Southern Finland

Transmission power model

To calculate μ_{RR}, we need the hydraulic transmission output power (P_t, kW) (Ala-Ilomäki et al. 2020). P_t can be determined directly using special hydraulic pressure sensors, or estimated from diesel engine power (P_e, kW) with a machine-specific semi-empirical model (Ala-Ilomäki et al. 2020). The latter approach has the advantage that μ_{RR} can be estimated from ARCNET data without additional instrumentation of the vehicle.

To build the model to predict P_t from diesel engine power, we performed a calibration test run on 9th December 2022. In the test, the harvester was driven uphill and downhill both on road and on forest terrain with the harvester boom facing forward. The bearing conditions were good with air temperature −5° C and forest terrain consisting mainly of dry mineral soil. This enabled focusing on the total engine power and transmission power ratio and made it possible to cover a wide range of power outputs in varying slopes, including engine braking in steep downhill and power cut-off in steep uphill. In a steep downhill, the machine commonly uses the transmission for braking, and the engine is consequently turned over by transmission return flow power. In a steep uphill, transmission pressure can reach a power

cut-off threshold which protects the transmission system from overloading.

P_t was determined based on the differential pressure over the hydraulic pump, motor volume per one rotation set by the transmission control unit, the hydraulic unit rotational speed, and the hydraulic efficiency using Eq. 1, supplied by Ponsse Plc.:

$$P_t = (P_{hdiff} \cdot (n_m \cdot 60 \cdot (I_m \cdot a_1 + b_1)/\eta_{mvol} \cdot 1000))/1000. \quad (1)$$

Here P_{hdiff} is the differential hydraulic pressure (kPa) measured with the pressure sensors in the test run, n_m is hydraulic motor rotational speed (s^{-1}), I_m is hydraulic motor control current (mA), $a_1 = (V_{mmax} - V_{mmin})/(I_{mmin} - I_{mmax})$, V_{mmax} is the hydraulic motor maximum displacement per revolution (cm^3), V_{mmin} is hydraulic motor minimum displacement per revolution (cm^3), I_{mmin} is the hydraulic motor minimum control current (mA), I_{mmax} is hydraulic motor maximum control current (mA), $b_1 = -a_1 \cdot I_{mmin} + V_{mmax}$, and η_{mvol} is the coefficient of motor volumetric efficiency. The calculated P_t was multiplied by −1, which gives the power a positive value when the engine is transmitting power to the wheels and the machine is moving with the harvester boom up front. In cases of steep downhill, the transmission power can get negative values, which indicates engine braking.

The effects of acceleration from standstill to the approximately constant driving speed were excluded from the transmission power model by removing data points where the speed of the harvester, which was taken from the ARCNET wheel rotational speed, was smaller than 0.5 ms^{-1}.

The resulting data was used to fit a piecewise function to model the transmission hydraulic power P_t as a function of the diesel engine power P_e (Eq. 2):

$$P_t(P_e) = \begin{cases} A \cdot P_e + B \cdot (P_e - P_{brake})^3 + D\,, & P_e \le P_{brake} \\ A \cdot P_e + C \cdot (P_e - P_{brake})^2 + D\,, & P_{brake} < P_e \le P_{cut} \\ P_C + (E - P_C) \cdot (1 - exp((P_{cut} - P_e)/F))\,, & P_e > P_{cut} \end{cases}, \quad (2)$$

where A, B, C, D, E, F, P_{brake} and P_{cut} are fitting parameters. P_{brake} and P_{cut} represent engine power values where engine braking and pressure cut-off begin to affect the transmission hydraulic power respectively. P_C, the hydraulic power at the point of pressure cut-off, is calculated as $P_C = A \cdot P_{cut} + C \cdot (P_{cut} - P_{brake})^2 + D$. The initial parameter values of the model were searched for using a differential evolution method (Storn and Price 1997). The final values were the result of a curve fit based on a non-linear least-squares method, which yielded the following parameter values: $A = 0.327$, $B = 0.005\ kW^{-2}$, $C = 0.004\ kW^{-1}$, $D = -12.662$ kW, $E = 116.949$ kW, $F = 138.331$ kW, $P_{brake} = 21.921$ kW, and $P_{cut} = 105.847$ kW. A heatmap of the calibration data consisting of the measured hydraulic transmission power P_t versus the diesel engine power P_e and the fitted model function are presented in Fig. 2. The root-mean-square-error (RMSE) of the fitted function was 7.37 kW and R^2 was 0.94.

The time series of the modelled P_t and the measured $P_{t,meas}$ from the calibration test drive are shown in Fig. 3. The modelled P_t significantly deviated from the measured one only when the harvester was driving steeply uphill and the power cut-off was activated and wheel slip occurred (Fig. 4). Wheel slip was indicated by differences between wheel speed and GNSS-measured vehicle speed logged along the ARCNET data and these were observed at slopes greater than 10°. When driving downhill with pitch < −10°, the wheel speed was lower than vehicle GNSS speed due to

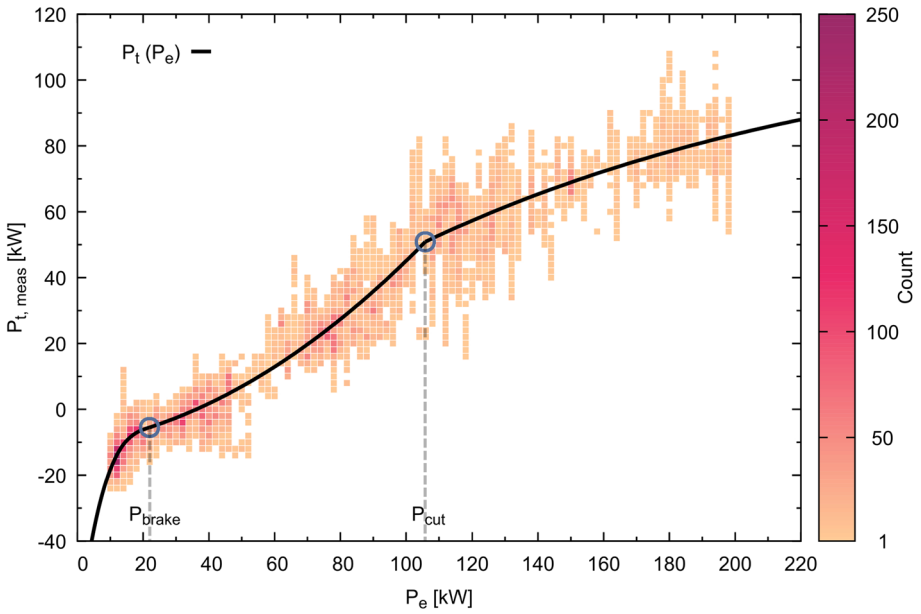

Fig. 2 Heatmap of the test calibration data with a 2 kW × 2 kW bin size, consisting of the transmission power P_t versus the diesel engine power P_e. The fitted Eq. 2 is plotted as the black line. P_{brake} and P_{cut} represent engine power values where engine braking and a pressure cut-off begin to affect the transmission power

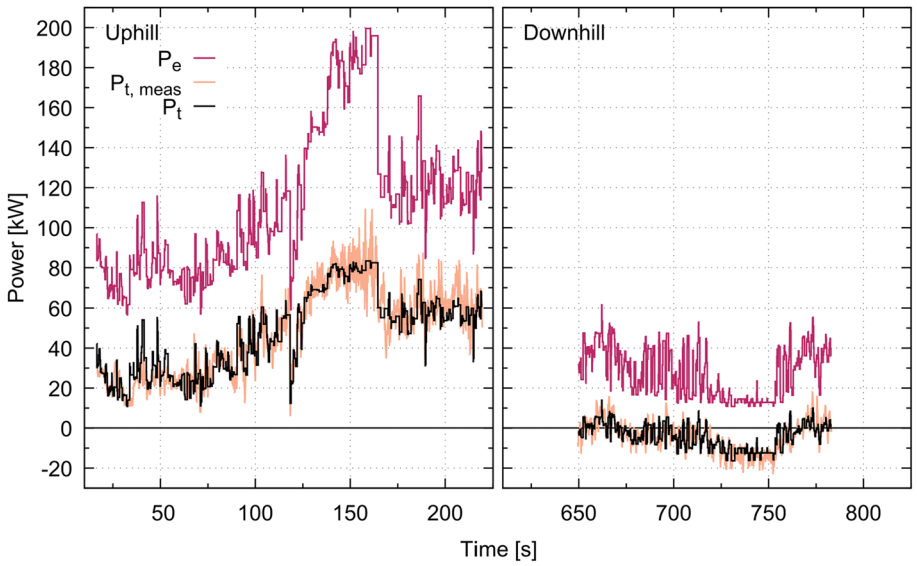

Fig. 3 Measured ($P_{t,meas}$) and modelled (P_t) transmission power and diesel engine power (P_e) when driving uphill (left) and when driving downhill (right) in the calibration test drive

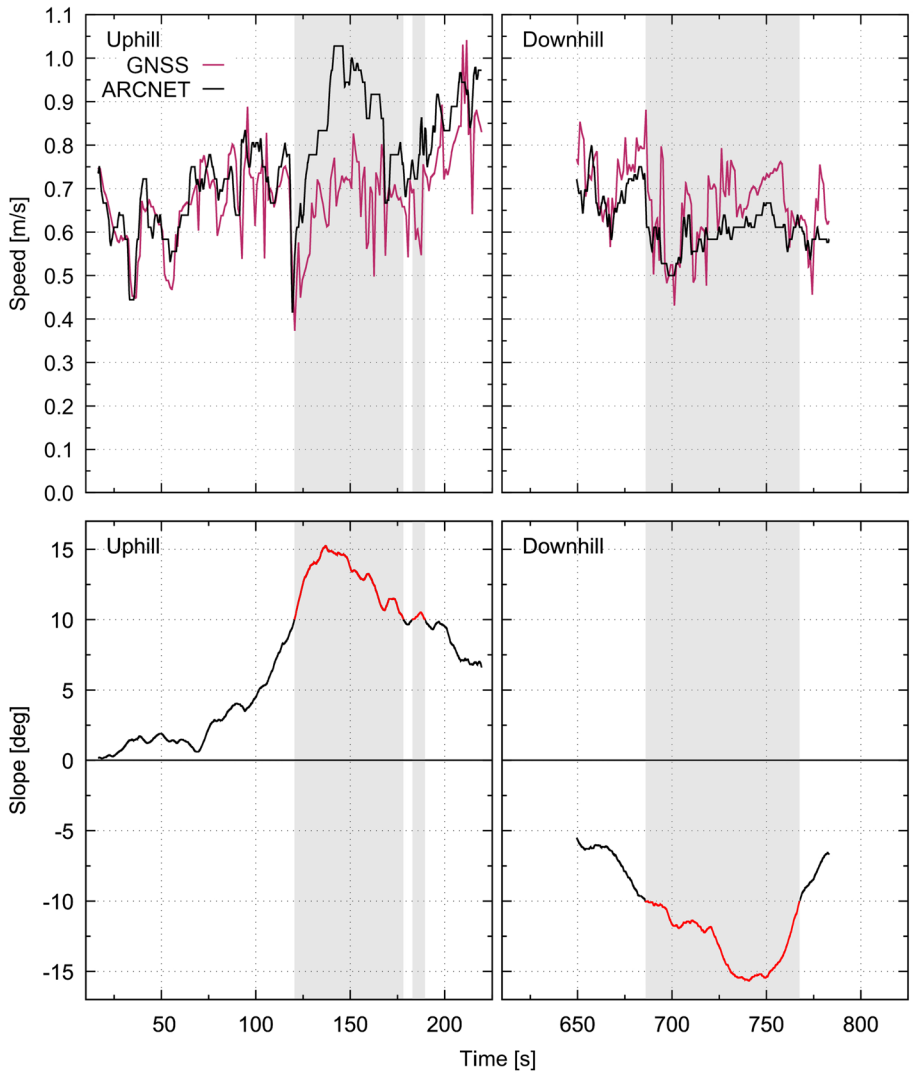

Fig. 4 Example of speed from the harvester GNSS device and the ARCNET wheel speed (top) when driving up and down the calibration test hill. Pitch, i.e., machine longitudinal inclination relative to flat ground (bottom). The grey range indicates times when pitch was outside of the [−10°, 10°] range

skidding of wheels. To avoid the conditions of wheel slip, skid and power cut-off, we limited all further analysis to the slope range [−10°, 10°].

Rolling resistance coefficient

The harvester rolling resistance coefficient (μ_{RR}) was the target variable of interest in this work. The fundamental basis for determining μ_{RR} is the equation of motion of the harvester,

$$\overline{F}_m + \overline{F}_s + \overline{F}_r = m\overline{a}, \tag{3}$$

where $\overline{F}_m$ is the motive force, $\overline{F}_s$ is the slope force, m is the total mass of the machine, and $\overline{a}$ is the acceleration of the machine. $\overline{F}_r$ is the total force acting on the machine after the effect of slope and motive force have been accounted for. By estimating $\overline{F}_m$ and $\overline{F}_s$ from the ARCNET data, inclinometer data and m, and assuming constant speed (i.e., $a = 0$), one can solve Eq. 3 for $\overline{F}_r$. We assume here that $\overline{F}_r$ is solely due to rolling resistance. Then, considering a normal force $\overline{F}_n$ acting on the machine as estimated from the inclinometer data and m, μ_{RR} can be calculated via

$$\mu_{RR} = F_r/F_n \tag{4}$$

The data collection from the harvester was automatised so that a new data file was created each time the harvester was started up. This led to hundreds of separate data files being created over the data collection period. The ARCNET data was logged at a frequency of 50 Hz. In practice, the collected variable values were mostly updated at a lower frequency into the data. A computation pipeline was established to process the raw data files into μ_{RR} and GNSS path data as follows.

First, each data file was cleaned by removing a few incomplete rows. Next, the GNSS path data, obtained along the ARCNET data file, was smoothed and output using a sliding average with a window size of 1.02 s.

Then, the machine data variables of diesel engine torque (τ, Nm), diesel engine rotational speed (n_{RPS}, s^{-1}), speed (v, ms^{-1}) as determined directly from the rotation of the wheels, as well as the pitch of the machine (α, °), were each smoothed using a custom-sized sliding average window. The window size (3.02 s for τ and n_{RPS}, 1.02 s for v, and 10.02 s for α) was determined by visually examining the time series of each variable. In the calibration test drive described earlier, the pitch sensor was found to have a systematic bias of −1.87°, and α was corrected for this bias.

After this, μ_{RR} was computed using the smoothed machine data variables. For each row of data, the power output by the diesel engine was first computed as $P_e = \tau \times 2\pi n_{RPS}$, after which power output by the hydraulic motor (P_t) was obtained using Eq. 2. The computation was restricted to the case of the harvester moving forward with the harvester boom up front, as this is the typical configuration for a harvester that is moving between logging positions. Four cases can be distinguished depending on the sign of the pitch and the power output by the hydraulic transmission:

(1) $\alpha \geq 0, P_t \geq 0$
(2) $\alpha < 0, P_t \geq 0$
(3) $\alpha \geq 0, P_t < 0$
(4) $\alpha < 0, P_t < 0.$

Here $\alpha \geq 0$ describes the situation where the machine is moving uphill or on level ground, and $\alpha < 0$ indicates that the machine is moving downhill. $P_t > 0$ means that the hydraulic transmission is feeding power to the machine wheels, whereas $P_t < 0$ indicates that the wheels are feeding power to the hydraulic transmission (i.e., the engine is acting as a brake).

The equation of motion for the harvester (Eq. 3) was formulated separately for each of the four cases presented above. To simplify the analysis, we assumed constant speed for the harvester, i.e., we set the right-hand side of Eq. 3 to zero. Then, we adopted the following definitions: $F_n = ||mg\cos(\alpha)||$ and $F_s = ||mg\sin(\alpha)||$, where g is the gravitational constant. The motive force F_m was computed as $||\varepsilon P_t/v||$ when $P_t \geq 0$, here ε being the efficiency of the mechanical transmission between the hydraulic motor and the wheels, and assumed here to be 0.74, similarly as in Ala-Ilomäki et al. (2020). In this case, the motive force was a force on the machine in the direction of forward motion, with energy flowing from the hydraulic motor to the wheels. In contrast, when $P_t < 0$, the motive force became a decelerating force due to engine braking and was computed as $||P_t/(v\varepsilon)||$ to describe the flow of energy from the wheels to the hydraulic motor. From each equation of motion, the corresponding expression for $\mu_{RR} = F_r/F_n$ was obtained:

(1) $\mu_{RR} = (F_m - F_s)/F_n$
(2) $\mu_{RR} = (F_m + F_s)/F_n$
(3) (Constant speed is not possible in this case)
(4) $\mu_{RR} = (F_s - F_m)/F_n$

For case (3), constant speed is not possible, as all force components on the machine act against the forward motion, and a balance of net zero force is unattainable. The value of μ_{RR} was computed for each row of the machine data, using the smoothed values of the variables. The result, including the location of the machine, was output for each row.

The series of smoothed GNSS path data was then concatenated chronologically to obtain a single GNSS path for the entire time interval of data collection. Similarly, the series of smoothed machine data along with the result for μ_{RR} for each

row was concatenated as ordered in time. The time series of μ_{RR} with the smoothed machine data variables was then filtered to include only rows with machine speed and pitch of $0.6\ \mathrm{ms}^{-1} \leq v \leq 1.2\ \mathrm{ms}^{-1}$ and $-10° \leq \alpha \leq 10°$, respectively. By visually inspecting the data, the v range from 0.6 to 1.2 ms^{-1} was deemed an interval in which the machine moved at approximately constant speed. We also assumed that in this speed range, the power consumption of other functions besides machine movement, e.g., boom usage, was negligible, and practically all engine power was used for machine movement. Impacts of this practical simplification are discussed later. By filtering the data by pitch value, we excluded cases of excessive slip when computing μ_{RR}. This was done because we did not collect the necessary data for taking into account slip resistance (Suvinen and Saarilahti 2006; Yong et al. 1984) in the total resistive force F_r (Eq. 3) which we used to compute μ_{RR} (Eq. 4). After this, rows corresponding to case (3) of the equation of motion were filtered out (covering only 0.001% of the remaining data). Finally, the data for each of the 11 test sites was extracted for detailed analysis by cutting both the GNSS and μ_{RR} data by a rectangular bounding box for each site.

Analysis of the rolling resistance data

The rolling resistance of a wheel is mainly due to two factors: 1) the deformation of the wheel 2) the deformation of the surface on which the wheel rolls (Suvinen and Saarilahti 2006; Wong 2001). In the case of a forest machine the latter component dominates (Kurjenluoma et al. 2009), and μ_{RR} therefore primarily describes machine rolling resistance due to the deformation of the terrain. Throughout this paper, we assume that the higher μ_{RR} is, the greater the deformation of the ground and the worse the terrain trafficability.

To ensure that μ_{RR} is representative of intact forest terrain, only data from off-road forest and from the first harvester pass was used. This was achieved by ordering data by the timestamp, and for data points in any given 8 m × 8 m spatial window of a regular grid, all subsequent data points occurring within a proximity of 4 m and at least 5 min later than the earliest timestamp were discarded from further analysis.

For deciphering the capability of μ_{RR} to provide information on trafficability conditions, we first made statistical tests for the entire data collected during the non-frost period in 2021 to ensure that μ_{RR} is lower on ground where bearing capacity is higher, as should be when driving on forest road versus driving off-road (Bygdén et al. 2003; Ala-Ilomäki et al. 2020). For this we statistically tested the homogeneity of variances (Levene's test, p value < 0.05 indicating confidence for rejecting the null hypothesis of equal variances) and further the differences in the means (t-test when variances are equal and Welch's t-test when variances are unequal, in both tests p value < 0.05 indicating confidence for declining the null hypothesis of means being equal).

The differences in μ_{RR} means across the 11 test sites were tested using the non-parametric Kruskal–Wallis test suitable for comparing multiple groups at once (p value < 0.05 suggesting confidence in rejecting the null hypothesis of all sites having equal means). To find out which sites differ from each other, Dunn's test was used to study each site's rate against the others through a matrix of p values (again value < 0.05 indicating confidence for rejecting the null hypothesis of equal means). The Bonferroni correction was applied to these p values to reduce the risk of false positives due to multiple comparisons. Exploring this matrix enables detecting groups of sites or individual sites differing significantly from other sites in their μ_{RR} means.

The relationship between μ_{RR} variability and local terrain characteristics was further studied within each site. First, for each 8 m × 8 m square, the mean or mode of values was retrieved for each feature in the original open spatial data (Supplementary Table S1) and for the SpaFHy-model outputs. We sought features that best explained the within-site variability of μ_{RR}. For this, we sub-sampled the μ_{RR} data and examined those grid cells where the mean of μ_{RR} was below the 10% quantile or above the 90% quantile of the site's μ_{RR} distribution. The median (of the mean or mode) of each spatial feature was then determined for these two subsets of data, and their differences were studied by the non-parametric Kruskal–Wallis test and Dunn's test with significance level $p < 0.05$. Similarly as in examination of the differences across sites, we wanted to see whether median feature values for these subsets were statistically significantly different and thus useful for indicating low or high μ_{RR}.

Results

Rolling resistance coefficient on forest road and on off-road forest terrain

The distributions of the rolling resistance coefficient (μ_{RR}) on forest road and on off-road forest terrain for the non-frost period of 2021 are shown in Fig. 5. For the whole data set, the variances between the two cases were not statistically different (Levene's test statistic was 0.076 with p value 0.783, null hypothesis of equal variances cannot be rejected), but the means (0.126 on forest road and 0.163 on off-road forest terrain) were, as the t-statistic was − 18.243 with a p value of 0.000.

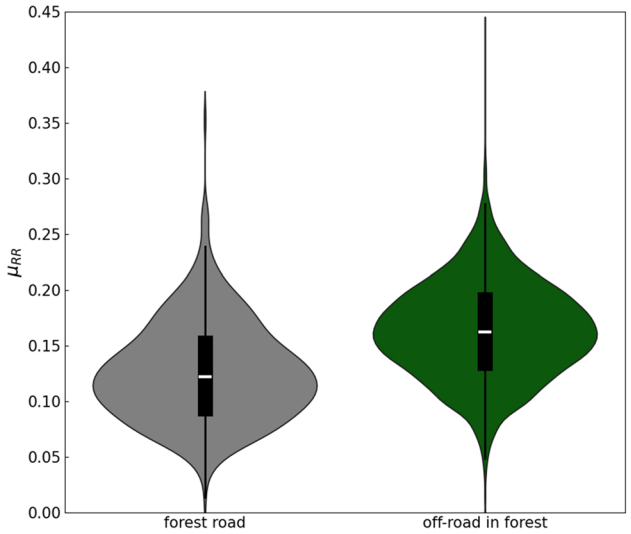

Fig. 5 The distribution of μ_{RR} on forest roads ($n = 490$ is the number of 8 m × 8 m grid cells) and in off-road forest conditions ($n = 5544$) across the non-frost period of 2021. The gray plot shows the kernel density estimation for the μ_{RR} distribution on forest roads and the green plot is the corresponding for off-road forest conditions. The median is indicated by the white line, and the black box represents the range between upper and lower quartiles. The vertical black line shows the minimum and maximum of the distribution excluding outliers

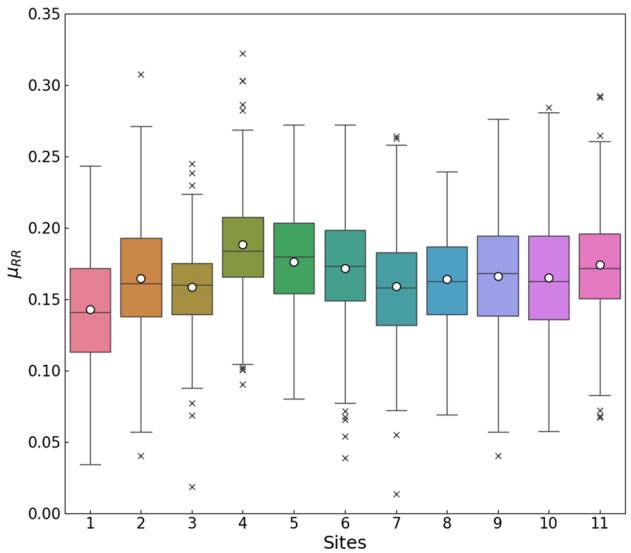

Fig. 6 Boxplot of the rolling resistance coefficient μ_{RR} per test sites. The median is indicated by the horizontal line in the box, the mean is indicated by a white circle, and the edges of the box are the upper and lower quartiles. The whiskers show the minimum and maximum of the distribution excluding the outliers, which are shown as crosses

Rolling resistance variability across and within the test sites

The distribution of μ_{RR} on each of the 11 test sites is shown in Fig. 6. The Kruskal–Wallis test indicated that the mean values of μ_{RR} differ between the sites, and Dunn's test (Supplementary material Table S3) showed how these differ from each other. The mean μ_{RR} on site 1 is clearly different from all the other sites (except from site 3 with p value 0.16 indicating that site 1 and 3 differ but only with a 84% confidence level). The mean μ_{RR} is 0.14 on site 1, clearly lower than on the other sites, where mean μ_{RR} is above 0.16. The highest mean μ_{RR} is found on site 4, which differs significantly from most of the other sites except site 5 (Supplementary material Table S3). On sites 4 and 5, μ_{RR} is on average higher (0.19 and 0.18, respectively) than on the other sites.

Detailed maps of the sites, ordered according to operation time, with μ_{RR} and the soil saturation deficit from SpaFHy shown, are presented in Fig. 7. The average site characteristics are shown in Table S2 of the Supplementary material. The two sites that represent the smallest and largest mean μ_{RR} (Fig. 6) have also been harvested at different times: site 1 at the beginning of May and site 4 in August. However, while they differ in μ_{RR}, the sites overall had rather similar hydrological conditions in terms of the saturation deficit, which on both sites was 0.03 on average.

Sites 1 and 2 are located next to each other (Figs. 1, 7) and were harvested at around the same time. Still, they differ in the mean μ_{RR} significantly. Site 2 has stronger variability in soil saturation deficit, and also wetter conditions overall (Fig. 7). Generally, sites 1–3, and 7–9 operated before mid-June show lower mean μ_{RR} than the other sites. Also site 10, operated in mid to late June, has lower μ_{RR} compared to sites 4, 5 & 6, which were operated in late June to mid-August.

Sites 4, 5 and 6 were harvested at around the same time. Site 6 has a mean μ_{RR} which is lower than on sites 4 and 5 but higher than on the rest of the sites. While mean μ_{RR} at sites 5, 6 and 11 cannot be separated statistically, the means of μ_{RR} on sites 4, 6 and 11 do differ significantly from each other. Site 11 was harvested earlier than sites 4–6. However, on all these sites, mean μ_{RR} is clearly above 0.17, and there were crossings across streams or ditches and operations close to the stream/ditch network. Also on site 1, where overall the μ_{RR} values were low, they were slightly higher in areas close to a stream in the western part of the target area (Fig. 7).

Relationship between rolling resistance and terrain features

Dunn's test performed for subsets of high and low μ_{RR} grid cells and the medians of the associated spatial data revealed terrain features that can best explain the within-site variation in μ_{RR}. The analysis (Table 1) enabled identifying those features which differ significantly ($p < 0.05$) between locations where high and low μ_{RR} was observed. Several topographical

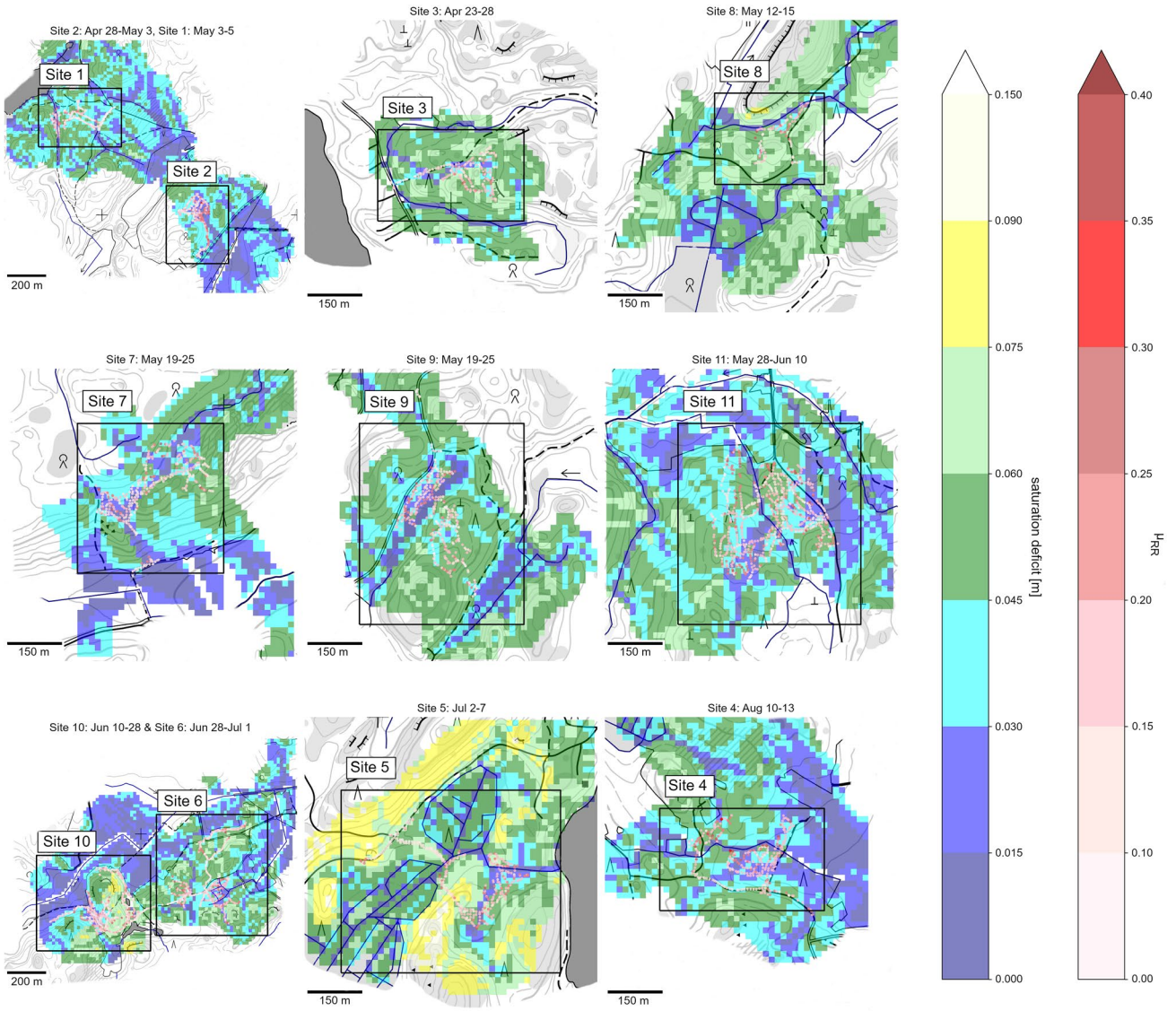

Fig. 7 Maps of μ_{RR} and hydrological model predicted soil saturation deficit during the operation of each test site. The lower the saturation deficit the wetter the soil. The harvester operating dates (in 2021) are given above each map. Sites in the upper two rows were operated before June 10 and sites in the bottom row June 10–August 13

wetness indices differ significantly between the locations corresponding to the extremes of μ_{RR}. High values of DTW with 4 ha threshold (indicating drier conditions) seem to contribute to low μ_{RR} values. High TWI values and also high root biomass or spruce volume values coincide with high μ_{RR} values. Table 1 is sorted so that the spatial features showing significant differences on several sites are shown at the top.

Discussion

Our study introduces a pioneering approach to compute the harvester rolling resistance coefficient (μ_{RR}) by utilising machine data collected from an operational Ponsse Ergo forest harvester over roughly 50 km of harvesting trails in actual working conditions. This extensive dataset encompasses a wide range of terrains and hydrometeorological conditions, representative of operational forest harvesting conditions from April 2021 to November 2021 in Southern Finland. To the best of our knowledge, this dataset is unprecedented in its scope and detail. The data provides a basis for predicting forest trafficability, aligning with the framework proposed in Salmivaara et al. (2020b). The collected dataset and our analysis of μ_{RR} variability support the previous hypothesis that forest harvester data can provide (indirect) information on soil strength for use in trafficability prediction (Ala-Ilomäki et al. 2020; Suvinen and Saarilahti 2006; Salmivaara et al. 2020b).

Table 1 The spatial features (explained in Table S1 of the Supplementary material) that relate to μ_{RR} extremes

Site Feature	1	2	3	4	5	6	7	8	9	10	11
aspect, slope aspect clockwise from north	**0.005**	**0.008**	nan	**0.025**	**0.011**	*0.003*	*0.046*	**0.042**	nan	*0.004*	nan
dtw04_0ha, depth-to-water index, 4 ha threshold	nan	nan	**0.022**	*0.000*	*0.000*	*0.001*	nan	*0.007*	*0.000*	nan	nan
aspect_sin,Sine of aspect	*0.001*	*0.026*	nan	nan	*0.003*	**0.000**	nan	*0.042*	nan	**0.001**	nan
dtw01_0ha,depth-to-water index, 1 ha threshold	nan	*0.000*	nan	*0.004*	*0.000*	*0.000*	nan	*0.012*	*0.000*	nan	nan
sat_deficit_**doy**,Saturation deficit, SpaFHy	nan	*0.000*	nan	nan	*0.001*	*0.000*	nan	nan	*0.004*	nan	*0.013*
bmroot_all, Root biomass	nan	**0.006**	nan	nan	**0.000**	**0.001**	**0.021**	nan	**0.001**	nan	nan
dtw00_5ha, depth-to-water index [m], 0.5 ha threshold	nan	*0.014*	nan	nan	nan	*0.001*	nan	*0.012*	*0.000*	nan	*0.019*
s_rootbm_ratio, Spruce root biomass per all root biomass	nan	**0.000**	nan	**0.034**	**0.000**	nan	nan	nan	**0.008**	*0.046*	nan
s_pulp_vol, Volume, spruce pulpwood [m^3ha^{-1}]	**0.004**	**0.000**	nan	nan	**0.000**	**0.001**	nan	nan	**0.003**	nan	nan
broadlv_vol, Volume, other broad-leaved trees	**0.031**	*0.010*	nan	nan	nan	**0.006**	**0.028**	**0.018**	nan	nan	nan
p_rootbm_**ratio**, Pine root biomass per all root biomass	*0.000*	nan	nan	nan	**0.000**	*0.000*	nan	*0.019*	*0.014*	nan	nan
dtw10_0ha, depth-to-water index, 10 ha threshold	*0.003*	nan	**0.022**	nan	**0.007**	nan	nan	*0.007*	*0.000*	nan	nan
twi, Topographic Wetness Index	nan	**0.000**	nan	nan	**0.001**	**0.000**	nan	nan	**0.004**	nan	**0.003**
tri, Terrain ruggedness index (TRI)	**0.020**	*0.006*	nan	*0.002*	nan	nan	nan	nan	*0.008*	nan	nan
topsoil, Coarseness class of superficial deposit	nan	nan	nan	nan	*0.000*	*0.000*	nan	*0.029*	*0.000*	nan	nan
slope, Slope calculated from dem	**0.022**	*0.006*	nan	*0.001*	nan	nan	nan	nan	*0.010*	nan	nan
s_vol, Volume, spruce	nan	**0.001**	nan	nan	**0.000**	**0.001**	nan	nan	**0.002**	nan	nan
p_timber_vol, Volume, pine timber	*0.027*	nan	nan	*0.016*	**0.000**	*0.026*	nan	nan	nan	nan	nan
p_pulp_vol, Volume, pine pulpwood	*0.001*	nan	nan	nan	**0.000**	*0.003*	nan	nan	*0.047*	nan	nan
s_timber_vol, Volume, spruce timber	nan	**0.014**	nan	nan	**0.000**	**0.031**	nan	nan	**0.002**	nan	nan
age, Stand age	nan	nan	**0.031**	nan	**0.000**	*0.020*	nan	nan	**0.001**	nan	nan
volume, Volume of the growing stock	nan	**0.003**	nan	nan	**0.000**	**0.003**	nan	nan	**0.004**	nan	nan
broadlv_pulp_vol, Volume, other broad-leaved pulpwood	**0.029**	*0.006*	nan	nan	nan	**0.004**	nan	**0.018**	nan	nan	nan
dem, Mean altitude above sea level	*0.005*	nan	nan	nan	**0.011**	nan	nan	*0.012*	*0.000*	nan	nan
diameter, Stand mean diameter	nan	**0.038**	nan	nan	**0.000**	*0.016*	nan	nan	**0.005**	nan	nan
p_vol, Volume, pine	*0.002*	nan	nan	nan	**0.000**	*0.012*	nan	nan	nan	nan	nan
vol_moisture_doy, Volumetric soil moisture, SpaFHy	nan	*0.005*	nan	nan	*0.013*	nan	nan	nan	**0.010**	nan	nan
b_pulp_vol, Volume, birch pulpwood	**0.016**	nan	nan	nan	nan	**0.000**	**0.001**	nan	nan	nan	nan
b_vol, Volume, birch	**0.029**	nan	nan	nan	nan	**0.000**	**0.001**	nan	nan	nan	nan
ba, Stand basal area	nan	**0.024**	nan	nan	**0.000**	nan	nan	nan	**0.003**	nan	nan
broadlv_timber_**vol**, Volume, other broad-leaved timber	**0.013**	nan	nan	nan	nan	*0.003*	nan	**0.011**	nan	nan	nan
diffMeanElev8, Local elevation difference	*0.024*	nan	nan	*0.028*	nan	nan	nan	nan	nan	*0.026*	nan
canopy_cov, Canopy cover	nan	nan	nan	nan	**0.000**	**0.000**	nan	nan	**0.039**	nan	nan
optimal_trafficability_**season**, classes	**0.013**	nan	nan	nan	*0.000*	nan	nan	nan	**0.002**	nan	nan
main_class_fertility_class_comb, combination	nan	nan	nan	nan	*0.010*	*0.000*	nan	nan	**0.034**	nan	nan
erosion_risk, Risk of surface soil to erosion	**0.014**	*0.022*	nan	nan	*0.000*	nan	nan	nan	nan	nan	nan
devMeanElev16, St. dev. of local elevation difference	*0.032*	nan	nan	nan	nan	*0.047*	nan	nan	*0.021*	nan	nan
c_rootbm_ratio, Conifer root biomass per all root biomass	nan	nan	nan	nan	**0.000**	*0.000*	nan	nan	nan	nan	nan
broadlv_rootbm_ratio, Broad-leaved root biomass ratio	nan	nan	nan	nan	*0.000*	**0.000**	nan	nan	nan	nan	nan
broadlv_canopycov, Canopy cover of broad-leaved trees	nan	nan	nan	nan	*0.001*	**0.000**	nan	nan	nan	nan	nan
standheight, Stand mean height	nan	**0.050**	nan	nan	**0.000**	nan	nan	nan	**0.001**	nan	nan
b_timber_vol, Volume, birch timber	nan	nan	nan	nan	nan	**0.001**	**0.030**	nan	nan	nan	nan
aspect_cos, Cosine of aspect	nan	nan	nan	nan	**0.000**	nan	nan	nan	*0.001*	nan	nan
roughness_scale, Roughness over a range of spatial scales	nan	*0.008*	nan	**0.047**	nan	nan	nan	nan	nan	nan	nan
roughness_mag, magnitude over a range of spatial scales	**0.009**	nan	nan	nan	nan	nan	nan	nan	nan	nan	nan

The Dunn's test *p* values are shown for those features which differ significantly ($p < 0.05$) between locations of low (below 10th percentile) and high (above 90th percentile) μ_{RR} at each site. The bold font indicates that an increase in the median of the spatial feature value leads to higher μ_{RR} (poorer trafficability), while italics indicates that an increase in the median feature value leads to lower μ_{RR} (better trafficability)

Refined transmission power model

In the model presented in Ala-Ilomäki et al. (2020), transmission power depended linearly on diesel engine power, and the model was developed using data from mostly level ground or slight uphill. Here, we created a refined model (Eq. 2) applicable to a wider range of engine power and enabling estimating transmission power also when the machine is driving downhill using engine braking. The values of μ_{RR} calculated with the Ala-Ilomäki et al. (2020) model were systematically lower than those calculated with our refined model. Apart from the different functional form, the previous model was developed for a Ponsse Scorpion harvester.

Up-scaling data collection: possibilities and limitations

Collecting μ_{RR} data from a fleet of forest harvesters as part of commercial forest operations was proposed by Salmivaara et al. (2020b) and Ala-Ilomäki et al. (2020). This would require calibrating the hydraulic transmission power model for different harvester models, and also possibly adjusting the machine parameters based on the status of their hydraulics maintenance program. Making such a calibration (Fig. 2) requires only installing pressure sensors to the harvester hydraulic transmission for the short driving tests on undulating terrain (Figs. 3, 4). Gathering such data would be crucial for developing a more comprehensive model to compute μ_{RR} from machine data of different harvester models and for different efficiency levels of their hydraulic transmission system. The P_t model developed in this work is machine-specific but provides a potential approach to develop similar P_t models for other harvester models.

In our approach to compute μ_{RR}, we assumed that in the 0.6 to 1.2 ms^{-1} speed range, diesel engine power was only consumed in machine movement, and power use in other functions such as boom usage was negligible. Our data collection plan did not include variables for enabling detailed analysis of boom usage power consumption. However, from a post-processing analysis of our data we estimated that boom usage can lead to a small overestimation of μ_{RR} (of the order of 0.01 to 0.02) compared to vehicle movement without any other power usage. This overestimation is rather small to significantly affect our interpretation of μ_{RR} variability. Also, the overestimation decreases with increasing power usage, and thus high μ_{RR}, which are of our main interest, are least affected. In future work, logging boom usage data would enable improved estimates of μ_{RR} distributions.

Furthermore, we assumed that wheel slip is negligible in the machine pitch range of −10 to 10° considered. Collecting both vehicle GNSS speed and wheel rotational speed and comparing them would enable detecting the wheel slip or skidding from machine data and subsequently removing these sections from the μ_{RR} calculation. Successful filtering requires, however, accurate GNSS speed, which may sometimes be difficult to obtain due to canopy cover (Kaartinen et al. 2015; Blum et al. 2016).

One further limitation of the demonstrated approach to estimating μ_{RR} is that information on the running gear the harvester was fitted with (tyres, tyres with chains, tyres with bogie tracks) at a given harvest site was not known. In the Nordic countries in snow-free conditions, the harvester is typically at least partly equipped with chained tyres, with bogie tracks being used only on the softest sites. However, changing the running gear will generally have an effect on the machine rolling resistance coefficient (Ala-Ilomäki 1993). Therefore, applying our approach to practical trafficability assessment and prediction in the future requires information on the fitted running gear.

Main factors contributing to variation in the rolling resistance coefficient

In our results, the mean of μ_{RR} varies in a narrower range across the 11 test sites (~ 0.14 to 0.19, Fig. 6) than the means across the test sites of Ala-Ilomäki et al. (2020) (~ 0.09 to 0.18). This difference can be attributed to the differences in scale and setup of data collection between the two studies. Our test sites, covering areas of approximately 400 m × 400 m, contrast with the smaller scale of the sites (5 m × 20 m) investigated by Ala-Ilomäki et al. (2020). In Ala-Ilomäki et al. (2020), a large variety of soil types and other site properties were purposefully chosen for the harvester to traverse with no regard to the produced rutting. In the present work, in contrast, the data was collected from regular, commercial forest operations. Most of the 11 sites were fully clear-cut, which means that the stands were of a mature age, resulting in a rather uniform forest structure across our test sites. Moreover, for minimising soil damage following normal guidelines for forest operations, the harvester routes were likely planned accordingly.

Creating a brush mat from logging residue on the logging trail is standard procedure in commercial harvesting operations such as those analysed in this study. On soft terrain, this improves the bearing capacity of the soil and reduces rut formation (Uusitalo and Ala-Ilomäki 2013; Poltorak et al. 2018), which could be expected to reduce μ_{RR}. However, at very good bearing conditions, the use of a brush mat may introduce additional motion resistance (Ala-Ilomäki et al. 2020), which would increase the computed μ_{RR}. The practically universal presence of brush mat in the present work may therefore be another factor contributing to the narrow range of μ_{RR} observed in our study.

We found significant differences in terrain trafficability, as described by μ_{RR}, between (Fig. 6) and particularly within

(Fig. 7) the test sites. Table 1 shows spatial features that differ significantly between locations where low or high (10th or 90th percentiles of the μ_{RR} distributions, respectively) rolling resistance coefficient values were observed. There are several features that link consistently to μ_{RR} extremes at most of the sites (Table 1).

Several topography-based wetness indices (depth-to-water index DTW, topographic wetness index `twi`) and the weather-history dependent soil saturation deficit (`sat_deficit_doy`) simulated by SpaFHy model are strong proxies of rolling resistance coefficient extremes. The rolling resistance coefficient had a significant negative relationship with `sat_deficit_doy` at five and with DTW indices `dtw01_0ha` and `dtw00_5ha` at 7 of the 11 test sites (Table 1), indicating that a high μ_{RR} was often associated with locations of high soil moisture (low DTW or saturation deficit). This is supported by the positive relationship between μ_{RR} and `twi`, as high `twi` indicates wet conditions. These findings agree with other studies that have found soil moisture and soil strength linkage (Schönauer et al. 2022; Campbell et al. 2013; Mohtashami et al. 2023). We also modelled the volumetric soil moisture content, which depends on antecedent weather conditions, stand attributes, soil hydraulic properties and topography (see Launiainen et al. (2019)). In the hydrological simulations, we used forest site fertility type dependent hydraulic properties (Launiainen et al. 2022), but still there was no consistent relation between `vol_moisture_doy` and μ_{RR} (Table 1). This is likely due to both high variability of soil hydraulic properties (Launiainen et al. 2022) and the partial inability of the SpaFHy-model to predict small-scale variability in soil water flow and soil moisture. Site-specifity is evident in our results, which agrees with other studies (Kemppinen et al. 2023; Schönauer et al. 2022), and certain sites show many more linkages to spatial features than other sites do (e.g., 3, 10 and 11). DTW maps with different thresholds have been estimated to represent different hydrological conditions, but it is clear that no single topographic wetness index outperforms the others. Our results show that DTW indices with smaller thresholds seem to work better on sites operated later in the year in more moist conditions (based on our hydrological model predictions).

While we did not assess rutting damage in this work, there is evidence that soil bearing capacity decreases and rut depth increases with increasing rolling resistance (Ala-Ilomäki et al. 2020; Salmivaara et al. 2020b; Kurjenluoma et al. 2009; Bygdén et al. 2003). Thus, our extensive high-resolution harvester data supports the use of topography-driven wetness indices as a practical means for trafficability prediction (Campbell et al. 2013; Heppelmann et al. 2022; Vega-Nieva et al. 2009; Salmivaara et al. 2020b; Mohtashami et al. 2023; Latterini et al. 2022b), particularly when the topographic wetness index is selected so that it reflects the landscape hydrological state during the operation (i.e., accounts for the weather history or variance between seasons (Kemppinen et al. 2023; Oltean et al. 2016; Niemi et al. 2017)). A practical means of including temporal information accompanied by consideration of vegetation, soil and climate on soil moisture is to use the soil saturation deficit simulated by a hydrological model, as suggested in Salmivaara et al. (2020b).

Unlike the wetness indices, the terrain ruggedness index (`tri`) and the difference of grid cell elevation relative to its neighbours (`diffMeanElev8`) reflect small-scale topographical variability without considering the upslope flow accumulation area of the cell or the position of the cell in the catchment (Riley et al. 1999; Lindsay 2023). The observed negative relationship between μ_{RR} and these local topography features suggests that rolling resistance increases at local surface depressions.

Rolling resistance variability was also strongly associated with stand attributes, here based on multi-source national forest inventory data (msNFI 2019; Mäkisara et al. 2022, Supplementary Table S1). High values of μ_{RR} were generally observed at locations where total tree root biomass `bmroot_all` (5 out of 11 sites), total tree `volume` (4 out of 11 sites) and/or basal area `ba` (3 out of 11 sites) were high (Table 1). Our data shows that μ_{RR} is high at locations with high spruce (`s_pulp_vol`, `s_timber_vol`, `s_rootbm_ratio`, on 4 to 5 out of 11 sites) or deciduous tree (`broadlv_vol`, `broadlv_pulp_vol`, on 3 to 4 out of 11 sites) biomass attributes. In contrast, locations where pine is the prevailing species (`p_rootbm_ratio`, `p_pulp_vol`, `p_timber_vol`, on 3 to 4 out of 11 sites) are often associated with low rolling resistance indicating terrain with good trafficability (Table 1). The data was collected mainly from clear-cut and some thinning operations conducted at closed-canopied, mature stands. Therefore, it is likely the variability in stand attributes, especially the species composition, reflects soil texture, site fertility and hydraulic regime. The high `volume` and root biomass (and high μ_{RR}) occur mainly at spruce and deciduous stands that typically dominate fertile and mesic locations, often on fine-textured soils with higher organic matter content (Launiainen et al. 2022; Muukkonen and Mäkipää 2006). This likely explains the high μ_{RR} observed at grid cells with large root biomass or `volume` (Table 1). However, due to denser forest structure, there might also be over-estimation in μ_{RR} due to more boom movement along the driving. We also observe a negative relationship between `topsoil` class (increasing from fine to coarse textured soils) and μ_{RR}, suggesting that high (low) extremes of rolling resistance coefficient occur more often at fine (coarse) soils, in line with Bygdén et al. (2003); Kurjenluoma et al. (2009); Ala-Ilomäki et al. (2020).

There are also features of which the relationship with μ_{RR} is not consistent across the sites (Table 1). Terrain slope `aspect` represents the facing of the slope (i.e., north at 0°, south at 180°). Thus, the metric is cyclic by nature, and requires a deeper analysis besides taking the sine and cosine of the aspect angle (also presented in Table 1). Figures S1 and S2 in the Supplementary material show, respectively, a radar plot of occurrence of aspects per each site and boxplots of μ_{RR} for each cardinal direction of the aspect for slopes of over 5° steepness, suggesting that south-facing slopes might have lower mean μ_{RR} than north-facing slopes. However, as can be seen from Table 1, the within-site effect of aspect can vary. While aspect seems to be a statistically significant factor for μ_{RR} variation, its effect on μ_{RR} is likely driven by a multitude of factors, and the interpretation requires future studies.

Lessons learnt and outlook

Our high-resolution harvester data collected during operational forest harvesting, supplemented by a specific one-day test run to calibrate the hydraulic power model parameters (Figs. 2, 3) enabled estimating the harvester rolling resistance coefficient for a range of sites and environmental conditions (Figs. 6, 7). The obtained μ_{RR} differed significantly between road and forest terrain (Fig. 5), and its variability was logically linked to variability of several terrain features obtained from open spatial data (Table 1). The results provide support for earlier studies proposing forest machines as cost-efficient platforms to collect extensive spatio-temporal data on trafficability and soil state (Ala-Ilomäki et al. 2020; Salmivaara et al. 2020b; Labelle et al. 2022). This concept would extend existing capabilities of collecting spatial data using forest harvesters, which include monitoring of thinning intensity (Möller et al. 2015), producing delineations for forest stands (Melkas et al. 2020), mapping the presence of root and butt rot disease (Räty et al. 2021), and collecting ground-truth data for remote-sensing forest inventory methods (Saukkola et al. 2019; Söderberg et al. 2021).

Across our test sites, the observed extremes of μ_{RR} were not consistently related to the static terrain trafficability map (`optimal_trafficability_season`, Table 1), which has been composed from topographic, site fertility type and vegetation attributes using expert judgement (Kankare et al. 2019; Finnish Forest Centre 2018). The pioneering dataset on vehicle ARCNET parameters and μ_{RR} (see details in Supplementary material Table S4) collected in this study opens interesting prospects for improved predictions of terrain trafficability. Salmivaara et al. (2020b) proposed that static harvest season classifications (Finnish Forest Centre 2018; Kankare et al. 2019) could be complemented with weather-dependent 'traffic light mapping'. The authors successfully demonstrated the approach by combining open spatial data, hydrological simulations and a harvester-measured rolling resistance coefficient from a tailored field experiment covering ca. 1.5 km of harvester tracks at a single site. The current study takes the first step in upscaling this approach to actual harvest operations, and to collect extensive μ_{RR} data to build data-driven models for trafficability forecasts (Labelle et al. 2022; Kemmerer and Labelle 2021; Hoffmann et al. 2022).

Forest operations are scheduled and planned based on multiple criteria, including economic, ecological and environmental (MacDicken et al. 2015; Latterini et al. 2022a) objectives. One of the goals is to minimise soil damages by careful planning of harvester and forwarder routes, on-site decisions by experienced harvester operators, and by use of brush mats on the softest areas (Uusitalo and Ala-Ilomäki 2013; Poltorak et al. 2018). The present work demonstrates the direct creation of a 'trafficability map' (Fig. 7) for the forwarder, which is an assessment of terrain trafficability produced from harvester data. Such a map could, in principle, be used to inform the planning of the subsequent forwarding operation in, e.g., on-site route planning. Furthermore, the *in situ* trafficability assessment provided by the harvester rolling resistance coefficient for a precise location and point in time could be used as a ground truth quantity in developing detailed trafficability predictions through any approach.

Our preliminary tests suggest that the tails of the μ_{RR} distribution can be distinguished (i.e., the extremes can be predicted) at the site-level, but not across the all sites combined (not shown). This may be due to coarser resolution of open data features compared to variability of rolling resistance, as well as low signal-to-noise ratio in both μ_{RR} and the predictor terrain features.

Conclusions

Our dataset of μ_{RR} collected from real commercial forest harvesting operations and machine parameters is unprecedented in extent and resolution, in both space and time. Our analysis demonstrates that operational harvester data can provide valuable information for trafficability assessment and prediction, in particular, for use in planning forwarding operations. We show that several wetness-related terrain indices, such as DTW with 4 ha and 1 ha thresholds and saturation deficit obtained from hydrological simulations, appear useful for predicting the occurrence of extreme μ_{RR} values at the site level. This enables identification of areas with poor and good trafficability conditions in a forest stand prior to any machine traffic. Using operational harvester data to infer and predict site trafficability is a useful addition to existing harvester data collection capabilities. In future work, we will

investigate the power of using μ_{RR} alongside open spatial data to predict rut depths resulting from forest operations.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10342-024-01717-6.

Acknowledgements The authors thank Metsätyö Eskola Oy for providing the Ponsse Ergo harvester for data collection, Ponsse Plc, particularly Mr. Simo Tauriainen, for technical assistance, Metsä Group for collaboration and Altisense Ltd. for collecting the point cloud data.

Author contributions AS, EH, SK, JA, SL, PV, PN and JH designed the research project and the analysis; AS, EH and JA performed data collection; data analysis and interpretation were performed by AS in collaboration with SK, PV, EH, JA and SL; all authors participated in writing/edition of the manuscript.

Funding Open access funding provided by Natural Resources Institute Finland. This research was funded by the Research Council of Finland (funding decisions 332172 and 332171) and the GreenFeedBack project from the EU Horizon Europe Framework Programme for Research and Innovation (Grant No. 101056921). In addition, this work was supported by the Research Council of Finland Flagship "Forest–Human–Machine Interplay—Building Resilience, Redefining Value Networks and Enabling Meaningful Experiences" (funding decision 337655).

Data availability The dataset is available as an electronic supplement at https://doi.org/10.5281/zenodo.11358276 under the terms and conditions of the Creative Commons Attribution (CC BY) licence (https://creativecommons.org/licenses/by/4.0). Sample of the data is shown in Table S4 of the Supplementary material.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Ågren A, Lidberg W, Strömgren M et al (2014) Evaluating digital terrain indices for soil wetness mapping—a Swedish case study. Hydrol Earth Syst Sci 18(9):3623–3634. https://doi.org/10.5194/hess-18-3623-2014

Ågren A, Lidberg W, Ring E (2015) Mapping temporal dynamics in a forest stream network-implications for riparian forest management. Forests 6(9):2982–3001. https://doi.org/10.3390/f6092982

Ågren A, Hasselquist E, Stendahl J et al (2022) Delineating the distribution of mineral and peat soils at the landscape scale in northern boreal regions. SOIL 8(2):733–749. https://doi.org/10.5194/soil-8-733-2022

Akumu C, Baldwin K, Dennis S (2019) GIS-based modeling of forest soil moisture regime classes: using Rinker Lake in northwestern Ontario, Canada as a case study. Geoderma 351:25–35. https://doi.org/10.1016/j.geoderma.2019.05.014

Ala-Ilomäki J (1993) A terramechanical approach for evaluating mobility and ground disturbance during skidding and forwarding: preliminary trials (les relations véhicule-sol dans l'évaluation de la mobilité et des perturbations du sol lors du débardage: essais préliminaires). Special Reports 86, FERIC

Ala-Ilomäki J, Salmivaara A, Launiainen S et al (2020) Assessing extraction trail trafficability using harvester CAN-bus data. Int J For Eng 31(2):138–145. https://doi.org/10.1080/14942119.2020.1748958

Arvidsson J, Sjöberg E, van den Akker J (2003) Subsoil compaction by heavy sugarbeet harvesters in southern Sweden: III. Risk assessment using a soil water model. Soil Tillage Res 73(1–2):77–87. https://doi.org/10.1016/S0167-1987(01)00168-4

Beven KJ, Kirkby MJ (1979) A physically based, variable contributing area model of basin hydrology/un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrol Sci J 24(1):43–69

Blum R, Bischof R, Sauter U et al (2016) Tests of reception of the combination of GPS and GLONASS signals under and above forest canopy in the Black Forest, Germany, using choke ring antennas. Int J For Eng 27(1):2–14. https://doi.org/10.1080/14942119.2015.1122945

Bygdén G, Eliasson L, Wästerlund I (2003) Rut depth, soil compaction and rolling resistance when using bogie tracks. J Terramech 40(3):179–190. https://doi.org/10.1016/j.jterra.2003.12.001

Campbell D, White B, Arp P (2013) Modeling and mapping soil resistance to penetration and rutting using LiDAR-derived digital elevation data. J Soil Water Conserv 68(6):460–473. https://doi.org/10.2489/jswc.68.6.460

Finnish Forest Centre (2018) Static trafficability maps. https://avoin.metsakeskus.fi/aineistot/Korjuukelpoisuus/. Accessed 01 Jan 2023

GSF, The Geological Survey of Finland (2015) Superficial deposits 1:20 000, 1:50 000 and 1:200 000. https://hakku.gtk.fi/en

Hansson L, Šimůnek J, Ring E et al (2019) Soil compaction effects on root-zone hydrology and vegetation in boreal forest clearcuts. Soil Sci Soc Am J 83:S105–S115. https://doi.org/10.2136/sssaj2018.08.0302

Heppelmann J, Talbot B, Antón Fernandéz C et al (2022) Depth-to-water maps as predictors of rut severity in fully mechanized harvesting operations. Int J For Eng 33(2):108–118. https://doi.org/10.1080/14942119.2022.2044724

Hoffmann S, Schönauer M, Heppelmann J et al (2022) Trafficability prediction using depth-to-water maps: the status of application in Northern and Central European forestry. Curr For Rep 8:55–71. https://doi.org/10.1007/s40725-021-00153-8

IPCC (2021) Climate change 2021: The physical science basis. Contribution of Working Group I to the sixth assessment report of the Intergovernmental Panel on Climate Change; Technical summary. In: Masson-Delmotte V, Zhai P, Pirani A, et al (eds) The Intergovernmental Panel on Climate Change AR6. https://elib.dlr.de/137584/

Jones MF, Arp P et al (2019) Soil trafficability forecasting. Open J For 9(04):296–322. https://doi.org/10.4236/ojf.2019.94017

Kaartinen H, Hyyppä J, Vastaranta M et al (2015) Accuracy of kinematic positioning using global satellite navigation systems under forest canopies. Forests 6(9):3218–3236. https://doi.org/10.3390/f6093218

Kankare V, Luoma V, Saarinen N et al (2019) Assessing feasibility of the forest trafficability map for avoiding rutting—a case study. Silva Fenn. https://doi.org/10.14214/sf.10197

Kemmerer J, Labelle E (2021) Using harvester data from on-board computers: a review of key findings, opportunities and challenges. Eur J For Res 140(1):1–17. https://doi.org/10.1007/s10342-020-01313-4

Kemppinen J, Niittynen P, Rissanen T et al (2023) Soil moisture variations from boreal forests to the tundra. Water Resour Res. https://doi.org/10.1029/2022WR032719

Kurjenluoma J, Alakukku L, Ahokas J (2009) Rolling resistance and rut formation by implement tyres on tilled clay soil. J Terramech 46(6):267–275. https://doi.org/10.1016/j.jterra.2009.07.002

Labelle E, Hansson L, Högbom L et al (2022) Strategies to mitigate the effects of soil physical disturbances caused by forest machinery: a comprehensive review. Curr For Rep 8(1):20–37. https://doi.org/10.1007/s40725-021-00155-6

Lagacherie P, McBratney A (2006) Chapter 1 Spatial soil information systems and spatial soil inference systems: perspectives for digital soil mapping. In: Lagacherie P, McBratney A, Voltz M (eds) Digital soil mapping, developments in soil science, vol 31. Elsevier, Amsterdam, pp 3–22. https://doi.org/10.1016/S0166-2481(06)31001-X

Latterini F, Stefanoni W, Venanzi R et al (2022) GIS-AHP approach in forest logging planning to apply sustainable forest operations. Forests 13(3):484. https://doi.org/10.3390/f13030484

Latterini F, Venanzi R, Tocci D et al (2022) Depth-to-water maps to identify soil areas that are potentially sensitive to logging disturbance: initial evaluations in the Mediterranean forest context. Land 11(5):709. https://doi.org/10.3390/land11050709

Launiainen S, Guan M, Salmivaara A et al (2019) Modeling boreal forest evapotranspiration and water balance at stand and catchment scales: a spatial approach. Hydrol Earth Syst Sci 23(8):3457–3480. https://doi.org/10.5194/hess-23-3457-2019

Launiainen S, Kieloaho AJ, Lindroos AJ et al (2022) Water retention characteristics of mineral forest soils in Finland: impacts for modeling soil moisture. Forests 13(11):1797. https://doi.org/10.3390/f13111797

Lidberg W, Nilsson M, Ågren A (2020) Using machine learning to generate high-resolution wet area maps for planning forest management: a study in a boreal forest landscape. Ambio 49(2):475–486. https://doi.org/10.1007/s13280-019-01196-9

Lindsay J (2023) WhiteboxTools 2.3.0, released 2023-03-25. https://www.whiteboxgeo.com/manual/wbt_book/preface.html

MacDicken K, Sola P, Hall J et al (2015) Global progress toward sustainable forest management. For Ecol Manag 352:47–56. https://doi.org/10.1016/j.foreco.2015.02.005

Mäkisara K, Katila M, Peräsaari J (2022) The multi-source national forest inventory of Finland—methods and results 2017 and 2019. Natural resources and bioeconomy studies 90/2022, Natural Resources Institute Finland (Luke), Helsinki

Melander L, Einola K, Ritala R (2020) Fusion of open forest data and machine fieldbus data for performance analysis of forest machines. Eur J For Res 139(2):213–227. https://doi.org/10.1007/s10342-019-01237-8

Melkas T, Riekki K, Sorsa JA (2020) Automated method for delineating harvested stands based on harvester location data. Remote Sens 12(17):2754. https://doi.org/10.3390/rs12172754

Mohtashami S, Bergkvist I, Löfgren B et al (2012) A GIS approach to analyzing off-road transportation: a case study in Sweden. Croat J For Eng 33(2):275–284

Mohtashami S, Eliasson L, Jansson G et al (2017) Influence of soil type, cartographic depth-to-water, road reinforcement and traffic intensity on rut formation in logging operations: a survey study in Sweden. Silva Fenn. https://doi.org/10.14214/sf.2018

Mohtashami S, Eliasson L, Hansson L et al (2022) Evaluating the effect of DEM resolution on performance of cartographic depth-to-water maps, for planning logging operations. Int J Appl Earth Obs Geoinf 108:102728. https://doi.org/10.1016/j.jag.2022.102728

Mohtashami S, Hansson L, Eliasson L (2023) Estimating soil strength using GIS-based maps—a case study in Sweden. Eur J For Eng 9(2):70–79. https://doi.org/10.33904/ejfe.1321075

msNFI (2019) Multi-source national forest inventory (MS-NFI) raster maps. Natural Resources Institute Finland. https://kartta.luke.fi/index-en.html

Murphy P, Ogilvie J, Meng FR et al (2011) Modelling and mapping topographic variations in forest soils at high resolution: a case study. Ecol Modell 222(14):2314–2332. https://doi.org/10.1016/j.ecolmodel.2011.01.003

Muukkonen P, Mäkipää R (2006) Empirical biomass models of understorey vegetation in boreal forests according to stand and site attributes. Boreal Environ Res 11:355–369

Möller J, Bhuiyan N, Hannrup B (2015) Development and test of decision-support tool for automated monitoring of thinning, arbetsrapport från Skogforsk, nr. 862-2015. Technical report, Skogforsk. https://shorturl.at/gwDI5

National Land Survey of Finland (2023) Topographic database. Data retrieved from https://asiointi.maanmittauslaitos.fi/karttapaikka/tiedostopalvelu/maastotietokanta

Niemi M, Vastaranta M, Vauhkonen J et al (2017) Airborne LiDAR-derived elevation data in terrain trafficability mapping. Scand J For Res 32(8):762–773. https://doi.org/10.1080/02827581.2017.1296181

Oltean G, Comeau P, White B (2016) Linking the depth-to-water topographic index to soil moisture on boreal forest sites in Alberta. For Sci 62(2):154–165. https://doi.org/10.5849/forsci.15-054

Poltorak B, Labelle E, Jaeger D (2018) Soil displacement during ground-based mechanized forest operations using mixed-wood brush mats. Soil Tillage Res 179:96–104. https://doi.org/10.1016/j.still.2018.02.005

Räty J, Breidenbach J, Hauglin M et al (2021) Prediction of butt rot volume in Norway spruce forest stands using harvester, remotely sensed and environmental data. Int J Appl Earth Obs Geoinf 105:102624. https://doi.org/10.1016/j.jag.2021.102624

Reeves D, Reeves M, Abbott A et al (2012) A detrimental soil disturbance prediction model for ground-based timber harvesting. Can J For Res 42(5):821–830. https://doi.org/10.1139/X2012-034

Riley S, DeGloria S, Elliot R (1999) Index that quantifies topographic heterogeneity. Intermt J Sci 5(1–4):23–27

Saarilahti M (2002) Soil interaction model. Development of a protocol for ecoefficient wood harvesting of sensitive sites (ECOWOOD), Deliverable D2, University of Helsinki

Salmivaara A, Launiainen S, Perttunen J et al (2020) Towards dynamic forest trafficability prediction using open spatial data, hydrological modelling and sensor technology. Forestry 93(5):662–674. https://doi.org/10.1093/forestry/cpaa010

Saukkola A, Melkas T, Riekki K et al (2019) Predicting forest inventory attributes using airborne laser scanning, aerial imagery, and harvester data. Remote Sens 11(7):797. https://doi.org/10.3390/rs11070797

Schönauer M, Hoffmann S, Maack J et al (2021) Comparison of selected terramechanical test procedures and cartographic indices to predict rutting caused by machine traffic during a cut-to-length thinning operation. Forests 12(2):113. https://doi.org/10.3390/f12020113

Schönauer M, Prinz R, Väätäinen K et al (2022) Spatio-temporal prediction of soil moisture using soil maps, topographic indices and SMAP retrievals. Int J Appl Earth Obs Geoinf 108:102730. https://doi.org/10.1016/j.jag.2022.102730

Shabani S, Najafi A, Majnonian B et al (2019) Spatial prediction of soil disturbance caused by forest logging using generalized additive models and GIS. Eur J For Res 138(4):595–606. https://doi.org/10.1007/s10342-019-01180-8

Söderberg J, Wallerman J, Almäng A et al (2021) Operational prediction of forest attributes using standardised harvester data and airborne laser scanning data in Sweden. Scand J For Res 36(4):306–314. https://doi.org/10.1080/02827581.2021.1919751

Storn R, Price K (1997) Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. J Glob Optim 11(11):341–359. https://doi.org/10.1023/A:1008202821328

Suvinen A (2006) A GIS-based simulation model for terrain tractability. J Terramech 43(4):427–449. https://doi.org/10.1016/j.jterra.2005.05.002

Suvinen A, Saarilahti M (2006) Measuring the mobility parameters of forwarders using GPS and can bus techniques. J Terramech 43(2):237–252. https://doi.org/10.1016/j.jterra.2005.12.005

Uusitalo J, Ala-Ilomäki J (2013) The significance of above-ground biomass, moisture content and mechanical properties of peat layer on the bearing capacity of ditched pine bogs. Silva Fenn. https://doi.org/10.14214/sf.993

Vega-Nieva D, Murphy P, Castonguay M et al (2009) A modular terrain model for daily variations in machine-specific forest soil trafficability. Can J Soil Sci 89(1):93–109. https://doi.org/10.4141/CJSS06033

Venäläinen A, Lehtonen I, Laapas M et al (2020) Climate change induces multiple risks to boreal forests and forestry in Finland: a literature review. Glob Change Biol 26(8):4178–4196. https://doi.org/10.1111/gcb.15183

Wong J (2001) Theory of ground vehicles. Wiley, New York

Yong R, Fattah E, Skiadas N (1984) Vehicle traction mechanics. Elsevier, Amsterdam