Bijlage IV TNO rapport onderzoek schadeprotocol

TNO report

TNO 2020 R 10620

A revision of ProRail's damage-protocol on train induced vibration in the light of new measurement data

Date	17 April 2020
Author(s)	C.P.W. Geurts, Á. Rózsás
Copy no	
No. of copies	
Number of pages	82 (incl. appendices)
Number of appendices	7
Customer	
Project name	ProRail trillingsprotocol
Project number	060.37686

Contents

Samenvatting

ProRail hanteert een protocol voor de afhandeling van schadeclaims als gevolg van trillingen door treinverkeer. In dit protocol wordt op basis van een aantal basisgegevens een afstand gegeven waarbuiten schade door treinverkeer niet aannemelijk is. Deze afstanden zijn gegeven voor verschillende liggingen van het spoor, de samenstelling van de ondergrond en voor de eigenschappen van het beschouwde gebouw (gebouwhoogte en gevoeligheid voor trillingen). Daarnaast wordt onderscheid gemaakt in situaties waarbij wel dan wel niet sprake is van discontinuiteiten in de buurt van het beschouwde gebouw.

Dit rapport presenteert de achtergronden bij de afleiding van de afstanden in het protocol. Deze zijn gebaseerd op analyses van recente datasets van trillingen ten gevolge van treinverkeer op meerdere locaties in Nederland.

Gebaseerd op deze nieuwe data is een update van de afstanden bepaald ten behoeve van het protocol. Deze nieuw afstanden zijn hieronder gegeven.

Tabel A: Afstanden waarbuiten schade door trillingen niet aannemelijk is (beoordelingsstap 1).

Gebied	**spoorbaan**	**Metselwerk, niet trillingsgevoelig**		**Metselwerk, trillingsgevoelig**	
		H < 12 m	*H* > 12 m	*H* < 12 m	*H* > 12 m
1	Aardebaan	34.4	16.4	61.9	34.4
1	Kunstwerk	12.2	4.91	26.9	12.2
2	Aardebaan	24.3	8.73	67.4	24.3
2	Kunstwerk	11.9	4.28	33.0	11.9

Tabel B: Afstanden waarbuiten schade door trillingen niet aannemelijk is (beoordelingsstap 2, aanwezigheid van discontinuiteiten is uitgesloten).

Gebied	**spoorbaan**	**Metselwerk, niet trillingsgevoelig**		**Metselwerk, trillingsgevoelig**	
		H < 12 m	*H* > 12 m	*H* < 12 m	*H* > 12 m
1	Aardebaan	20.8	9.03	41.7	20.8
1	Kunstwerk	6.52	2.50	15.7	6.52
2	Aardebaan	8.73	3.14	24.3	8.73
2	Kunstwerk	4.28	1.54	11.9	4.28

In hoofdstuk 5 van dit rapport is een voorstel gedaan voor afronding van deze getallen en is de vergelijking met de afstanden in het huidige protocol gegeven.

Deze tabellen zijn gebaseerd op een meetduur van 7 dagen. In dit rapport is een beschouwing gegeven van de invloed van meetduur op de voorspelling van de trillingsniveaus op basis van lange duur data van drie meetlocaties. Uit deze analyses volgt dat de topwaarde bij een meetperiode van 1 maand ongeveer 10% hoger ligt dan bij een meetperiode van 1 week. Bij langer wordende meetperiodes

wordt een verdere verhoging gevonden. De mate van verhoging vlakt snel af voor meetperiodes van 3 maanden of langer.
Geadviseerd wordt om nader onderzoek te verrichten naar deze effecten om op basis hiervan te besluiten of aanpassingen aan het protocol wenselijk zijn.

1 Introduction

This introductory section provides the motivation of the study, delineates the research questions, outlines the adopted approach, and the scope.

1.1 Motivation and problem statement

Railway traffic is a source of vibrations which are felt in the direct surroundings of railway tracks. ProRail, as the owner of the railway infrastructure, frequently receives damage claims from house owners. To process these claims, ProRail uses a damage claim protocol. Within this protocol, a threshold distance has been defined. This distance is chosen so that outside the threshold damage due to vibrations is very unlikely. Within the threshold the probability of damage caused by the railway is not negligible, and a further investigation of the damage claim is initiated.
The original protocol is based on the backgrounds of SBR guideline A from 2002. In 2017 this guideline has been revisited. Following this revision, ProRail has asked TNO to update the analysis to define the threshold distances. The SBR thresholds correspond to a probability of damage of 1%. The goal of this work is to find a threshold distance for which the probability of damage equal to 1 %. This distance is related to the SBR vibration velocity thresholds.

1.2 Approach

The approach used in this study is a statistical one with the following main steps (see also Figure 1-1):

1. Develop a peak(top) velocity attenuation function: fit a mathematical model to the measurements. The model accounts for:
    a. uncertainties (aspects/phenomena which are not explicitly included in the model);
    b. the differences in measurement duration: bring all measurements to a common denominator: same measurement duration.
2. Establish and/or select the peak(top) velocity threshold (v_{th}) and target exceedance probability of peak(top) velocity threshold (P_{th}).
3. Using the model from 1. and the threshold from 2., find the threshold distance (R_{th}) that corresponds to P_{th} exceedance probability.

The approach is largely based on the previous revision of the protocol (Roos, 2004a, 2004b) but extends it in a few ways:

- Explicit consideration of the measurement duration. The velocities over time are modelled as a stochastic process. Due to this the longer the exposure period to the vibration, the more likely the higher vibration velocities are.
- The uncertainty component in the vibration velocity vs distance to track model is changed from Rayleigh to Gumbel distribution that is supported by extreme value theorem (additionally GEV and Normal distributions are compared). One reason is for this is that the single parameter Rayleigh distribution has a fixed coefficient of variation (0.52), hence cannot follow the measurements if they showcase a different variability.

- The vibration velocity vs distance to track model is formulated in a physically more consistent way in order to avoid negative vibration velocity values.
- The model is explicitly based on vibration measurements and it shows how new measurements can be incorporated into the model. This aspect is not clear in (Roos, 2004a).

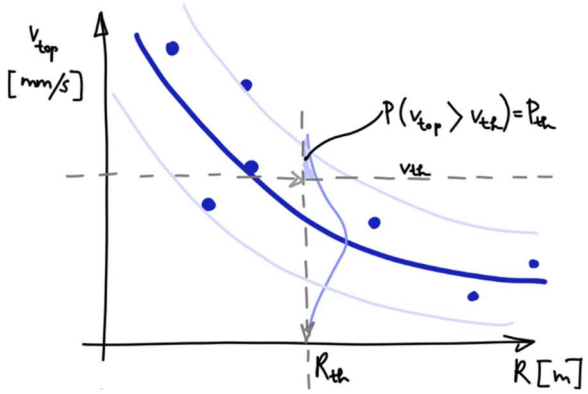


Figure 1-1: A schematic overview of the approach used for deriving threshold distances for train induced vibrations.

1.3 Scope and limitations

The following main assumptions are made in this study:

- The Barkan formula is used to describe the maximum vibration velocity attenuation over distance.
- Vibration maxima per time intervals, such as one-day or entire measurement period, are used for fitting.
- All fitted models and derived distance thresholds are based on the data available to us and should not be used outside of the coverage of the data. For example this means that due to lack of data we could not differentiate between passenger and cargo trains, although the used model could accommodate that.
- The measurement data available to us does not allow for such a differentiation between threshold distances as in (Roos, 2004a), e.g. distinguishing between cargo and passenger trains.
- Vibration maxima over time, e.g. daily maxima, are assumed to be independent. As far as the available data allowed it we checked this assumption and found it to be likely valid.
- Gumbel, GEV, and Normal distributions are considered to describe the uncertainty component in the maximum velocity attenuation model.
- Vibration threshold values are adopted from the SBR Guideline A (SBRCURnet, 2017): Design values are $v_{th} = 2$ mm/s and $v_{th} = 1.2$ mm/s.
- Target probability of vibration induced damage over the vibration exposure period is adopted from SBR Guideline A: $P_d = 1$ %.
- The vibration as an action is the non-dominant component in the damage reliability ($\alpha_{vibration} = 0.4$, α is the sensitivity factor).

2 Data

This section presents an overview of the available and used data.

2.1 Overview

Based on their source and temporal resolution, five different databases (data sources) are considered in this work. An overview of them is provided in Table 2-1. It is salient that the temporal resolutions and the available amount of data differ significantly for the different data sources. By far the most data is available for the N-P-SM data set, about 50000 data points. However, that data is coming from only three different locations hence provides limited information on the attenuation of velocities over distance to the track. This already highlights one of the major challenges: the data is available on different temporal scales and each is limited, hence their combination is needed to derive reliable and accurate models.

As a first step the data from the considered five sources (Table 2-1) is collected and cleaned, and put into databases which allow for more convenient computer based analyses. The conversion steps are documented in the Glossary.

The following sections provide additional information about the five data sources. Moreover, they present some exploratory data analyses mostly in the form of descriptive statistics and visuals in order to gain insight into the data and to inspire subsequent modelling decisions.

Table 2-1: Overview of the vibration measurement data considered and their IDs used in this report.

Data source†	Block size*	Data from block	Section	ID‡	Amount of data
old (Heijnen, 1996)	measurement duration	single maximum per block	2.2	O-M-SM	in total 36 maxima
new	measurement duration	single maximum per block	2.3	N-M-SM	in total 110 maxima
new	day	single maximum per block	2.4	N-D-SM	7 locations, in total 60 maxima
new	passage	single maximum per block	2.5	N-P-SM	3 locations, in total 51733 maxima
new	passage	time series	2.6	N-P-TS	in total 10 time series (10-40s long)

* Refers to the time interval from which individual data “units” are available. For maximum type of data the unit (data from block) is a single scalar value, while for others it is an entire time series corresponding to the block size.

† New refers to data that was provided to TNO for this project.

‡ The IDs used in this report to concisely refer to the databases/data sources.

2.2 O-M-SM: measurement duration maxima

In this section the old measurements concerning a single maxima per measurement duration/campaign are described and looked into in more detail. The data is obtained from Heijnen (1996) by digitizing the figures and calculating the marker coordinates from the digitized images. Due to this, currently only the maximum vibration velocity and distance to track pairs are available for this database. The 36 velocity maxima—distance to track pairs are shown in Figure 2-1. The measurements are accompanied with allowed maximum velocities from van Staalduinen (1999). With the exception of a single measurement point all measurements fall below both allowed maximum velocity envelopes.

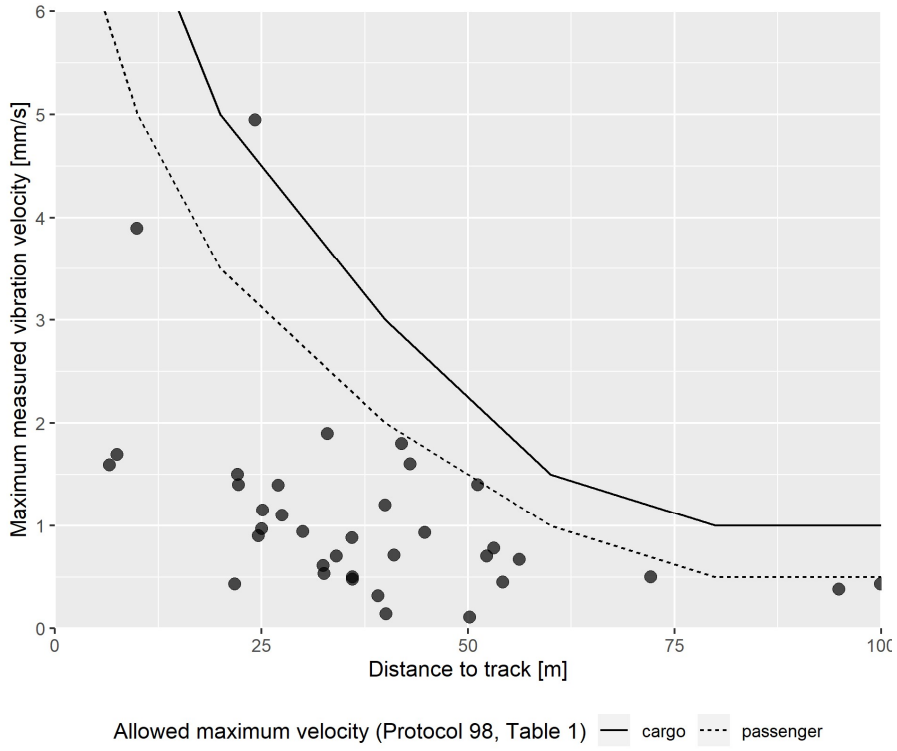


Figure 2-1: Old maximum measured vibration velocities and allowed maximum velocity envelopes from van Staalduinen (1999), Table 1 (gebied I). Data: O-M-SM.

2.3 N-M-SM: measurement duration maxima

Updating the protocol started with collecting a new series of measurement outcomes from different situations. Based on information made available by ProRail, Cauberg Huygen, and Quattro Expertise, a database was built in the form of an Excel sheet. Every line in the sheet contains information on a separate project/measurement. This database is here referred to as N-M-SM.

Per measurement, the following information was included in the database:

Column 1: Party who performed the measurements
Column 2: Identifier of the measurement; a project number or other, assigned by the party that performed the measurements.

Column 3: Duration of the measurement (in calendar days; the first and last day are counted as 1 day)
Column 4: Maximum observed velocity Vtop in mm/s during the full duration
Column 5: The dominant frequency at Vtop
Column 6: Date
Column 7: Time of occurrence of Vtop
Column 8: Is there a crossing or civil structure nearby (Yes or No); If Yes, these data have been used to analyze the effect of a structure. If No, the data were related to the distance to the track.
Column 9: If Yes to 8, at what distance is the structure from the measurement.
Column 10: Distance to the track
Column 11: Type of track
Column 12: Type of train (passenger, cargo, both) during the measurements
Column 13: Maximum allowed train speed
Column 14: Exact speed, if known
Column 15: Type of building
Column 16: Vibration sensitive or not ?
Column 17: Type of soil
Column 18: Damage (Yes or No)

In total 110 observations have been collected. 7 of these measurements concerned with only passenger trains, 3 with only cargo and 100 were mixed.
For none of these measurement campaigns, the question whether there was damage was answered with Yes. 48 were unknown, 62 were answered with No. This made it impossible to find direct relations between probability of damage and vibration velocities.
No information is available in the database regarding the exact measurement locations of the measurements. Its representativeness has been judged from proxies such as soil type, building type, track type, train type, etc. (see for example Figure 2-2 and Annex B.1).

This section deals with these new measurements concerning a single maxima per measurement duration/campaign. See Table Glossary-0-1 for the used data source names and for the interpretation of terms (particularly the translation of the Dutch terms to English).

Here only a few selected figures are presented which are deemed to be informative and supporting the modelling decision in later sections. Additional figures are presented in Annex B.1.

First the new measurements are plotted against the old ones to see if changes have occurred in measured vibration velocities over the decades separating the two data sources; see Figure 2-2. The old and new measurements are in a good agreement, no salient differences are observable. In addition to the measurement points the allowed maximum vibration velocities according to the 1998 Protocol (van Staalduinen, 1999) are also presented. With the exception of one measurement from each dataset all datapoints fall below the cargo threshold line.

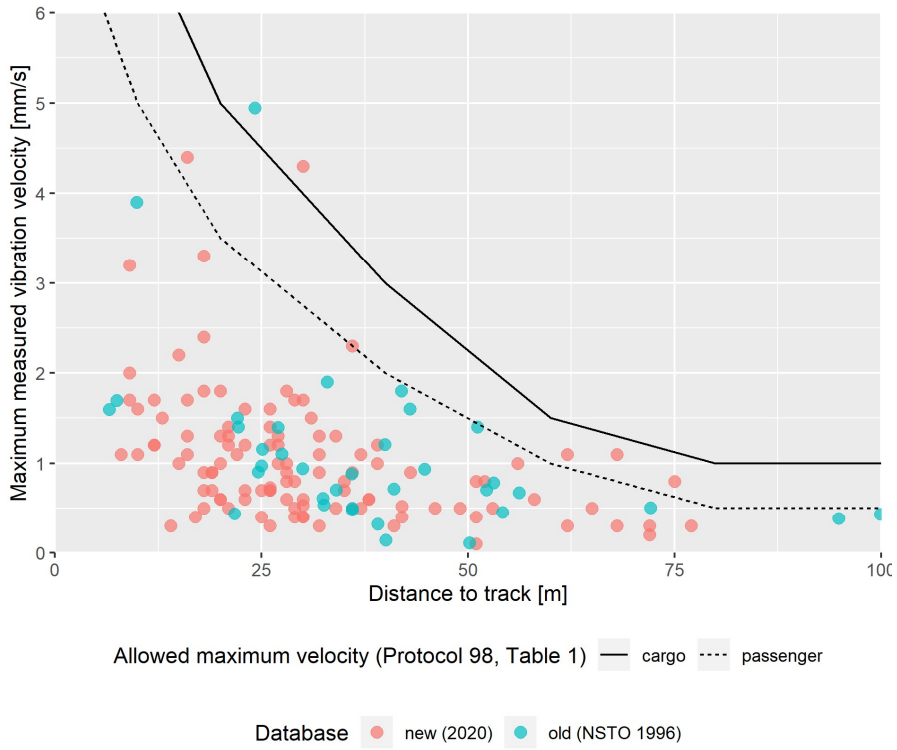


Figure 2-2: Old and new maximum measured vibration velocities. Data: O-M-SM and N-M-SM. The lines are the allowed maximum velocity envelopes from van Staalduinen (1999), Table 1 (gebied I).

In case of the N-M-SM data, we have considerable additional information about the measurements and their location. The rest of this section explores the influence of these additional factors on the maximum measured vibration velocity.
Since the measurement durations vary from 7 to 40 days, the maximum measured values are plotted against the distance to track while coloring the markers according to the measurement duration; see Figure 2-3. The figure shows that the velocity maxima from longer measurement durations, on average are higher than those of shorter measurement duration. However, clear conclusions are hard to draw since the number of longer measurement durations is relatively low. Note that the observed average behavior is the expected one: the more time available, the higher the observed velocities, or equivalently the higher the probability of seeing a given velocity.

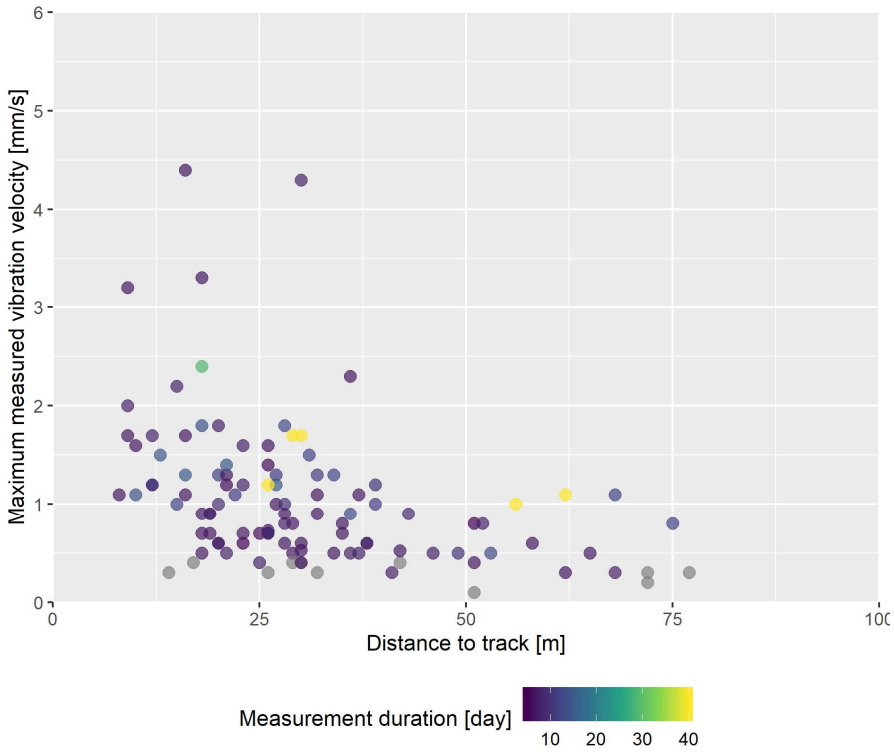


Figure 2-3: The impact of track distance and measurement duration on the maximum measured vibration velocity. The markers are colored according to their measurement duration in day. The grey markers are measurements for which no duration is available in the database. Data: N-M-SM.

Since the old protocol differentiates between and provides different distance thresholds based on soil type (x2), railway truck substructure (x2), train type (x2), and all their possible combinations, we visualized the N-M-SM dataset in light of these properties to see if visually discernible groups emerge. The plot are presented in Figure 2-4 and Figure 2-5 (combined). Figure 2-5 shows only one clear clustering: measurements from locations with structure nearby (Overweg of Kw = true) are in general higher than those without a structure nearby. It is a less clear clustering but on average the maximum velocities on soft soils (clay and loam) seem to be higher than those of on the stiff sand. In terms of train type no clear pattern is recognizable, this is largely due to that 93% of trains were classified as ‘cargo and passenger’, i.e. there is not enough variability in the data. Due to this lack of variability in N-M-SM, this dataset is insufficient in itself to derive distance thresholds with such a resolution (differentiation between soils, substructures, and trains) as in the previous protocols. To our knowledge, fewer measurements were available for deriving the recommendations of the old protocol.

Figure 2-4: Visualization of the effect of soil type and train type on the attenuation of maximum vibration velocity over distance. The black dots inside the markers indicate if there is a structure nearby (Overweg of Kw = true), the absence of a dot indicates that there is no structure nearby. The two plots are combined in Figure 2-5. Data: N-M-SM.

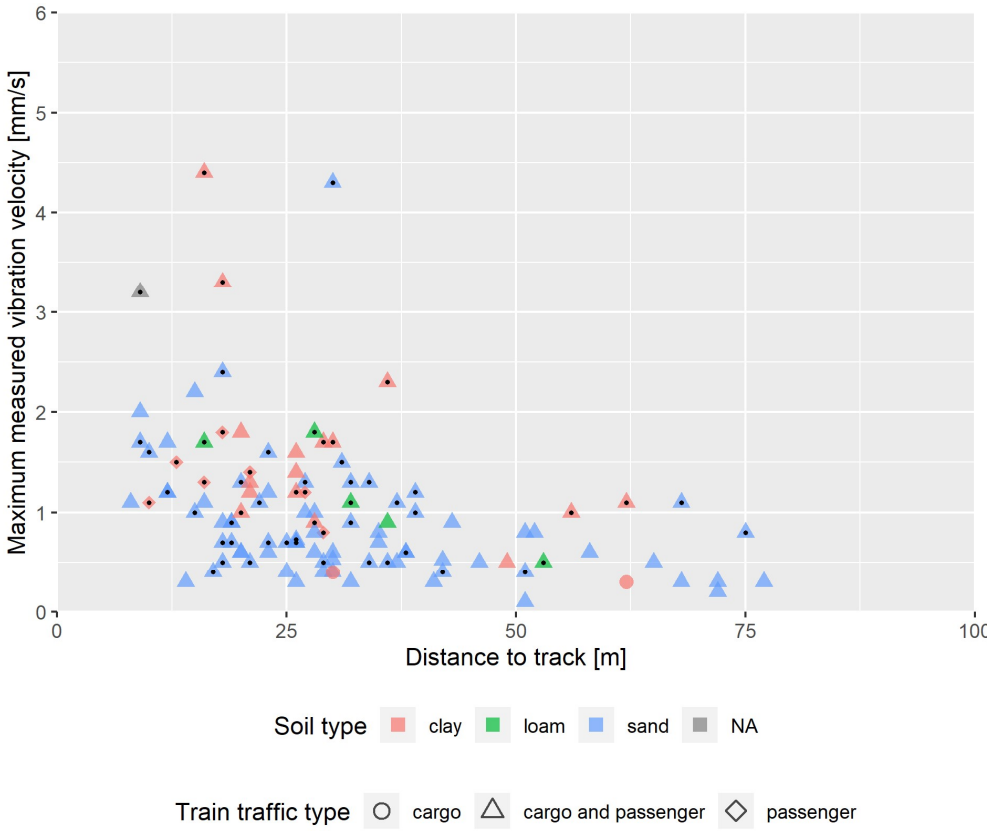


Figure 2-5: Visualization of the effect of soil type and train type on the attenuation of maximum vibration velocity over distance. The black dots inside the markers indicate if there is a structure nearby (Overweg of Kw = true), the absence of a dot indicates that there is no structure nearby. The point with and without a structure nearby are separately plotted in Figure 2-4. Data: N-M-SM.

2.4 N-D-SM: daily maxima

For seven locations not only the maximum over the entire measurement duration but the daily maxima are also available. To show the variability in daily maxima they are plotted over time per location in Figure 2-6. It is visible that for most locations the daily maxima exhibit low daily variability. This daily variability and correlation between daily maxima are further analyzed in the next section which concerns measurement data of three locations with longer measurement duration and has maxima per passage.

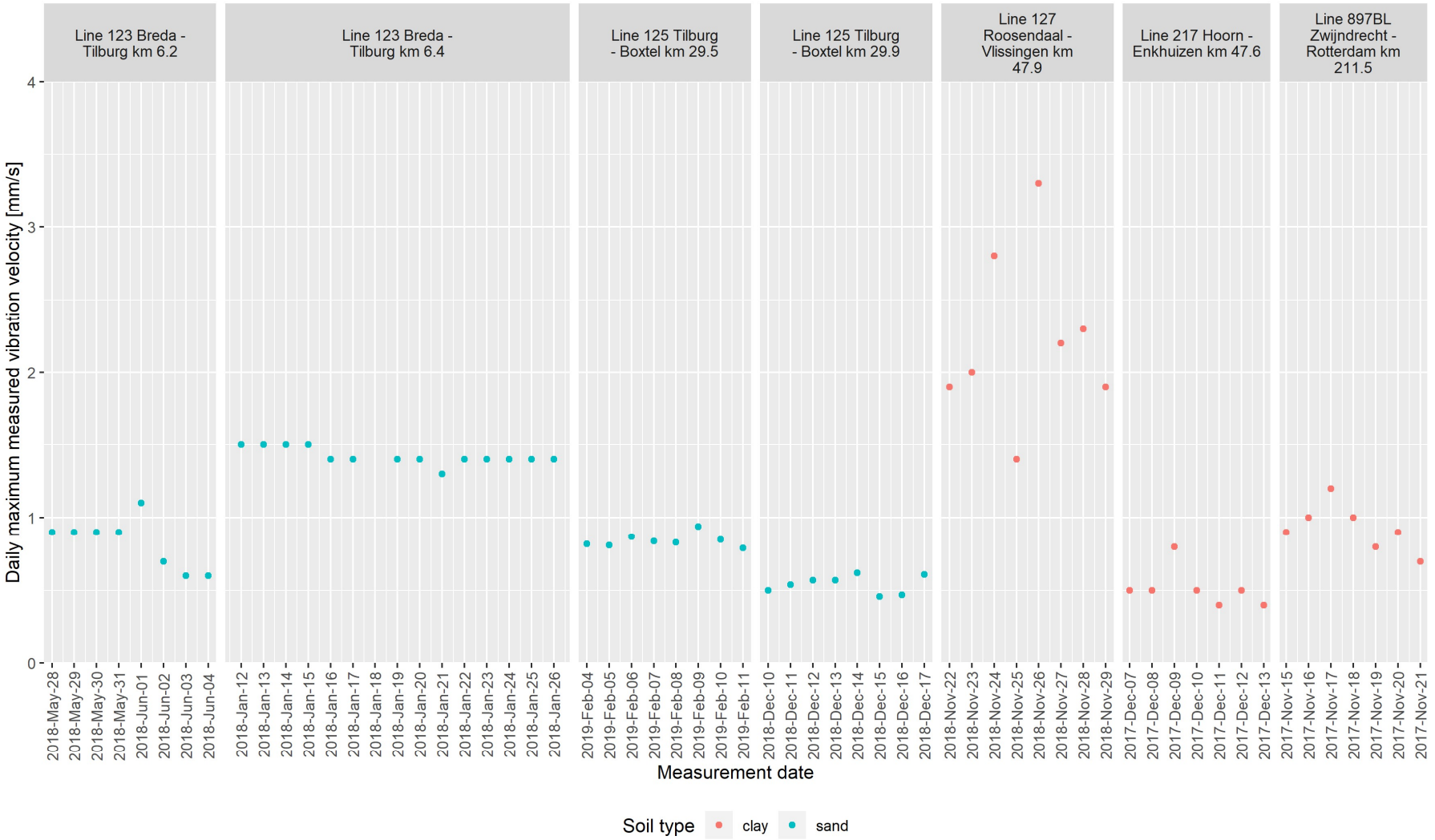


Figure 2-6: Daily maximum vibration velocities for seven locations. Data: N-D-SM.

2.5 N-P-SM: passage maxima

Another database that became available for this project is the vibration velocity maxima in three orthogonal directions from each train passage for three locations. The distances of the measuring locations to track 1 and 2 are summarized in Table 2-2. The measurements cover about 9 months and record about 50000 train passages. For simplicity and in line with the modelling decisions, in this report we focus on the daily maxima but it should be noted that the database may also serve as a starting point for later studies accounting for the number of train passages and train types, i.e. cargo or passenger.

Table 2-2: Summary of the distance of the three measurement locations from the two tracks in each location. Data: N-P-SM.

Measurment location's distance [m] to ↓	Dorst	Oisterwijk	Rijen
track 1	34	23	31
track 2	38	27	39

Pre-processing
See Table Glossary-0-2 for the used data source names and for the interpretation of terms (particularly the translation of the Dutch terms to English).

For each passage we calculated the maximum of the velocities in three orthogonal directions:

$$v_{\text{top,passage}} = \max(v_{\text{top},x,\text{passage}}, v_{\text{top},y,\text{passage}}, v_{\text{top},z,\text{passage}}) . \quad (2.1)$$

Then the maximum for each day is calculated by taking the maximum of all passages in a given day ($\mathbf{v}_{\text{top,passage}}$):

$$v_{\text{top,daily}} = \max(\mathbf{v}_{\text{top,passage}}) . \quad (2.2)$$

Descriptive statistics
A numerical overview of the N-P-SM database – focusing on the passage maxima – is presented in Table 2-3. It shows that the three locations are comparable in terms of average maximum velocities per passage although Dorst had a considerably lower (about half) recorded traffic frequency than the other two locations.

Visualization
In order to explore potential patterns in the measurements: the daily maxima ($v_{\text{top,daily}}$) are plotted over time in Figure 2-7 and the passage counts are plotted in Figure 2-8. No clear pattern is visible on either figure. There is no weekly periodicity, nor are the weekends clearly different from weekdays. The only visible pattern is the increase of traffic frequency and passage induced vibration velocity at the end of the summer months and beginning of fall. However, the periodicity of this pattern cannot be checked as only one summer is covered by the database.

Table 2-3: Summary of selected descriptive statistics of the passage maxima database per location. Data: N-P-SM.

	Dorst	Oisterwijk	Rijen
max($v_{top,passage}$) [mm/s]	2.02	1.78	2.20
min($v_{top,passage}$) [mm/s]	0.461	0.855	0.543
mean($v_{top,passage}$) [mm/s]	1.31	1.21	1.37
median($v_{top,passage}$) [mm/s]	1.29	1.19	1.40
mean(daily train passage count)	58	121	104
total number of train passages recorded	11691	22336	17706

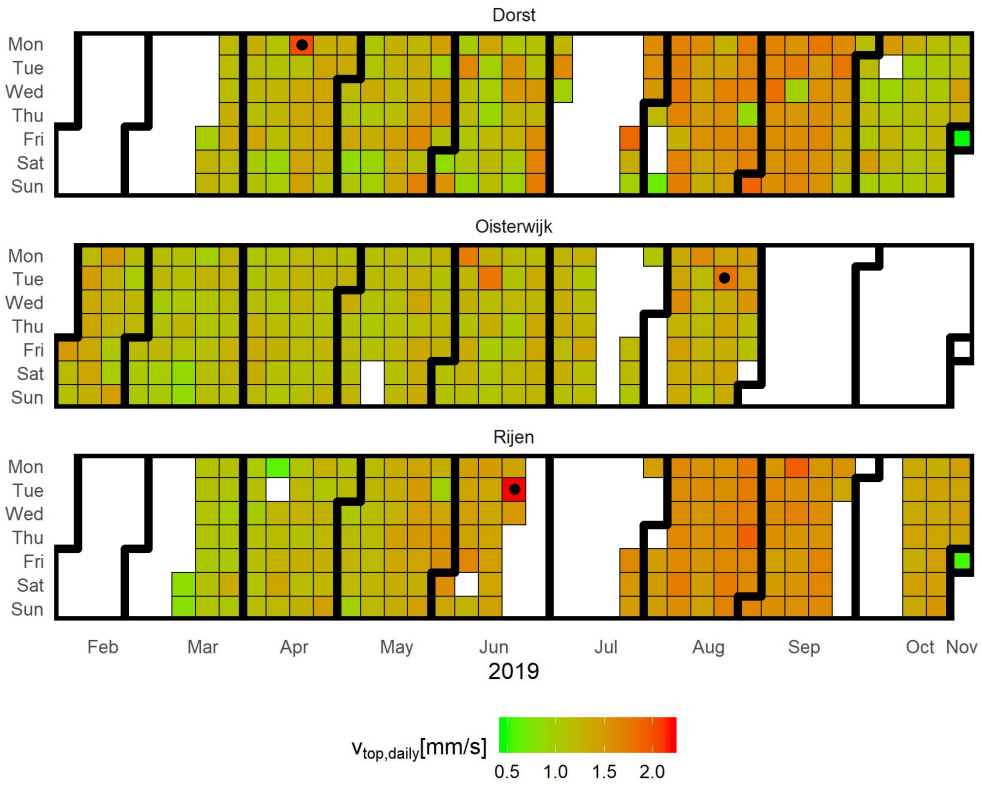


Figure 2-7: Overview of the daily vibration velocity maxima for the three locations with data available. The black circle indicates the largest value for each location. No fill indicates no data/measurement. Data: N-P-SM.

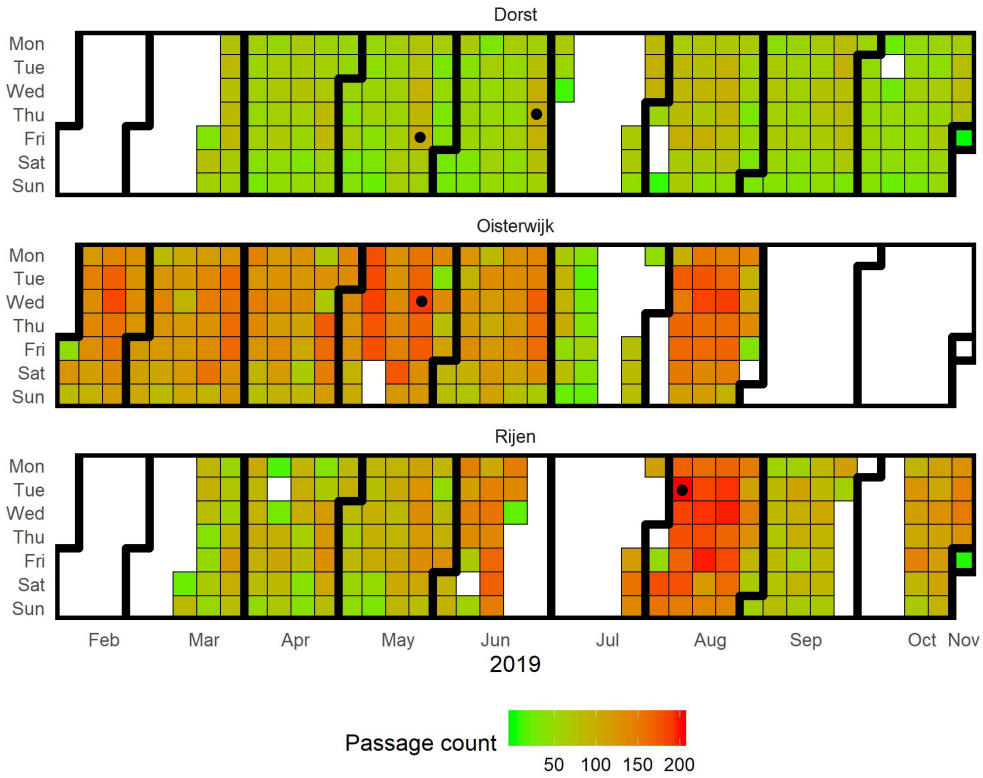


Figure 2-8: Overview of the daily train passage count for the three locations with data available. The black dot indicates the largest value for each location. No color indicates no data/measurement. Data: N-P-SM.

2.6 N-P-TS: passage time series

Ten train passages are available as quasi-continuous time series, their duration ranges from 10 to 40 seconds. For illustrative purposes an example passage with its three velocity time series is presented in Figure 2-9. It is visible that the maxima per measurement direction occur at different times. Due to the very few and very short time series this dataset is not used in later modelling steps. However, if more time series data became available they could be used for modelling. For example, peak over threshold models (Generalized Pareto distribution) could be fitted to the maxima and used for estimating distance thresholds.
The time series data also allows us to check an assumption in the old protocol's background document (Roos, 2004a) which assumes 100 peaks per train passage. Considering the N-P-TS data and assuming that the received time series contain only the passage time: the mean peak count is 421 from the ten passages and three directions for each passage. A more granular representation of the peak counts is shown in Figure 2-10. As expected, the measurement duration (assumed to be equivalent to passage duration) has a considerable effect on the number of peaks. Since counting peaks in time series measurements is ambiguous and domain specific, we provide a few sample plots of the outcome of our counting algorithm to illustrate what we considered as peaks and as not peaks, see Figure 2-11, Figure 2-12, and Figure 2-13. In comparison with the 100 peaks/passage used in Roos (2004a) the N-P-TS dataset has more than four times as many peaks.

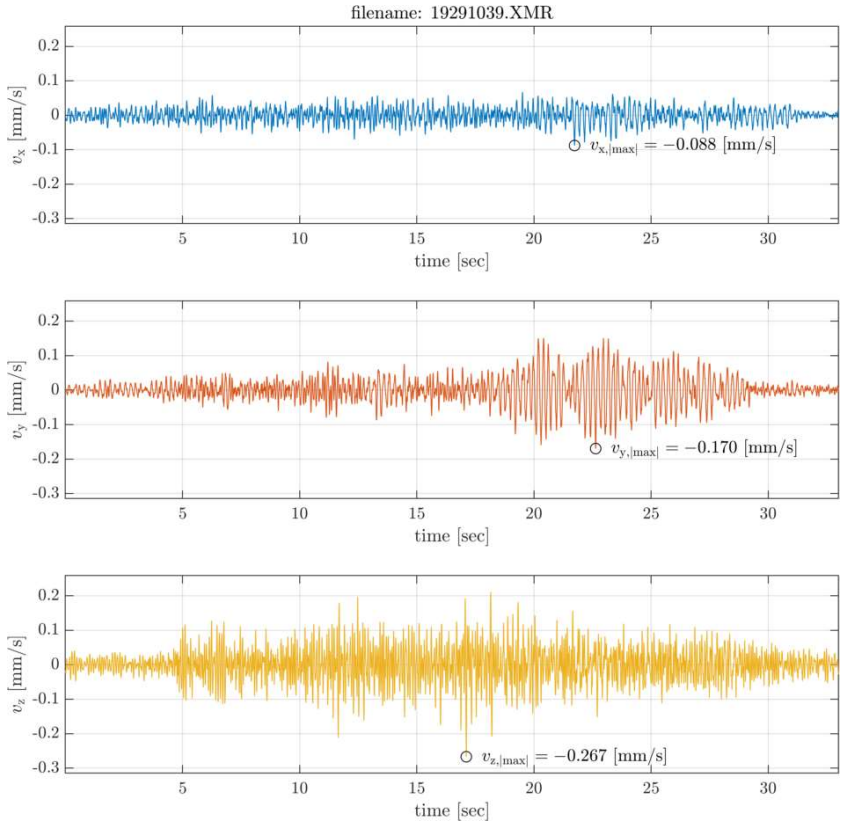


Figure 2-9: An example time series data from the 7 available passage time series. Data: N-P-TS.

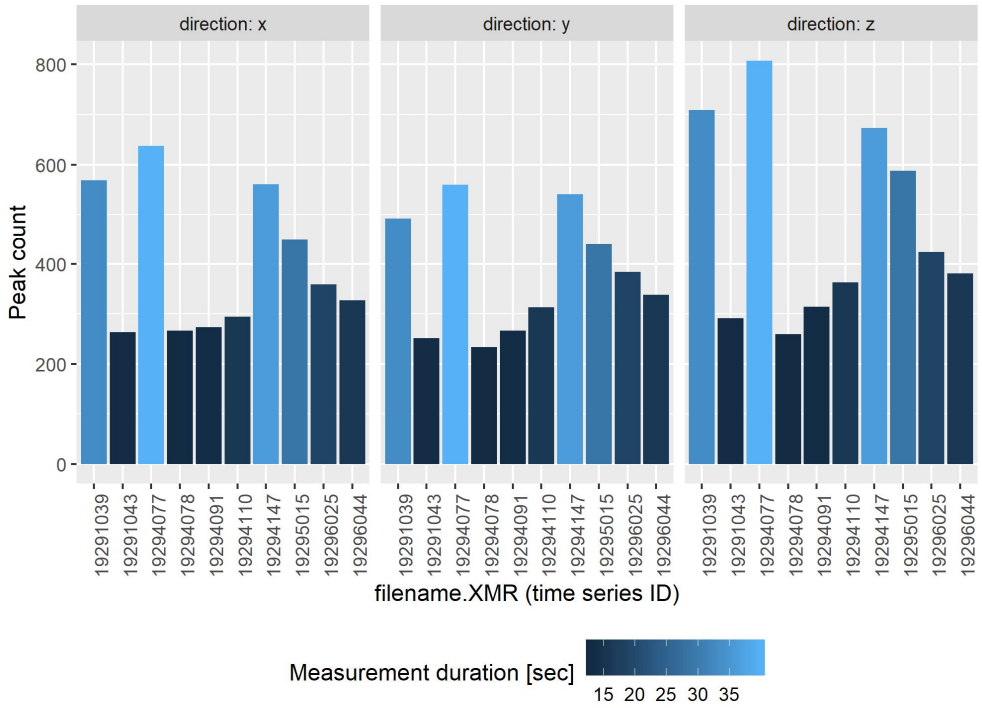


Figure 2-10: Peak counts per train passage. The peaks are counted in velocity time series in three orthogonal directions for 10 passages. The measurement duration corresponding to a

train passage is indicated with the color fill. See also Figure 2-11, Figure 2-12, and Figure 2-13 for illustrations of the identified and counted peaks. Data: N-P-TS.

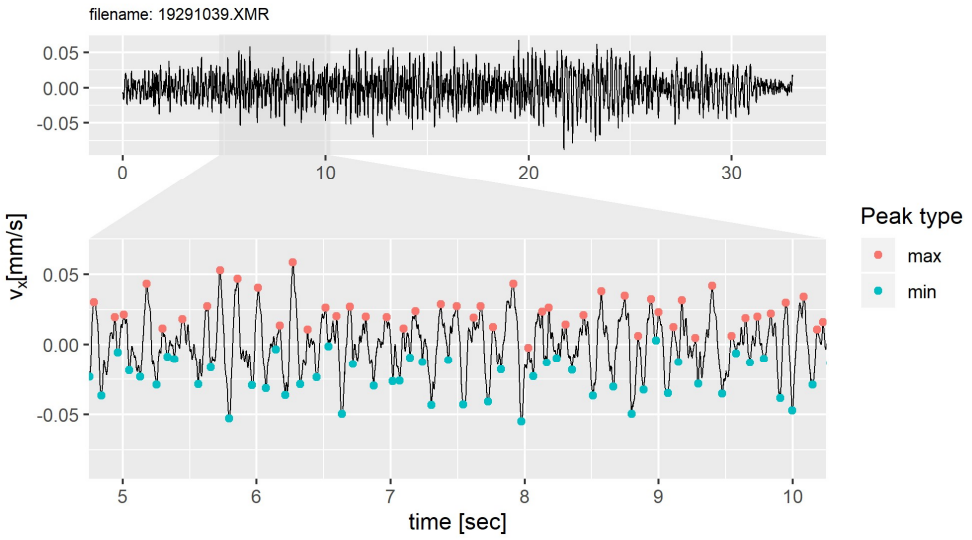


Figure 2-11: Illustration of identified peaks for time series 19291039.XMR in x direction. Data: N-P-TS.

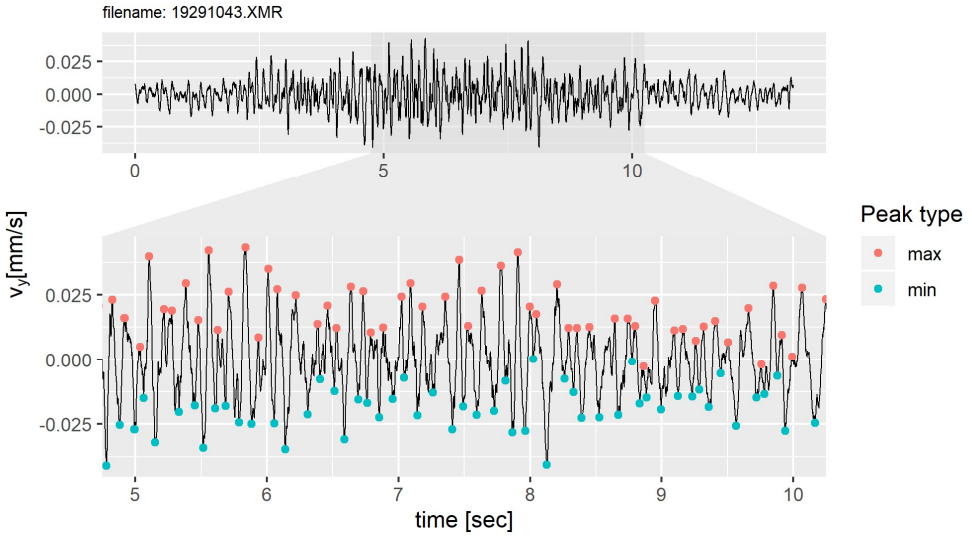


Figure 2-12: Illustration of identified peaks for time series 19291043.XMR in y direction. Data: N-P-TS.

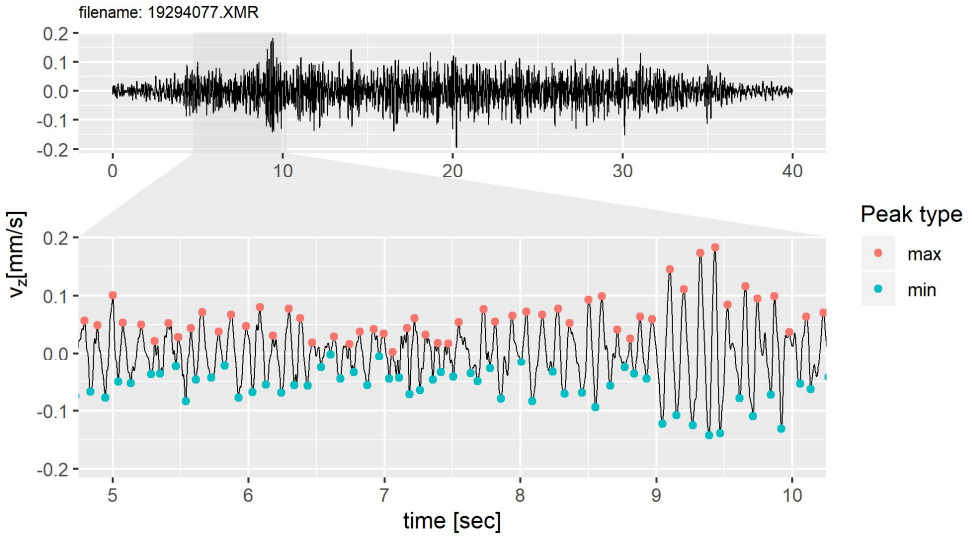


Figure 2-13: Illustration of identified peaks for time series 19294077.XMR in z direction. Data: N-P-TS.

2.7 Comparison to previous protocol

The current protocol (of 2004) differentiates between the following groups:

- soil type (gebied): soft (gebied I), stiff (gebied II);
- type of the railway track substructure: earthwork (aardenbaan), structure (kunstwerk);
- type of train: passenger, cargo;
- building categories: (cat 2, masonry in good state, cat 3, masonry in bad state or a monument); this defines the velocity threshold.

Differentiation means that distance threshold values are provided for all possible combinations, in total 16 (2x2x2x2).

For two reasons we cannot derive distance thresholds with such a resolution based only on the measured data:

i. The underlying assumptions for the calculation steps in (Roos, 2004a) are not always clear. We could not decipher how they are based on measurements and what measurements/post-processed values are used.
ii. The measurement data available to us in this study does not allow for such a differentiation as described in (Roos, 2004a). This will be further discussed when the results are described.

The measurement data available for this study poses some limitations:

- There is limited variability in the new databases:
  - N-M-SM and N-D-SM cover a wide range of track distances but there is no or not enough daily maxima to analyze the effect of vibration exposure period (time).
  - N-P-SM and N-P-TS are longer measurements, particularly N-P-SM can be used to analyze the effect of vibration exposure period (time); however data is only available for a few locations, which do not cover all situations in the protocol (e.g. soil type, distance, presence of discontinuities).

- There is a lack of information in respect of the old data, e.g. train types, measurement duration. The missing measurement duration does not allow to include the old data in our models that account for vibration exposure period (see section 3.2).

3 Methods and tools

This section outlines the tools, methods used during the revision.

3.1 Overview

A schematic overview of the general approach taken in this report is shown in Figure 1-1. It builds largely on the previous revision of the protocol (Roos, 2004a) but extends it in a few ways which are explained in the coming sections.

The main components of the approach:

- Peak (top) velocity attenuation curve (solid dark blue line).
- Uncertainty quantification of the attenuation curve (solid light blue).
- Peak velocity threshold (v_{th}).
- Target exceedance probability of peak velocity threshold ($P_{th} = P(v_{top} > v_{th})$).

These main components are the same as in the previous revision of the protocol but certain modelling decisions are different, the reasons are explained in later sections. An extension of the previous revision is the explicit consideration of the vibration exposure period. The velocities over time are modelled as a stochastic process. Due to this the longer the exposure period, the more likely the higher vibration velocities are.

The approach follows the following steps:

1. Develop a peak velocity attenuation function: fit a mathematical model to the measurements. The model accounts for:
    a. uncertainties (aspects/phenomena which are not explicitly included in the model);
    b. the differences in measurement duration: bring all measurements to a common denominator: vibration exposure period.
2. Establish and/or select the peak velocity threshold (v_{th}) and target exceedance probability of peak velocity threshold (P_{th}).
3. Using the model from 1. and the threshold from 2. find the threshold distance (R_{th}) that corresponds to P_{th} exceedance probability.

Note that due to the explicit consideration of the measurement duration and converting to a single common vibration exposure period the threshold distance is also conditioned on the exposure period, i.e. time dependent.

The subsequent sections detail these steps and the corresponding modelling decisions.

3.2 Peak velocity attenuation

The general structure of the model is presented in Figure 3-1 and can be written as:

$$v_{top,t} = f(R) + E\,, \tag{3.1}$$

where

$v_{\text{top},t}$ peak velocity at distance R, subscript t emphasizes the dependence on time;
R distance from vibration source;
E a random variable that represents uncertainty.

Since most of our currently available data is in a format of a maximum per measurement block it seems natural to use extreme value theory to describe the distribution of maxima (Coles, 2001). For convenience, the Gumbel distribution (2-parameter distribution, easy to change block size, see Annex C) is selected for modelling the uncertainty in extremes (later Normal and generalized extreme value distributions are also checked). However, the Gumbel distribution's domain covers all real numbers which could lead to negative model predictions for small peak velocities. To avoid this Eq.(3.1) is reformulated as:

$$\log(v_{\text{top},t}) = \log(f(R)) + E \tag{3.2}$$

then the prediction of the peak velocity: $v_{\text{top},t} = \exp\left[\log(f(R)) + E\right]$. Hence E can be defined over all real values while $v_{\text{top},t}$ is still constrained to non-negative values. This Eq.(3.2) formulation is used from here on.

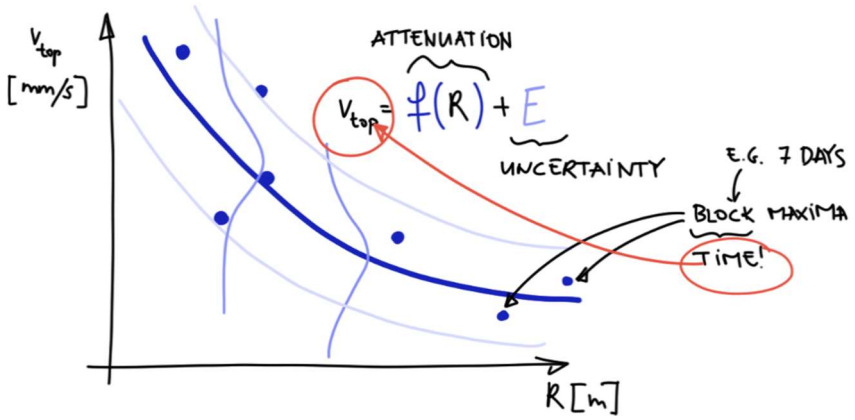


Figure 3-1: A schematic overview of the peak velocity attenuation curve.

3.2.1 *The base function (model)*

In line with Roos (2004a), the Barkan formula is used to describe the attenuation:

$$v_R = v_0 \cdot \left(\frac{R_0}{R}\right)^n \cdot e^{-\alpha\cdot(R-R_0)}, \tag{3.3}$$

where

v_R peak velocity at distance R;
R distance from vibration source;
v_0 peak velocity at reference distance R_0;
R_0 reference distance from vibration source (fixed at value 20 m);
α model parameter, can be interpreted as the damping coefficient;
n exponent, 0.5;

Given the structure of the formula (pair-wise formulation), the reference point location (R_0) can be arbitrary. R_0 is selected to be 20 m. Note that this decision has no impact on the fitted model or the derived distance threshold but only on the parametrization of the attenuation base function. The free parameters in Eq.(3.3) that are inferred during model fitting: $[v_0, \alpha]$. Both parameters' support is constrained to non-negative numbers.

3.2.2 The uncertainty model

E is assumed to be Gumbel distributed in Eq.(3.2). This means that the log of peak velocities are modelled as Gumbel distributed. This is possible because log(.) is a monotonic function. To investigate the sensitivity of the results to this decision generalized extreme value (GEV) and normal distributions are also used instead of Gumbel (Coles, 2001). GEV encompasses Gumbel as a special model so it can be used to test if Gumbel is the right choice among the distribution families in GEV. The normal distribution is used as a reference.

The following additional modelling decisions are made:

- *E* is independent of *R*.
- The mode of *E* is zero, i.e. the mode of $\log(v_{\text{top},t})$ is $\log(f(R))$.

This means that the uncertainty model can be described by a single parameter, we have chosen the standard deviation of the Gumbel (GEV, normal) distribution (σ_G) for this purpose. This also means that the total number of inferred parameters is three: $[v_0, \alpha, \sigma_G]$.

3.2.3 Time dependence

The impact of vibration exposure time is accounted for by assuming independent block maxima, i.e. independent maximum velocities per day. The independence assumption is supported by the analyses presented in section 2.5; however, it is not decisive as that dataset is not complete (unknown number of missing passages at unknown times). Moreover, the block maxima are assumed to follow the same distribution. This impact of exposure period is accounted for in model fitting and for prediction as well. This is necessary for the former as the measurement data are coming from different measurement durations. The mathematical details of accounting for vibration exposure period are summarized in Annex C.

3.3 Peak velocity threshold

The peak velocity threshold (v_{th}) is adopted from (SBRCURnet, 2017). Moreover, for rail traffic induced vibrations, usually a measurement period of 7 days (or more) is chosen. Therefore this vibration exposure period was selected as a common basis for all subsequent analyses. Unless otherwise stated if the vibration exposure period or vibration measurement duration is not mentioned it is 7 days. The threshold value was chosen, based on masonry buildings. A value for the vibration limit of 5 mm/s is the initial value based on the values given in SBR-guideline A for dominant frequencies lower than 10 Hz. For higher frequencies this limit value is higher, so 5 mm/s in that case is a safe choice (Figure B-2 shows that the majority of measurements falls into the <10 Hz category). A train passage is assumed to be long enough in time to be classified as a continuous vibration. The vibration limit is therefore divided by a safety factor 2.5 as specified in SBR-guideline A. This leaves:

$v_{th} = 2$ mm/s.

For masonry which is classified as sensitive to vibrations, or if the masonry belongs to a monumental building, a safety factor of 1.7 is applied to the vibration limit, giving a threshold of:

$v_{th} = 1.2$ mm/s.

If the vibration values are below this value, both vibration sensitive buildings, monuments and non-vibration sensitive buildings are sufficiently safe against damage.

Please note that SBR-guideline A has no specific provisions related to the thresholds to account for the vibration exposure period or the number of peaks occurring during a certain lifetime of the structure. For railway traffic, SBR guideline A specifies a minimum measurement duration of 1 week (preferably 2 weeks). SBR Guideline A also provides a statistical method for analyzing data. However, this does not include a procedure how to account for the 'design life' for which the vibration levels are analyzed.

3.4 Target probability

The value and the derivation of the target probability of the peak velocity threshold exceedance is identical of that of the previous revision of the protocol (Roos, 2004a). For convenience, the assumptions and steps are recapitulated here:

- Target probability of vibration induced damage over the vibration exposure period is adopted from SBR Guideline A: $P_d = 1$ %.
- We assume that the vibration as an action is the non-dominant component in the damage reliability ($\alpha_{vibration} = 0.4$, alpha is the sensitivity factor). For non-dominant components in reliability analysis and standardization this is a widely accepted and recommended value, see for example (CEN, 2002).

From these two decisions the target probability of the peak velocity threshold exceedance is:

$$P_{th} = \Phi\left(0.4 \cdot \Phi^{-1}(0.01)\right) = 0.176 \,. \tag{3.4}$$

Where $\Phi(.)$ is the standard normal cumulative distribution.

3.5 Additional consideration and decisions

- Measurements farther than 100 m from the track were not considered. In the available data, this concerns a single measurement 200 m from the track.
- Parameter x_0 in Eq.(3.3) is set to 20 m (it does not have an effect on the final fitted model but on the model parameters)
- The α parameter in Eq.(3.3) is constrained to non-negative values. The motivation for this is twofold:
  - α can be interpreted as damping which is not negative.

  - For certain cases the mode of the fitted attenuation model was curving backward (not monotonic decreasing).

  However, due to this decision sometimes the maximum likelihood estimate is at the lower boundary of α.
- Train load frequency per track is not taken into account.
- Train type (cargo, passenger) composition is not taken into account.
- Frequentist approach is adopted (deemed to be sufficient given the accuracy the subject matter allows)
  - Maximum likelihood method for parameter estimation.
  - Parameter estimation uncertainty is not considered.
- The impact of model selection decisions is not explored.

3.6 Multiplication factors

Since the available data is not covering all relevant situations we cannot fully rely on it to derive threshold distances for all the categories that are present in the old protocols. This situation is handled by using multiplication factors applied to v_{top} values. The multiplication factors are taken from Roos (2004a) that can be considered as approximations of the effect of soil type, substructure type, etc. References to the sources of the values in these sources are given in Roos (2004a). These factors have not been further analyzed.

The factors applied are:

1. The first factor, applied to all data is the factor 1.6 for an indicative measurement. For all data in our database, only the measurements at 1 stiff point at the foundation was present.
2. For buildings with a height of 12 meters or more, the data is additionally multiplied by 0.6.
3. If the track runs over a civil structure (concrete instead of earthwork), a factor 0.5 is applied for soft soils and a factor 0.7 for stiff soils.

These factors are combined, if needed, so for a building with H > 12 m, and a train running over a civil structure in Gebied I (soft soil), a total factor of 1.6 x 0.6 x 0.5 (resulting in approximately 0.5) is applied.

The factors are summarized in Table 3-1 and the procedure goes as follows:

1. Select the multiplication factor from Table 3-1 that fits the considered situation in terms of soil, track substructure, building height H, and v_{th}.
2. Select a subset of the v_{top} dataset based on the presence of discontinuity: yes, not, yes or no.
3. Multiply each v_{top} selected in 2. with the multiplication factor from 1.
4. Proceed with as described in section 3.1-3.5.

Table 3-1: Multiplication factors applied to v_{top} values for fitting a model and deriving threshold distances. The same factors as presented here apply for locations with, without, and with or without discontinuities.

Gebied (soil)	Spoorbaan (track substructure)	Metselwerk, niet trillingsgevoelig v_{th} = 2.0 mm/s		Metselwerk, trillingsgevoelig v_{th} = 1.2 mm/s	
		H<12m	H≥12m	H<12m	H≥12m

1 (soft)	Aardebaan (earth/soil)	1.6	1.6×0.6	1.6	1.6×0.6
1 (soft)	Kunstwerk (structure)	1.6×0.5	1.6×0.5×0.6	1.6×0.5	1.6×0.5×0.6
2 (stiff)	Aardebaan (earth/soil)	1.6	1.6×0.6	1.6	1.6×0.6
2 (stiff)	Kunstwerk (structure)	1.6×0.7	1.6×0.7×0.6	1.6×0.7	1.6×0.7×0.6

*H: building height.

In the protocol, two steps in the assessment are defined. In each step a table is used. The first step provides the outer distances outside which damage is unlikely, and includes data from all sources. In the second step, a shorter distance is defined for the situations that discontinuities are not present in the direct vicinity of the building. In the former protocol, an additional multiplication factor was defined, to be applied to the factors in Table 3-1, thus giving a table with shorter distances. This factor is 0.7 for soft soil and 0.6 for stiff soil.
In this report, sufficient data for the situations where a discontinuity was present, and for situations there was not, were available. The enabled us to directly find the threshold distances based on the data and to compare that result with the one obtained by applying the additional multiplication factors.

4 Results

This section provides the details of the proposed statistics-based simulation approach. The approach is used to generated synthetic data that is used for testing some selected similarity measures.

4.1 Overview

The model described in the preceding sections is fitted to different subsets of the dataset that consists a single maximum for each measurement site (data: N-M-SM), i.e. the maximum over the entire measurement period. In order to obtain threshold distances various multiplication factors (Section 3.6) are applied to the measured maxima and the statistical model is fitted to this scaled data.
Section 4.2 presents the results for 7-day vibration exposure period, it is accompanied by additional analysis that are only briefly reported in this section and documented in more detail in Annex D, E, F, and G. The section is concluded with a discussion (section 4.3).

4.2 Threshold distances for a 7-day exposure period

The results in terms of threshold distances for a 7-day exposure period along with those of the current protocol (2004) are also summarized Table 4-1, Table 4-2, Table 4-3, and Table 4-4. For convenience the same values are also plotted in Figure 4-1, Figure 4-2, Figure 4-3, and Figure 4-4. Moreover, for illustration some of the fitted models are displayed in Figure 4-5 and Figure 4-6. These figures show the curves corresponding to the 18% target probability and the measurement data. For the mode, mean, and prediction uncertainty estimate of the fitted models see Annex G and Figure G-1, Figure G-2, and Figure G-3; these figures also allow for a visual assessment of the goodness of the fit (for all the figures see the extended digital version of Annex G provided separately to this report).

In Figure 4-1, Figure 4-2, Figure 4-3, and Figure 4-4 a comparison is presented for the distances obtained for the three situations (1) only data with discontinuities, (2) all data and (3) only data without discontinuities. Logically, there should be a decrease in distance from 1 to 3. This is observed for the cases stiff soil and the cases with a earthwork substructure. For the combination soft soil and civil structure as substructure, the opposite trend was observed. In paragraph 4.3 this will be discussed further.

Table 4-1: Summary of all threshold distances derived in this report (data: N-M-SM, with or without discontinuities) and their comparison to the previous protocol. The values for TNO 2004 are for cargo and passenger trains (in this order: cargo/passenger).

			TNO 2020 – this report				TNO 2004 (Roos, 2004a)(cross ref. Table 7 and 9)			
	Track	v_{th}=	2 mm/s		1.2 mm/s		2 mm/s		1.2 mm/s	
Soil	substructure		*H*<12m	*H*>12m	*H*<12m	*H*>12m	*H*<12m	*H*>12m	*H*<12m	*H*>12m
Data: with or without discontinuities (as in Table 5-1)										
Area 1										
soft	earthwork		34.4	16.4	61.9	34.4	30/20	15/10	60/40	35/20
soft	structure		12.2	4.91	26.9	12.2	20/10	6/6	25/20	10/6
Area 2										
stiff	earthwork		24.3	8.73	67.4	24.3	20/10	6/6	30/20	15/10
stiff	structure		11.9	4.28	33.0	11.9	10/6	6/6	20/15	10/6

Table 4-2: Summary of all threshold distances derived in this report (data: N-M-SM, with discontinuities) and their comparison to the previous protocol. The values for TNO 2004 are for cargo and passenger trains (in this order: cargo/passenger).

			TNO 2020 – this report				TNO 2004 (Roos, 2004a)(cross ref. Table 7 and 9)			
	Track	v_{th}=	2 mm/s		1.2 mm/s		2 mm/s		1.2 mm/s	
Soil	substructure		*H*<12m	*H*>12m	*H*<12m	*H*>12m	*H*<12m	*H*>12m	*H*<12m	*H*>12m
Data: with discontinuities (as in Table 5-2)										
Area 1										
soft	earthwork		35.5	13.3	90.4	35.5	NA	NA	NA	NA
soft	structure		9.27	3.37	25.1	9.27	NA	NA	NA	NA
Area 2										
stiff	earthwork		27.7	9.98	77.0	27.7	NA	NA	NA	NA
stiff	structure		13.6	4.89	37.7	13.6	NA	NA	NA	NA

NA: not available.

Table 4-3: Summary of all threshold distances derived in this report (data: N-M-SM, without discontinuities) and their comparison to the previous protocol. The values for TNO 2004 are for cargo and passenger trains (in this order: cargo/passenger).

			TNO 2020 – this report				TNO 2004 (Roos, 2004a)(cross ref. Table 7 and 9)			
	Track	v_{th}=	2 mm/s		1.2 mm/s		2 mm/s		1.2 mm/s	
Soil	substructure		*H*<12m	*H*>12m	*H*<12m	*H*>12m	*H*<12m	*H*>12m	*H*<12m	*H*>12m
Data: without discontinuities (as in Table 5-3)										
Area 1										
soft	earthwork		34.0	21.6	48.6	34.0	NA	NA	NA	NA
soft	structure		17.8	9.39	29.3	17.8	NA	NA	NA	NA
Area 2										
stiff	earthwork		21.3	8.38	49.1	21.3	NA	NA	NA	NA
stiff	structure		11.2	4.22	27.8	11.2	NA	NA	NA	NA

NA: not available.

Table 4-4: Summary of all threshold distances derived in this report (data: N-M-SM, without discontinuities (with factors)) and their comparison to the previous protocol. The values for TNO 2004 are for cargo and passenger trains (in this order: cargo/passenger).

			TNO 2020 – this report				TNO 2004 (Roos, 2004a)(cross ref. Table 7 and 9)			
	Track	v_{th}=	2 mm/s		1.2 mm/s		2 mm/s		1.2 mm/s	
Soil	substructure		*H*<12m	*H*>12m	*H*<12m	*H*>12m	*H*<12m	*H*>12m	*H*<12m	*H*>12m
Data: without discontinuities (with factors) (as in Table 5-4)										
Area 1										
soft	earthwork		20.8	9.03	41.7	20.8	20/15	10/6	40/25	20/10
soft	structure		6.52	2.50	15.7	6.52	10/6	6/6	15/10	6/6
Area 2										
stiff	earthwork		8.73	3.14	24.3	8.73	10/6	6/6	20/10	6/6
stiff	structure		4.28	1.54	11.9	4.28	6/6	6/6	10/6	6/6

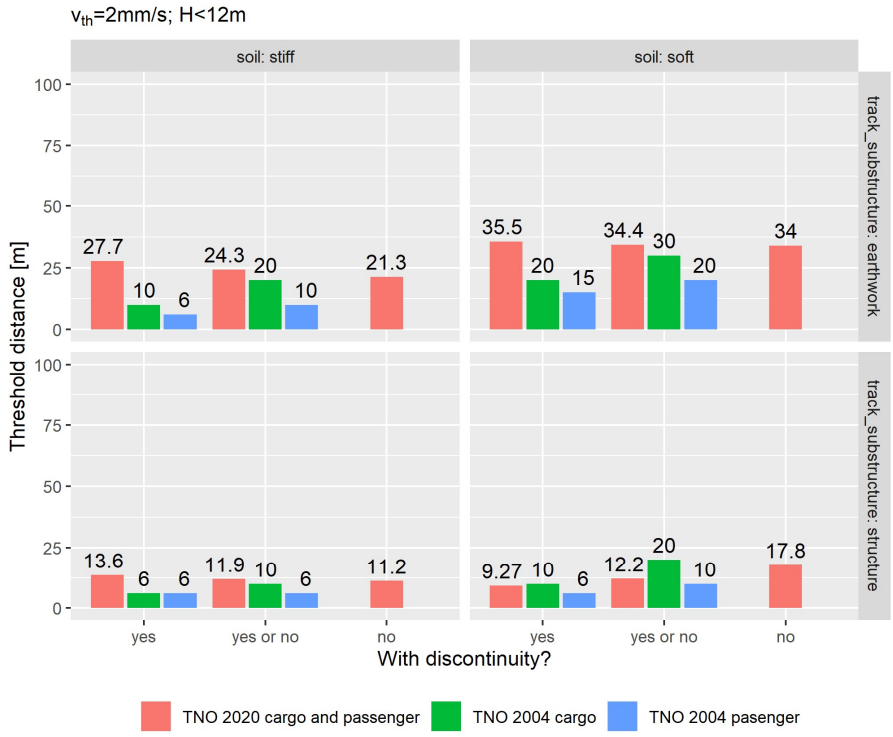


Figure 4-1: Visual overview of the threshold distances derived in this report and those of the previous protocol (Roos, 2004a) for v_{th}=2 mm/s and H<12 m. Data: N-M-SM.

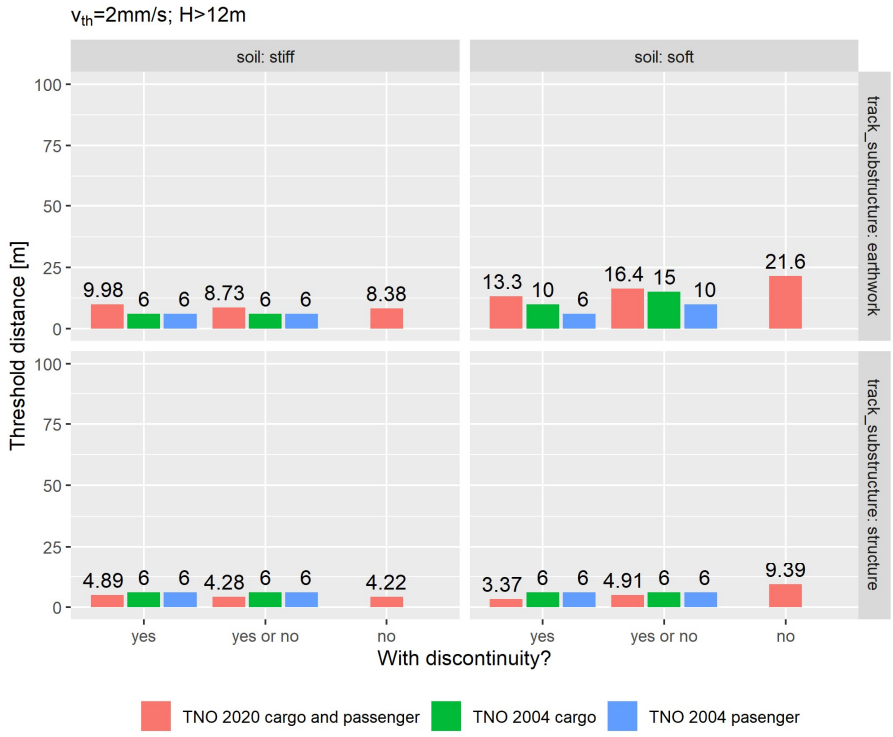


Figure 4-2: Visual overview of the threshold distances derived in this report and those of the previous protocol (Roos, 2004a) for v_{th}=2 mm/s and H>12 m. Data: N-M-SM.

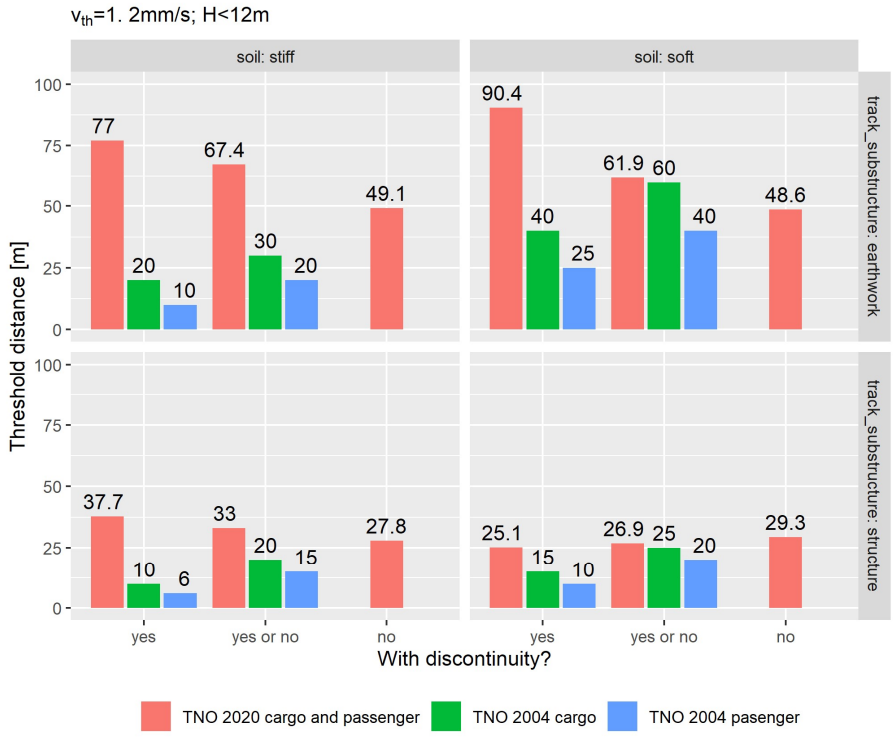


Figure 4-3: Visual overview of the threshold distances derived in this report and those of the previous protocol (Roos, 2004a) for v_{th}=1.2 mm/s and H<12 m. Data: N-M-SM.

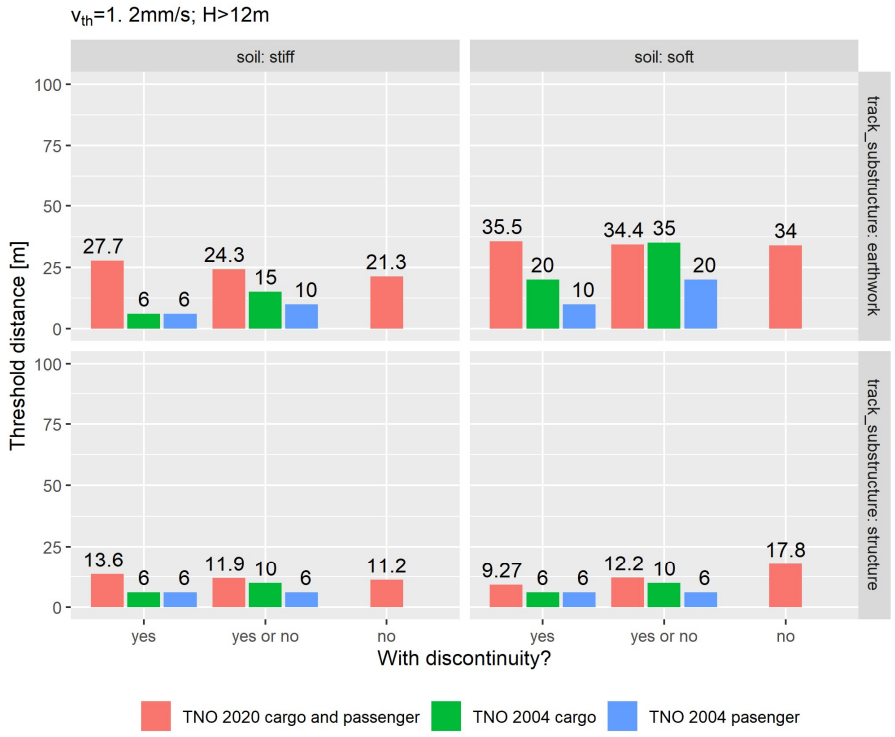


Figure 4-4: Visual overview of the threshold distances derived in this report and those of the previous protocol (Roos, 2004a) for v_{th}=1.2 mm/s and H>12 m. Data: N-M-SM.

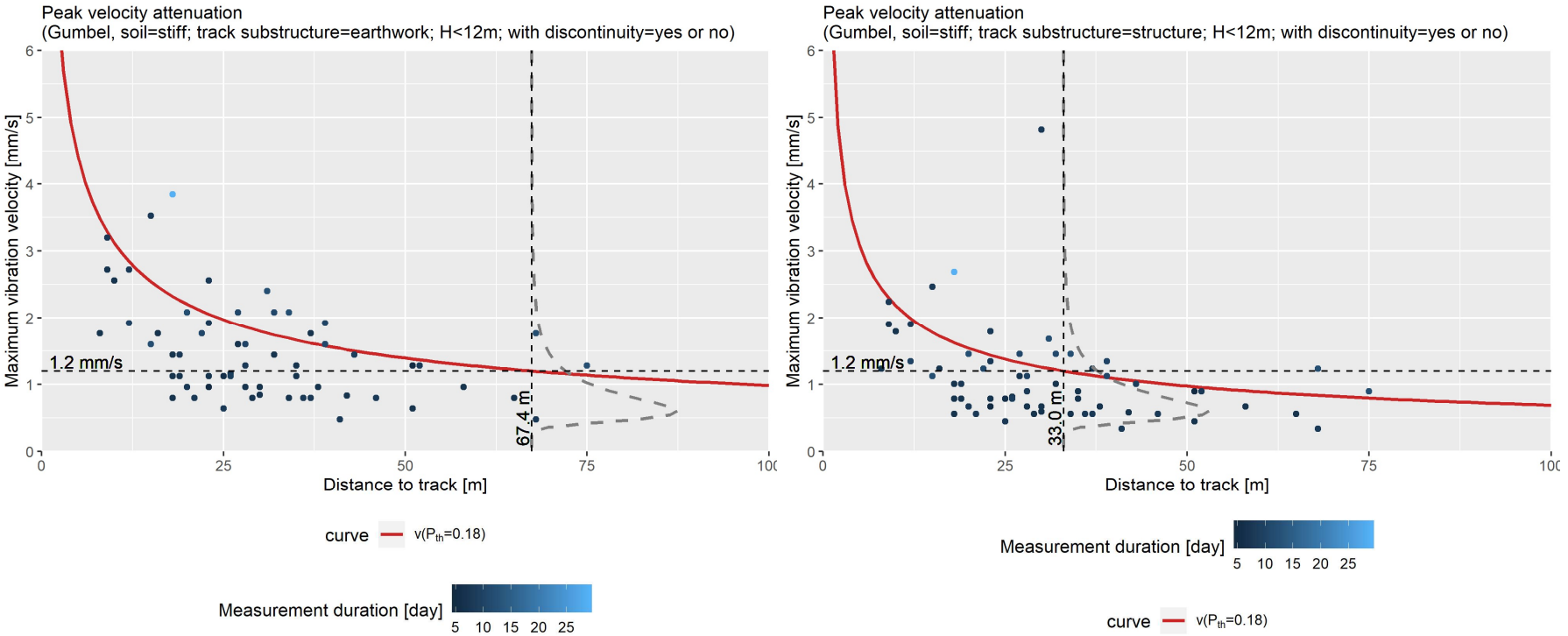


Figure 4-5: Fitted 7-day peak velocity attenuation curves and measurements for stiff soils. Solid brown line: quantile of the peak velocity corresponding to 18% non-exceedance probability; dashed gray line: probability density of the model prediction at the threshold distance; gray dashed lines: threshold peak velocity (v_{th}) and calculated threshold distance (R_{th}). Mind the inconsistency in the time duration of the attenuation model and the measured maxima which can distort the visual comparison. Data: N-M-SM.

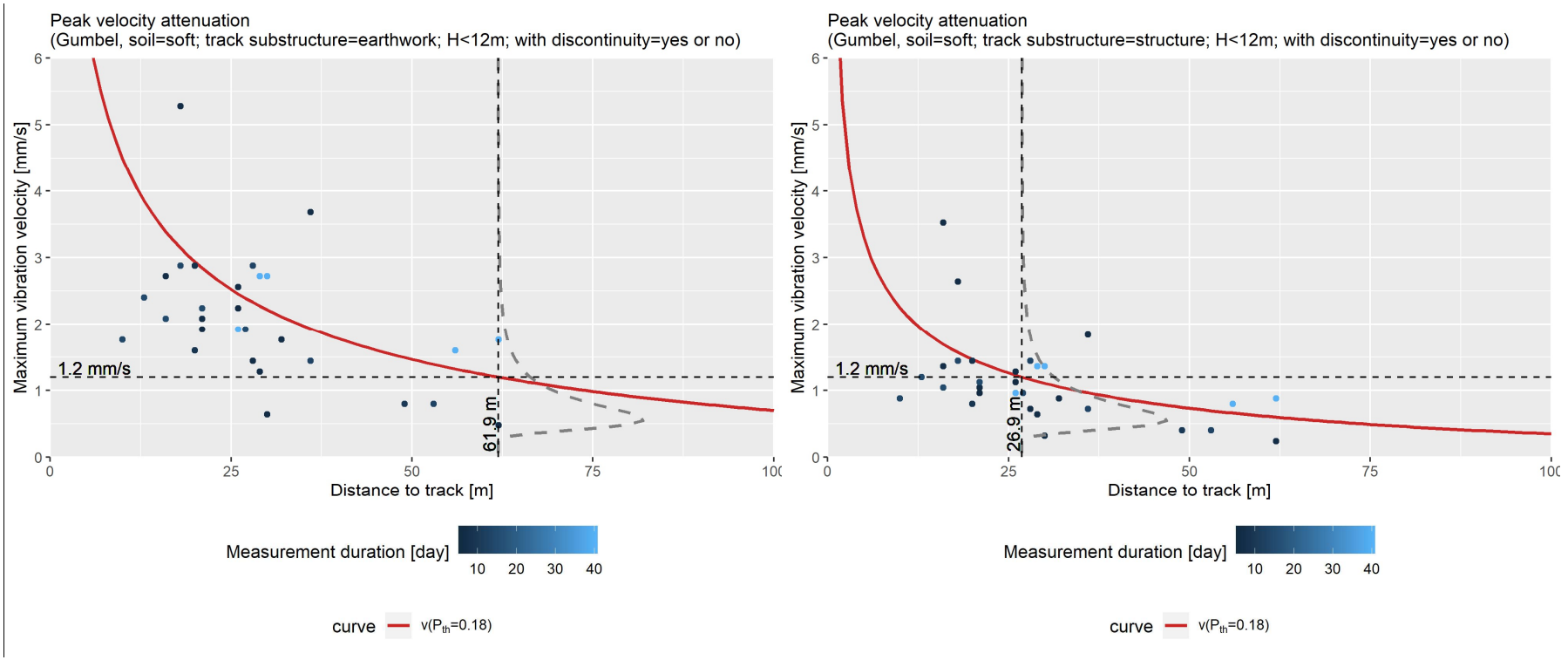


Figure 4-6: Fitted 7-day peak velocity attenuation curves and measurements for soft soils. Solid brown line: quantile of the peak velocity corresponding to 18% non-exceedance probability; dashed gray line: probability density of the model prediction at the threshold distance; gray dashed lines: threshold peak velocity (v_{th}) and calculated threshold distance (R_{th}). Mind the inconsistency in the time duration of the attenuation model and the measured maxima which can distort the visual comparison. Data: N-M-SM.

4.3 Analysis of assumptions

We analyzed the impact of several modelling assumptions and summarize here only the conclusions, the details can be found in relevant annexes.

4.3.1 Distribution type

The impact of the distribution type (statistical distribution: probability density function) of *E* in Eq.(3.2) is analyzed considering Gumbel, normal, and generalized extreme value (GEV) distributions. It is found the distribution type has a small influence on the 7-day threshold distance (for the details of the analysis see Annex D.2). This is attributed to the large target probability (18%), at which level the difference between distributions is small. The distributions are fitted to the same data hence their first few moments are expected to be similar. larger differences would be expected farther at the tails. The differences due to the distribution type start to manifest when other than 7-day exposure periods are considered (see Annex E.1).

4.3.2 Exponent in the Barkan formula

Two exponents (*n*) in the Barkan formula (Eq.(3.3)) are compared: 0.5 and 1.0. It is found that the small (<10m) and large (>50m) threshold distances are considerably affected by the choice of the exponent. The change is opposite for the small and large distances. Comparing the goodness of the fit of the two models demonstrate that the model with n=0.5 provides substantially better fit than n=1.0. Based on the data, among the considered two models the n=0.5 one is clearly better. The details of the analysis can be found in Annex D.1.

4.3.3 Vibration exposure period

The analysis of considered other than 7-day vibration exposure period indicates that the exposure duration can have a substantial impact on the threshold values (Annex E and F). However, further analysis is needed, which is outside the scope of this study, to explicitly model between site and within site variability. Only the latter would contribute to the shift of the attenuation curve over time.

4.4 Discussion

The results presented in section 4.2 form the basis for an update of the track distances for the new ProRail Protocol. There are some limitations to the results, and therefore some of these results need some consideration. These are discussed below. For the protocol, the distances have also been rounded off. These are discussed further in chapter 5.

One of the limitations of the tables is that there is not sufficient data to directly derive all the distances from the data alone. Factors have been applied, which have been chosen for the 2004-protocol, and have not been changed here.
The only exception is the difference between the situations where we have all data included versus the situation where we have data without nearby discontinuities. From our dataset, these situations have been analyzed separately. This analysis enables us to compare it with the results obtained when the factors from the 2004 protocol have been applied. This leads to considerable differences in results, see Table 4-5.

Table 4-5: Difference in threshold distances obtained using two approaches: model 1 – model 2. Model 1 (Table 4-3) using only data without discontinuities and the factors in Table 3-1. Model 2 (Table 4-4) using all data (with and without discontinuities) and the factors in Table 3-1 and an additional 0.7 and 0.6 multiplication factor for soft and stiff soils respectively. Data: N-M-SM.

			TNO 2020: Table 4-3 - Table 4-4			
	Track	v_{th}=	2 mm/s		1.2 mm/s	
Soil	substructure		*H*<12m	*H*>12m	*H*<12m	*H*>12m
Area 1						
soft	earthwork		13.2	12.6	6.90	13.2
soft	structure		11.3	6.89	13.6	11.3
Area 2						
stiff	earthwork		12.6	5.24	24.8	12.6
stiff	structure		6.92	2.68	15.9	6.92

There is a number of possible reasons which may cause the differences in these tables, but these could not be checked during the current research:

- The input in the database was to put either *yes* or *no* for a discontinuity nearby. There is no other information available. It was assumed that a minimum distance of 75 meters was used, but this was not made explicit in the Excel file.
- The values for the factors applied from the 2004- protocol have not been clearly motivated in the background report.
- The number of data analyses for this category is a subset of all data available. The final result might therefore be very sensitive to single outliers. To check this, this requires a larger amount of data for this subset.

5 Summary and recommendations

This final section provides a summary and it recapitulates the main conclusions along with practical recommendations.

5.1 Summary

The threshold distances determined in this project and given earlier in Table 4-1 are summarized numerically in Table 5-1, Table 5-2, Table 5-3, and Table 5-4.

Table 5-1: Threshold distances for various groups and threshold velocities. Data: N-M-SM, with or without discontinuities.

		[mm/s]	TNO 2020 – this report			
	Track	v_{th}=	2		1.2	
Soil	substructure		*H*<12m	*H*>12m	*H*<12m	*H*>12m
Area 1						
soft	earthwork		34.4	16.4	61.9	34.4
soft	structure		12.2	4.91	26.9	12.2
Area 2						
stiff	earthwork		24.3	8.73	67.4	24.3
stiff	structure		11.9	4.28	33.0	11.9

Table 5-2: Threshold distances for various groups and threshold velocities. Data: N-M-SM, only locations with discontinuities.

			TNO 2020 – this report			
	Track	v_{th}=	2 mm/s		1.2 mm/s	
Soil	substructure		*H*<12m	*H*>12m	*H*<12m	*H*>12m
Area I						
soft	earthwork		35.5	13.3	90.4	35.5
soft	structure		9.27	3.37	25.1	9.27
Area 2						
stiff	earthwork		27.7	9.98	77.0	27.7
stiff	structure		13.6	4.89	37.7	13.6

Table 5-3: Threshold distances for various groups and threshold velocities. Data: N-M-SM, only locations without discontinuities.

			TNO 2020 – this report			
	Track	v_{th}=	2 mm/s		1.2 mm/s	
Soil	substructure		*H*<12m	*H*>12m	*H*<12m	*H*>12m
Area I						
soft	earthwork		34.0	21.6	48.6	34.0
soft	structure		17.8	9.39	29.3	17.8
Area 2						
stiff	earthwork		21.3	8.38	49.1	21.3
stiff	structure		11.2	4.22	27.8	11.2

Table 5-4: Threshold distances for various groups and threshold velocities. Data: N-M-SM, with or without discontinuities and additional 0.7 and 0.6 multiplication factor for soft and stiff soils respectively, accounting for the effect of discontinuities. Compare the results of this table with Table 5-3.

			TNO 2020 – this report			
	Track	v_{th}=	2 mm/s		1.2 mm/s	
Soil	substructure		*H*<12m	*H*>12m	*H*<12m	*H*>12m
Area 1						
soft	earthwork		20.8	9.03	41.7	20.8
soft	structure		6.52	2.50	15.7	6.52
Area 2						
stiff	earthwork		8.73	3.14	24.3	8.73
stiff	structure		4.28	1.54	11.9	4.28

5.2 Advice for the distances in the protocol

For the application in the Prorail protocol, table 5.1 is the basis for step 1 (beoordelingsstap 1). In line with the protocol, the distances will be rounded off to factors of 5 meters (and below 10 meters rounded off to meters. The distances from the current protocol are given between brackets.

Table 5-5: Proposed threshold distances for various groups and threshold velocities for beoordelingsstap 1

		[mm/s]	TNO 2020 – this report			
	Track	v_{th}=	2		1.2	
Soil	substructure		*H*<12m	*H*>12m	*H*<12m	*H*>12m
Area 1						
soft	earthwork		35 (30)	20 (15)	65 (60)	35 (35)
soft	structure		15 (20)	5 (6)	30 (25)	15 (10)
Area 2						
stiff	earthwork		25 (20)	9 (6)	70 (30)	25 (15)
stiff	structure		15 (10)	5 (6)	35 (20)	15 (10)

For step 2 (beoordelingsstap 2), the following distances are derived, two tables are constructed.

Table 5-6: Threshold distances for various groups and threshold velocities for beoordelingsstap 2

			TNO 2020 – this report			
	Track	v_{th}=	2 mm/s		1.2 mm/s	
Soil	substructure		*H*<12m	*H*>12m	*H*<12m	*H*>12m
Area I						
soft	earthwork		35 (20)	25 (10)	50 (40)	35 (20)
soft	structure		20 (10)	10 (6)	30 (15)	20 (6)
Area 2						
stiff	earthwork		25 (10)	9 (6)	50 (20)	25 (6)
stiff	structure		15 (6)	5 (6)	30 (10)	15 (6)

The difference between the former protocol and this table is large and the distances do not differ much from the distances in step 1. Therefore it is recommended to base the values in the protocol on the distances given in Table 5-7. The differences between the two table have been given in Table 4-5.

Table 5-7: Threshold distances for various groups and threshold velocities for beoordelingsstap 2 (based on the full data set and accounted for by factors)

			TNO 2020 – this report			
	Track	v_{th}=	2 mm/s		1.2 mm/s	
Soil	substructure		*H*<12m	*H*>12m	*H*<12m	*H*>12m
Area 1						
soft	earthwork		25 (20)	10 (10)	45 (40)	25 (20)
soft	structure		7 (10)	3 (6)	20 (15)	7 (6)
Area 2						
stiff	earthwork		9 (10)	4 (6)	25 (20)	9 (6)
stiff	structure		5 (6)	2 (6)	15 (10)	5 (6)

In the update of the 2004 protocol relative to the 1998 protocol it was decided not to use shorter distances than already included. If this is the preference now also, the higher value of the two values in the cells should be chosen as distances in the updated protocol.

5.3 Recommendations

The following aspects might be worthwhile to explore in later studies:

- The effect of seasonal trends, e.g. apparent increase in peak velocity during the summer period, could be analyzed via longer measurements (a few years long).
- The effect of vibration exposure duration on the threshold exceedance probability and in turn on the threshold distances.
- The effect of considering all train passages rather than using an incomplete dataset.
- The effect of dependence between block maxima could be further analyzed by:
  - Obtaining a complete dataset of train passages and analyzing autocorrelation over time.
  - If considerable autocorrelation is found then the uncertainty model could be extended to account for dependent extremes.

References

Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference. A Practical Information-Theoretic Approach* (second ed.). New York: Springer-Verlag.

CEN. (2002). Eurocode 0: Basis of structural design.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. London: Springer-Verlag.

Dorfman, R. (1938). A Note on the δ-Method for Finding Variance Formulae. *The Biometric Bulletin, 1*, 129-137.

Heijnen, G. W. J. (1996). *Overzicht van in Nederland verrichte onderzoeken naar optredende trillingen in woningen veroorzaakt door passerend trainverkeer*. NS Technisch Onderzoek. Report number: 9620031.

Roos, W. (2004a). *Beslisdocument behorend bij het schadeprotocol (rapportnummer: 2004-CI-R0082)* TNO. Report number: 2004-CI-R0082 (TNO reference: 2004-CI-B0500/RSW).

Roos, W. (2004b). *Protocol voor het beoordelen van klachten of claims inzake schade door trillingen afkomstig van railverkeer*. TNO. Report number: 2004-CI-R0082.

Rózsás, Á. (2016). *Snow extremes and structural reliability.* (PhD), Budapest University of Technology and Economics, Budapest, Hungary.

SBRCURnet. (2017). *SBR Trillingsrichtlijn A : Schade aan bouwwerken*. Delft.

van Staalduinen, P. C. (1999). *Protocol voor het beoordelen van klachten of claims inzake schade door trillingen afkomstig van railverkeer*. Report number: 98-CON-R1403.

Glossary

The terminology and notation used here are based on the terminology accepted and used in (*i*) structural engineering; and (*iii*) mathematical statistics and probability theory. The roman numerals indicate precedence in case of conflicting terminology or notation in different fields. For clarity, the definition of some key terms are given here.

Model: a mathematical representation of selected characteristics of an object or phenomenon.

Physical model: a deterministic model which describes/represents a physical phenomenon. Note that it can be empirical, first principles based, analytical (symbolic), numerical, etc. Provided with the same inputs it always yields the same outputs. Examples of physical models:

- a standardized, symbolic formula to calculate the shear resistance of a reinforced concrete beam;
- a nonlinear finite element model.

Real world data: data that is collected from cases in real-world settings, opposed to controlled experiments and *synthetic data*.

Synthetic data: data that is generated from a fully known model opposed to real data that is measured on a real system. The model used to generate synthetic data is often meant to represent a real system. In this work, synthetic data refers to data that is generated (simulated) from a finite element model and contaminated with random realizations from a known probabilistic model to add uncertainty.

Table Glossary-0-1: Recording of the changes in reformatting the passage maxima database: N-M-SM.

Dutch (original)	English (used in this report and project in general)
opleverversie QE 1912282 Database ProRail 2004 - 2019 V1.1.xlsx	vibration_data_all_time_max_2019.csv
Firma	company
Dossier nr. firma	document_id
Meetduur	measurement_duration
Vtop in mm/s	vtop mm/s
Hz	dominant_freq_vtop Hz
Datum	date_vtop
Tijd	time_vtop
Overweg of Kw	structure_nearby
Afstand tot overweg / KW	distance_to_structure
Afstand tot spoor	distance_to_track
Spoorbaan type	sleeper_type
betonnen bielzen op aardebaan	concrete on soil
kunstwerk	structure
houten bielzen op aardebaan	timber on soil
Type trein	train_type
Goederen- en reizigerstreinen	cargo and passenger
Alleen Reizigers	passenger
Maximale Rijsnelheid	max_travel_speed

Exacte rijsnelheid	travel_speed_vtop
Type gebouw/object	object_type
vrijstaand	freestanding
half vrijstaand	semi-freestanding
anders	other
tussenwoning	terrace house
Trilling- gevoelig	
Grondslag	soil
zand	sand
leem	loam
klei	clay
Schade	damage
nee	F
Kunstwerken	structure / civil engineering work
Aardebaan	earthwork

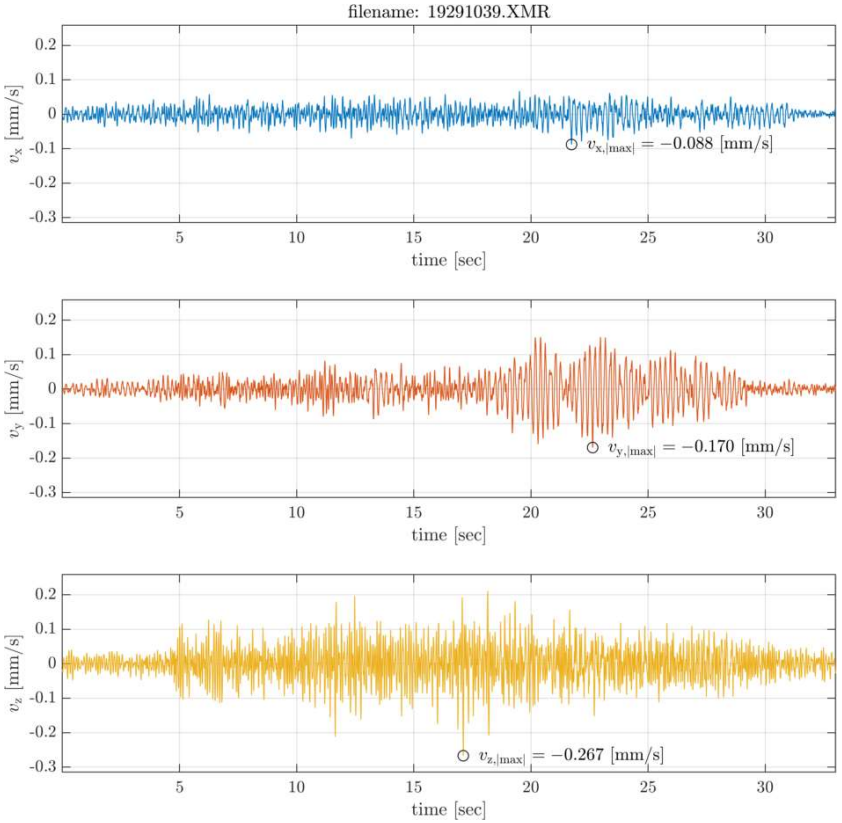
Table Glossary-0-2: Recording of the changes in reformatting the passage maxima database: N-P-SM.

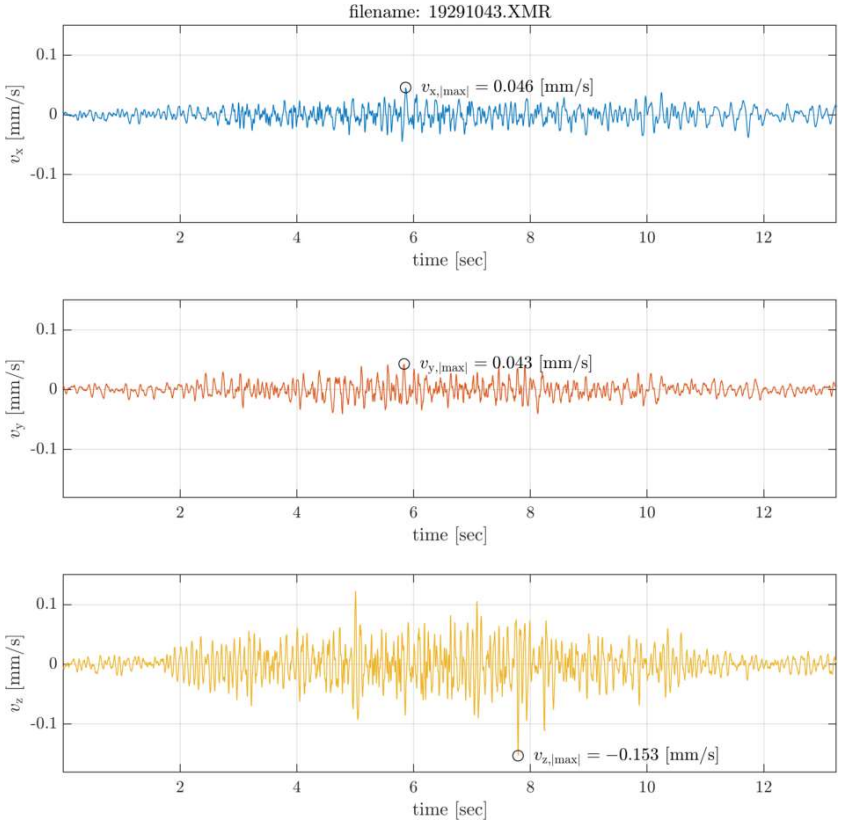
Dutch (original)	English (used in this report and project in general)
ProRail_Brabantroute_snelheden.xlsx	vibration_data_passage_max_oisterwijk_2019.csv vibration_data_passage_max_rijen_2019.csv vibration_data_passage_max_dorst_2019.csv
Datum	date_time
Treintype	train_type
Spoor	track
Snelheid	speed
Treinnummer	train_number
Type trein	train_type
21X v_top.mm/s	v_top_x mm/s 21
21X acc.m/s/s	accm_x mm/s^2 21
21X v_eff	v_eff_x mm/s 21
21X Hz	freq_x Hz 21

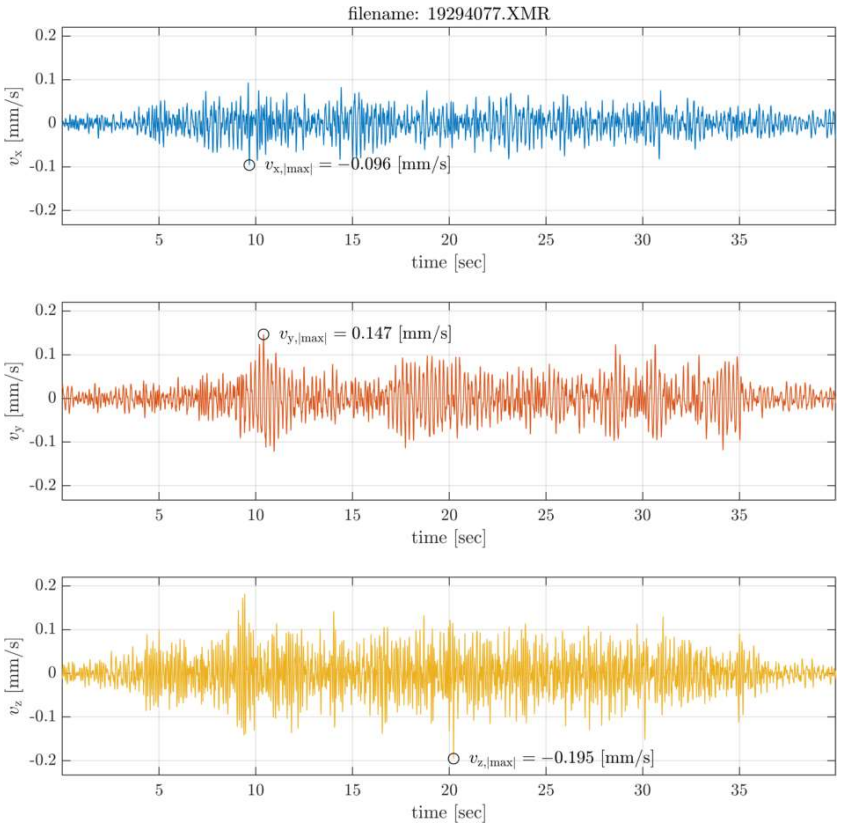
The same conversion is applied to all directions (X, Y, Z) and locations (Oisterwijk, Rijen, Dorst).

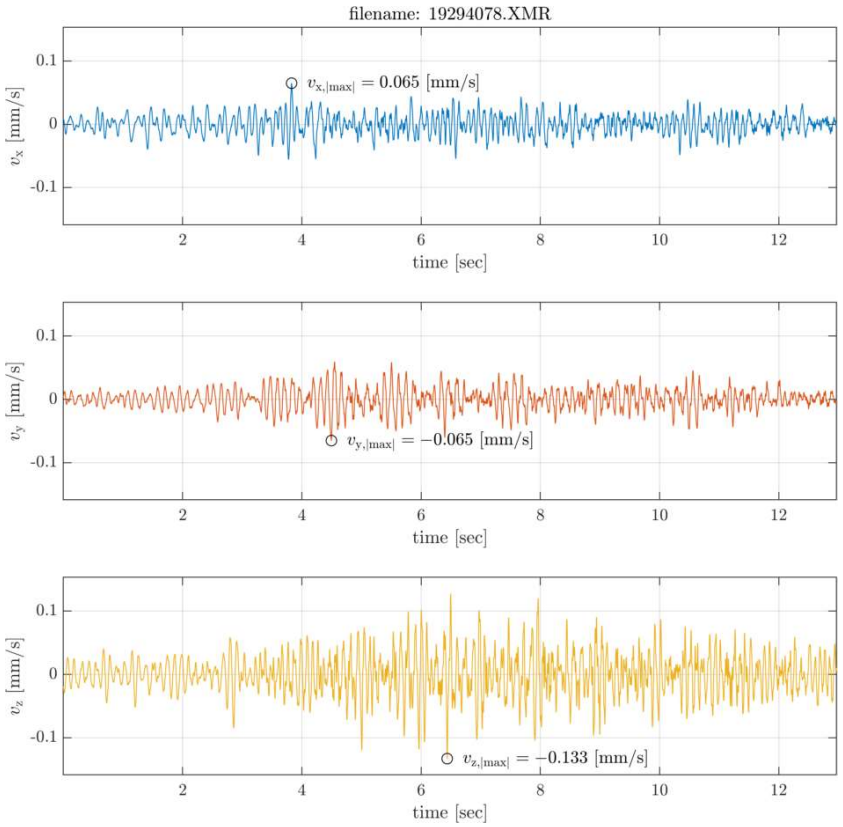
A Velocity time series

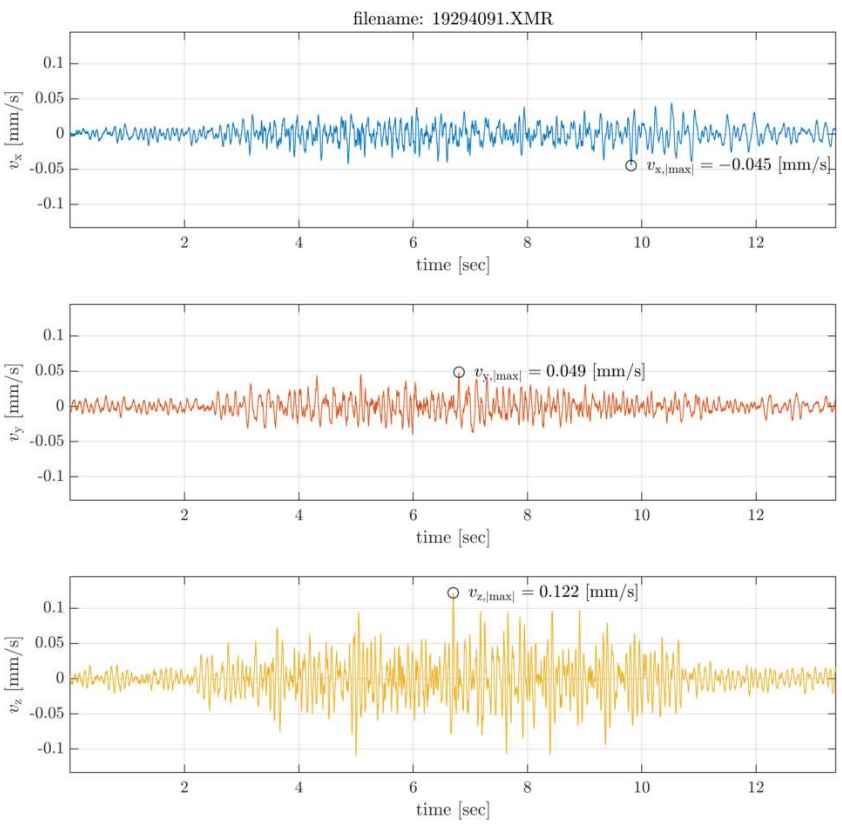
For convenience the velocity time series provided by the client are visualized in the figures below. x and y are perpendicular axes in the horizontal plane, and z is the vertical direction. The absolute maximum value are indicated with a black circle on each time trace. Data: N-P-TS.

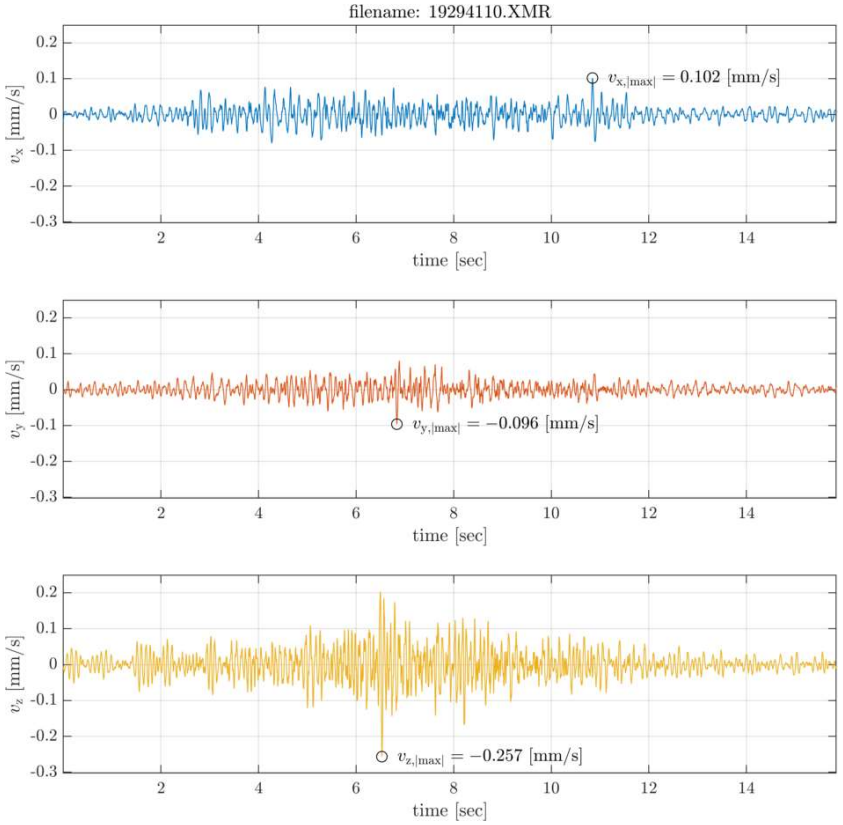


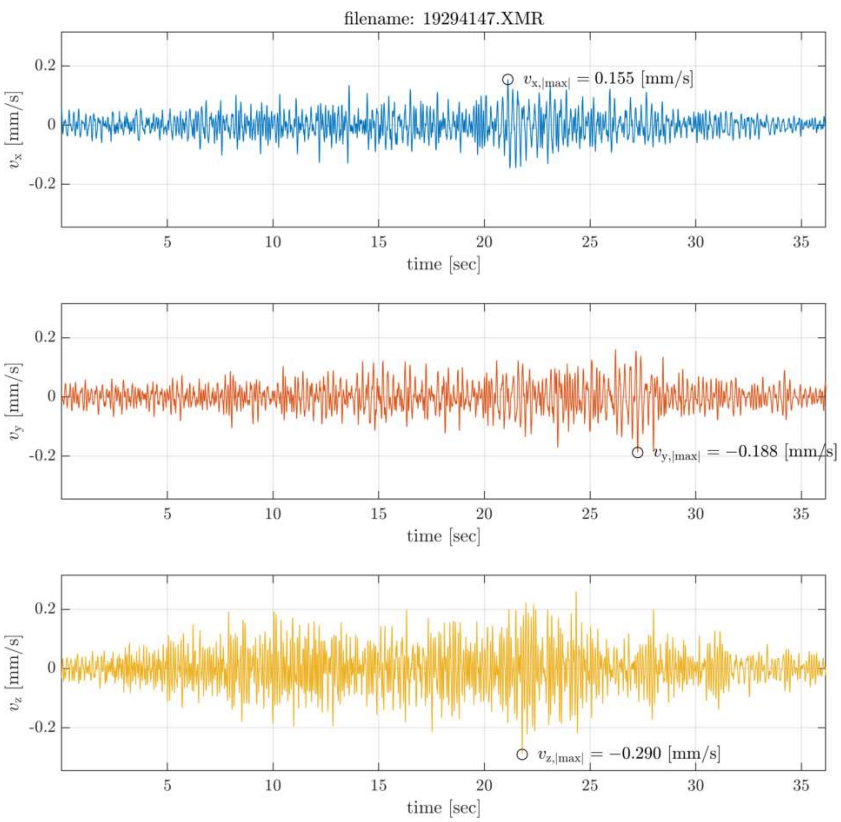


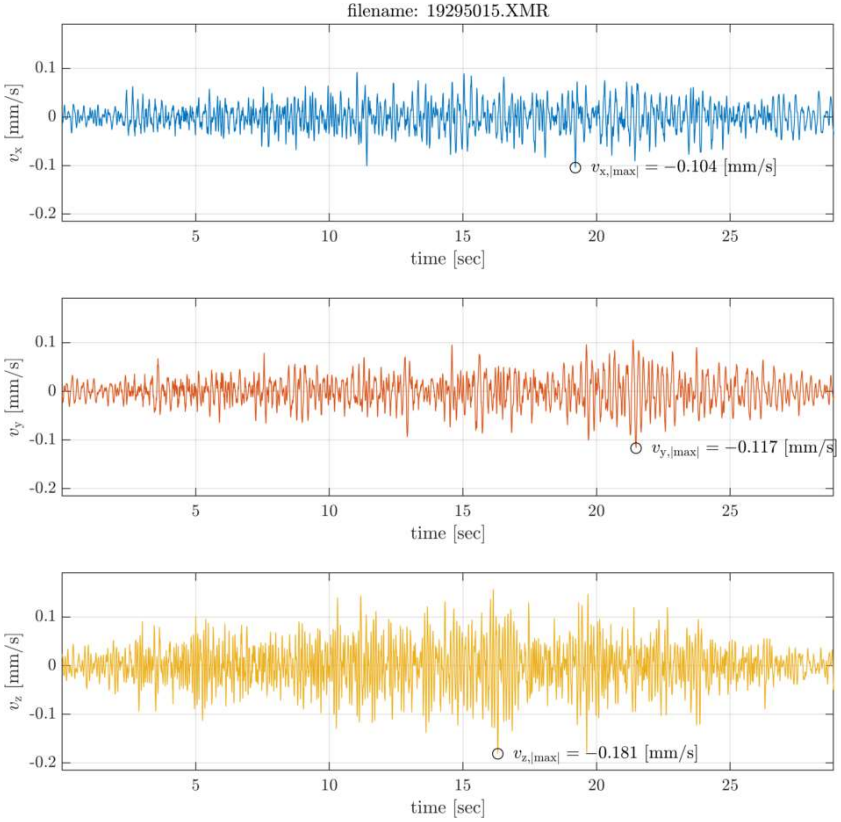

filename: 19294077.XMR
$v_{x,|max|} = -0.096$ [mm/s]
$v_{y,|max|} = 0.147$ [mm/s]
$v_{z,|max|} = -0.195$ [mm/s]
v_x [mm/s]
v_y [mm/s]
v_z [mm/s]
time [sec]

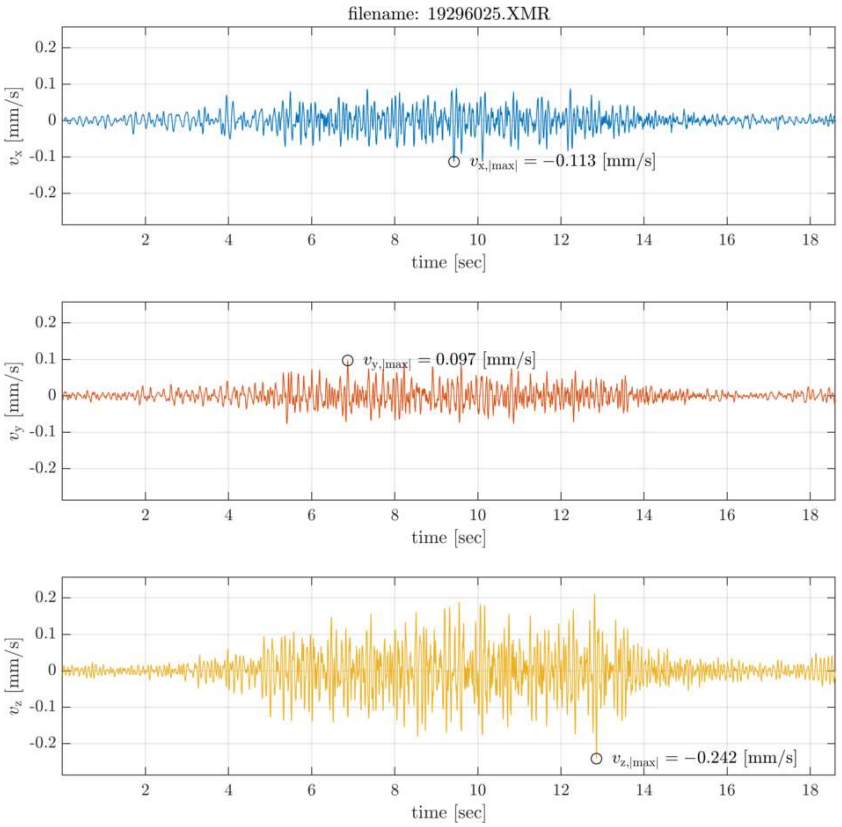

filename: 19294078.XMR
$v_{x,|max|} = 0.065$ [mm/s]
$v_{y,|max|} = -0.065$ [mm/s]
$v_{z,|max|} = -0.133$ [mm/s]
v_x [mm/s]
v_y [mm/s]
v_z [mm/s]
time [sec]

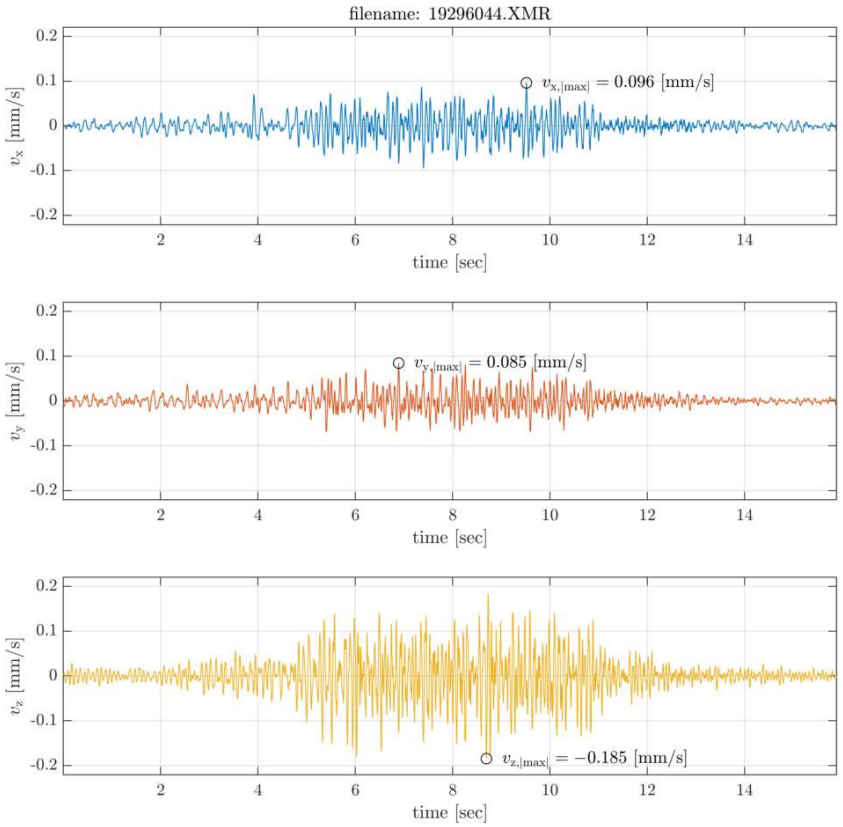

filename: 19294091.XMR
$v_{x,|max|} = -0.045$ [mm/s]
$v_{y,|max|} = 0.049$ [mm/s]
$v_{z,|max|} = 0.122$ [mm/s]
v_x [mm/s]
v_y [mm/s]
v_z [mm/s]
time [sec]


filename: 19294110.XMR
$v_{x,|max|} = 0.102$ [mm/s]
$v_{y,|max|} = -0.096$ [mm/s]
$v_{z,|max|} = -0.257$ [mm/s]
v_x [mm/s]
v_y [mm/s]
v_z [mm/s]
time [sec]


filename: 19294147.XMR
$v_{x,|max|} = 0.155$ [mm/s]
$v_{y,|max|} = -0.188$ [mm/s]
$v_{z,|max|} = -0.290$ [mm/s]
v_x [mm/s]
v_y [mm/s]
v_z [mm/s]
time [sec]


filename: 19295015.XMR
$v_{x,|max|} = -0.104$ [mm/s]
$v_{y,|max|} = -0.117$ [mm/s]
$v_{z,|max|} = -0.181$ [mm/s]
v_x [mm/s]
v_y [mm/s]
v_z [mm/s]
time [sec]


filename: 19296025.XMR
$v_{x,|max|} = -0.113$ [mm/s]
$v_{y,|max|} = 0.097$ [mm/s]
$v_{z,|max|} = -0.242$ [mm/s]
v_x [mm/s]
v_y [mm/s]
v_z [mm/s]
time [sec]


filename: 19296044.XMR
$v_{x,|max|} = 0.096$ [mm/s]
$v_{y,|max|} = 0.085$ [mm/s]
$v_{z,|max|} = -0.185$ [mm/s]
v_x [mm/s]
v_y [mm/s]
v_z [mm/s]
time [sec]

B Additional visualization of maxima

B.1 N-M-SM: measurement duration maxima

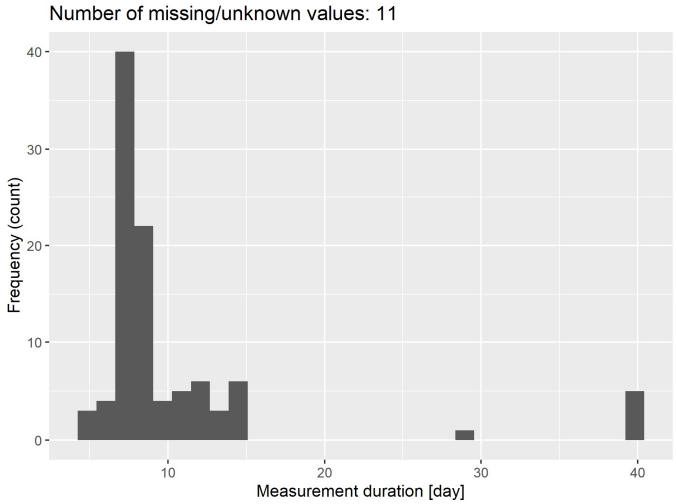


Figure B-1: Measurement duration distribution in the new dataset: N-M-SM.

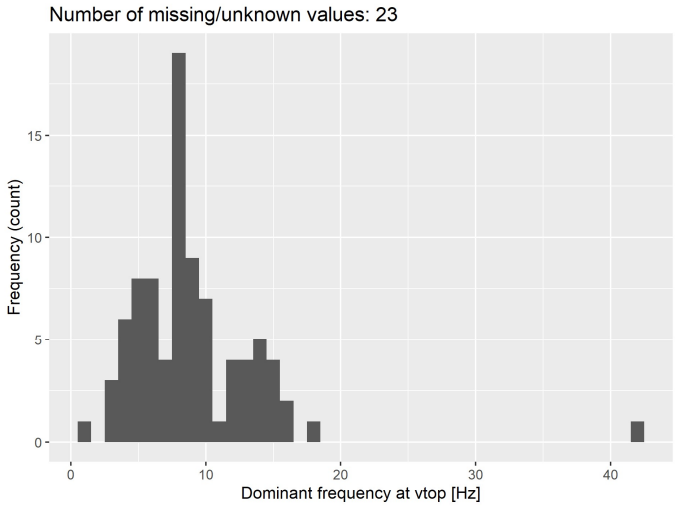


Figure B-2: Dominant frequency distribution in the new dataset: N-M-SM.

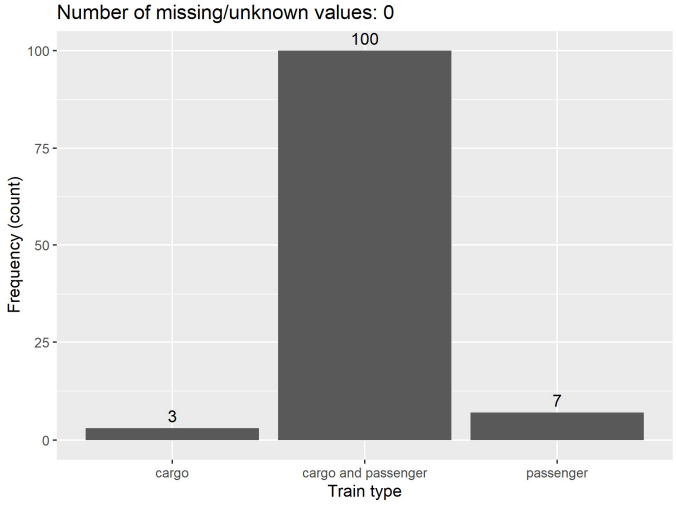


Figure B-3: Train type distribution in the new dataset: N-M-SM.

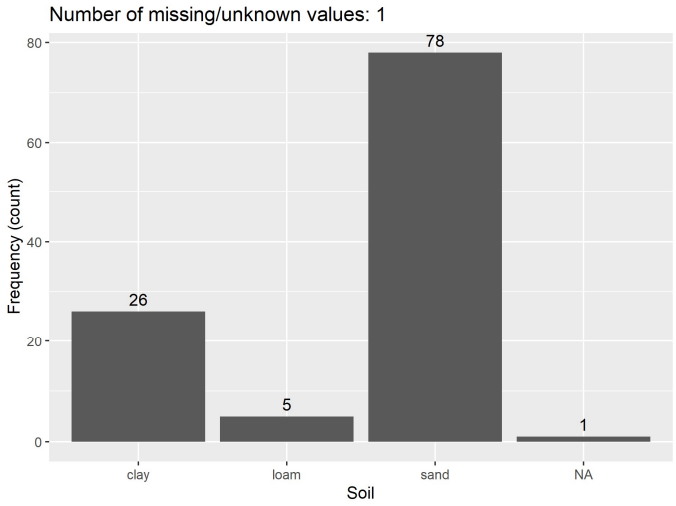


Figure B-4: Soil type distribution in the new dataset: N-M-SM.

Figure B-5: Structure nearby distribution in the new dataset: N-M-SM: N-M-SM.

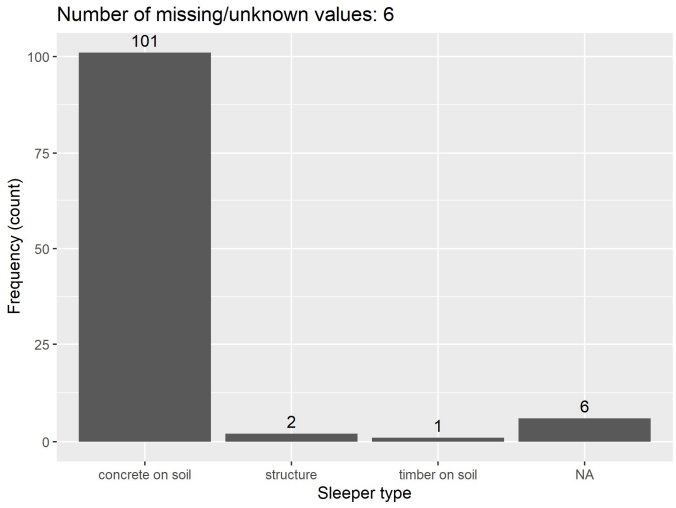


Figure B-6: Sleeper type distribution in the new dataset: N-M-SM.

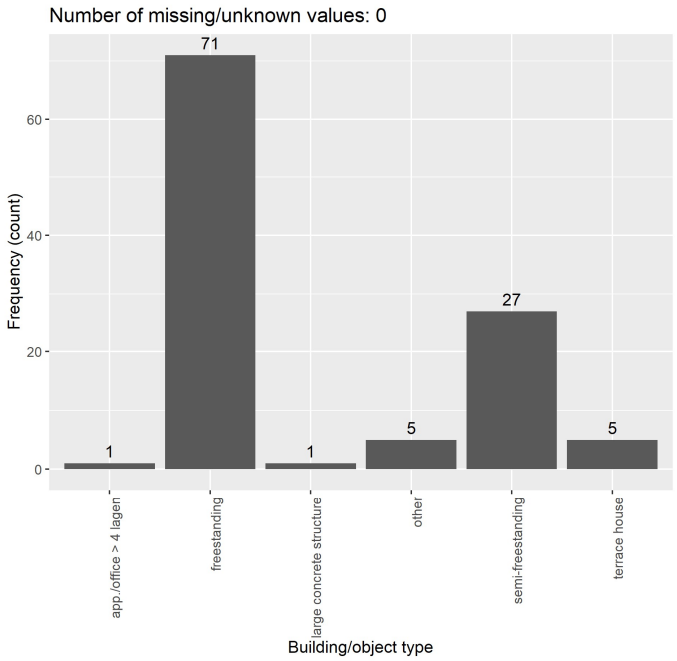


Figure B-7: Building/object type distribution in the new dataset: N-M-SM.

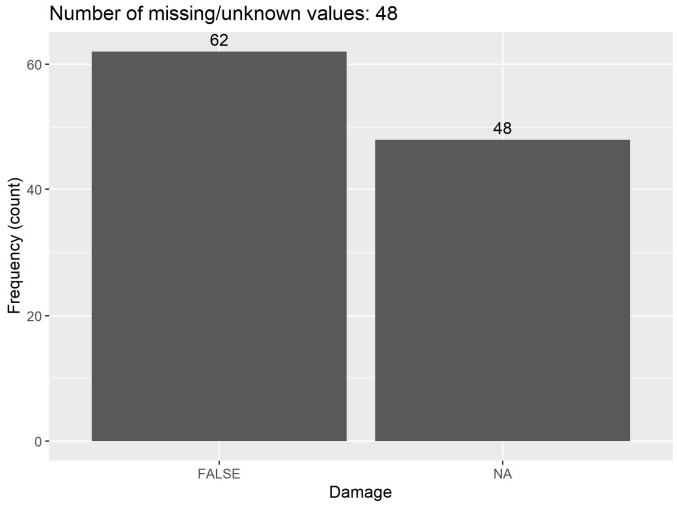


Figure B-8: Damage distribution in the new dataset: N-M-SM.

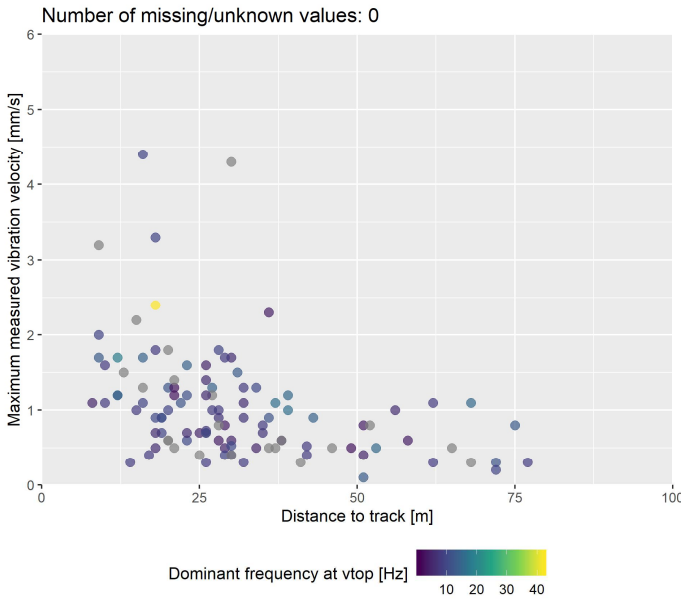


Figure B-9: Maximum velocity vs distance to track. Data: N-M-SM.

C Changing block size extremes

C.1 Independent extremes

C.1.1 General

Extreme value distributions are inseparably connected to time. For example in case of block maxima (block indicates the time interval from which the extreme value is taken). If the maxima are available for different block duration or one want so change to another block duration the transformation of the extreme value distribution is needed. For example going from 1-year extremes to 50 extremes. This transformation can be accomplished by assuming that the **block maxima are independent**, then the cumulative distribution function (cdf) of *n*-block maxima is:

$$F_n(x) = F_1(x)^n \,. \tag{C.1}$$

Since in general $F_n(.)$ is not a special distribution we have to use numerical methods to obtain most of the properties of the distribution. For this is we often need the probability density function which can be obtained as:

$$f_n(x) = \frac{\mathrm{d}F_n(x)}{\mathrm{d}x} = \frac{\mathrm{d}F_1(x)^n}{\mathrm{d}x} = n \cdot F_1(x)^{n-1} \cdot f_1(x) \,. \tag{C.2}$$

For numerical reasons it is also advantageous to have the log of the probability density function:

$$\log\left(f_n(x)\right) = \log(n) + (n-1) \cdot F_1(x)^{n-1} + \log\left(f_1(x)\right) \,. \tag{C.3}$$

C.1.2 Gumbel distribution

The Gumbel distribution has the nice property that its cdf preserves its Gumbel type by raising it to a power, i.e. $F_n(.)$ is also Gumbel distributed. The standard deviation and mean of *n*-block maxima Gumbel from that of the annual maxima can be derived using Eq.(C.1), omitting the details and just presenting the results:

$$\begin{aligned} \mathrm{std}_n &= \mathrm{std}_1 = \mathrm{std} \\ \mathrm{mean}_n &= \mathrm{mean}_1 + \frac{\sqrt{6}}{\pi} \cdot \ln(n) \cdot \mathrm{std} \end{aligned} \,.$$

Using these the coefficient of variation can be expressed as:

$$\mathrm{cov}_n = \frac{\mathrm{std}_n}{\mathrm{mean}_n} = \frac{1}{\frac{1}{\mathrm{cov}_1} + \frac{\sqrt{6}}{\pi} \cdot \ln(n)} \,.$$

In a more compact form for one -block and *n*-block reference periods:

$$\mathrm{cov}_1 = \left(\frac{1}{\mathrm{cov}_n} - \frac{\sqrt{6}}{\pi} \cdot \ln(n) \right)^{-1}$$

$$\mathrm{cov}_n = \left(\frac{1}{\mathrm{cov}_1} + \frac{\sqrt{6}}{\pi} \cdot \ln(n) \right)^{-1} .$$

D The impact of some modelling assumptions

This annex presents the analysis of the impact of some modelling assumptions. The impact is mainly investigated on the threshold distances as the most important final outcome of this study.

D.1 Exponent in the Barkan formula

Theoretically the exponent (*n*) in the Barkan formula (Eq. (3.3)) is 0.5 for point sources and the assumption of Rayleigh waves. Although the real world is expected to exhibit a more complex behavior than that is assumed to derive the Barkan formula and the above exponents, we tested the impact of the exponent on the threshold distances and checked which model fits the data better (*n*=0.5 vs. *n*=1.0).

The same analysis as presented in section 4 is performed for *n*=1.0 for the case of discontinuity = `yes or no`, and the results are summarized in Table D-1. The following trend is observable by switching from *n*=0.5 to n=1.0: for *n*=0.5 threshold distances below 30m the *n*=1.0 threshold distance is larger, while for >30m it is smaller. For smaller threshold distances the difference due to the different exponent can be as large as two fold.

To assess which model provides a better fit to the data the Akaike information criterion[1] (AIC) is calculated for each model that is used to obtain the threshold distances in Table D-1. The AICs are summarized in Table D-2. The smaller the AIC value, the better the model, if the AIC difference between two models -- fitted to the same data -- is larger than 10 that means that the model with larger AIC has essentially no empirical support over the other model (Burnham & Anderson, 2002). Table D-2 shows that the *n*=0.5 is substantially better than model *n*=1.0.

[1] AIC is an asymptotic information criterion that is based on the premise that the model with the smallest information loss (Kullback-Leibler divergence) should be preferred. In the absence of the true model, the information loss cannot be calculated in absolute terms; however, the models can be compared and their relative “strength” can be expressed by the difference in AICs. AIC penalizes model complexity (Burnham & Anderson, 2002).

Table D-1: Threshold distances for various groups and threshold velocities using exponent 0.5 and 1.0 in the Barkan formula. Data: N-M-SM, with or without discontinuities.

		[mm/s]	TNO 2020 – this report			
	Track	v_{th}=	2		1.2	
Soil	substructure		*H*<12m	*H*>12m	*H*<12m	*H*>12m
n=0.5 in the Barkan formula (Eq. (3.3))						
Area 1						
soft	earthwork		34.4	16.4	61.9	34.4
soft	structure		12.2	4.91	26.9	12.2
Area 2						
stiff	earthwork		24.3	8.73	67.4	24.3
stiff	structure		11.9	4.28	33.0	11.9
n=1.0 in the Barkan formula (Eq. (3.3))						
Area 1						
soft	earthwork		32.5	19.5	54.2	32.5
soft	structure		16.3	9.75	27.1	16.3
Area 2						
soft	earthwork		26.6	16.0	44.4	26.6
soft	structure		18.6	11.2	31.1	18.6

Table D-2: Akaike information criteria (AIC) for models fitted to various groups and threshold velocities using exponent 0.5 and 1.0 in the Barkan formula. The same models as those were used to obtain the threshold distances in Table D-2. Data: N-M-SM, with or without discontinuities.

	Track	TNO 2020 – this report	
Soil	substructure	*H*<12m	*H*>12m
n=0.5 in the Barkan formula (Eq. (3.3))			
Area 1			
soft	earthwork	44.1	44.1
soft	structure	44.1	44.1
Area 2			
stiff	earthwork	71.5	71.5
stiff	structure	71.5	71.5
n=1.0 in the Barkan formula (Eq. (3.3))			
Area 1			
soft	earthwork	52.0	52.0
soft	structure	52.0	52.0
Area 2			
stiff	earthwork	92.6	92.6
stiff	structure	92.6	92.6

D.2 Distribution type

The impact of using different distributions for *E* in Eq.(3.2): Gumbel, normal, and generalized extreme value (GEV) distributions is analyzed. The results in terms of threshold distances are summarized in Table D-3 and Table D-4. The results show that the distribution type has a relatively small influence on the 7-day threshold distance. This is attributed to the large target probability (18%), at which level the difference between distributions is small. The distributions are fitted to the same

data hence their first few moments are expected to be similar. larger differences would be expected farther at the tails. The differences due to the distribution type start to manifest when other than 7-day exposure periods are considered, that is analyzed in Annex E.

Table D-3: Threshold distances for various groups and threshold velocities. Comparison of Gumbel, normal, and GEV distributions. Data: N-M-SM, with or without discontinuities.

		[mm/s]	TNO 2020 – this report			
	Track	v_{th}=	2		1.2	
Soil	substructure		*H*<12m	*H*>12m	*H*<12m	*H*>12m
E~Gumbel (Eq.(3.2))						
Area 1						
soft	earthwork		34.4	16.4	61.9	34.4
soft	structure		12.2	4.91	26.9	12.2
Area 2						
stiff	earthwork		24.3	8.73	67.4	24.3
stiff	structure		11.9	4.28	33.0	11.9
E~Normal (Eq.(3.2))						
Area 1						
soft	earthwork		35.2	16.7	63.7	35.2
soft	structure		12.4	5.00	27.5	12.4
Area 2						
stiff	earthwork		25.2	9.08	70.0	25.2
stiff	structure		12.4	4.45	34.3	12.4
E~GEV (Eq.(3.2))						
Area 1						
soft	earthwork		35.1	16.9	62.8	35.1
soft	structure		12.5	5.08	27.5	12.5
Area 2						
stiff	earthwork		24.0	8.64	66.7	24.0
stiff	structure		11.8	4.24	32.7	11.8

Table D-4: Threshold distances for various groups and threshold velocities. Comparison of Gumbel, normal, and GEV distributions. Data: N-M-SM, without discontinuities.

		[mm/s]	TNO 2020 – this report			
	Track	v_{th}=	2		1.2	
Soil	substructure		*H*<12m	*H*>12m	*H*<12m	*H*>12m
E~Gumbel (Eq.(3.2))						
Area 1						
soft	earthwork		34.0	21.6	48.6	34.0
soft	structure		17.8	9.39	29.3	17.8
Area 2						
stiff	earthwork		21.3	8.38	49.1	21.3
stiff	structure		11.2	4.22	27.8	11.2
E~Normal (Eq.(3.2))						
Area 1						
soft	earthwork		30.1	19.0	43.0	30.1
soft	structure		15.7	8.25	25.9	15.7
Area 2						
stiff	earthwork		19.5	7.56	45.9	19.5
stiff	structure		10.1	3.80	25.6	10.1
E~GEV (Eq.(3.2))						
Area 1						
soft	earthwork		NC	18.7	NC	27.6
soft	structure		17.3	9.05	26.2	16.2
Area 2						
stiff	earthwork		20.3	8.09	45.7	20.3
stiff	structure		10.8	4.10	26.3	10.8

NC: no convergence.

E Time dependence analysis for N-M-SM

E.1 Exploring the effect of the length of the exposure period

Using the model: Gumbel, soil=all, structure=all, no multiplication factor, mode predictions per different exposure periods are given in Figure E-1. The equally spaced exposure periods show how the impact of exposure period is diminishing over time. Although the difference between the lines at equal time distance is monotonic decreasing these differences still add up to a relatively large total difference over longer times, see the bottom plot of Figure E-1. The diminishing differences are in in line with our expectations; however, the total difference (“shift”) over longer periods such as 1 year goes against our intuition and the frequency of damage claims. To study to what extent the long terms “shift” can be attributed to our modelling decisions we replaced the Gumbel distribution with the GEV distribution. The latter contains the former as a special case and allows for heavier and lighter tail distributions than Gumbel. Hence, if the data favors a Gumbel distribution that will emerge from the analyses, i.e. using GEV does not equate to swapping one assumption to another but more to loosening an assumption. For a selected subset of data – where the effect is the most prominent -- the results of the Gumbel and GEV models are displayed in Figure E-3 and Figure E-4 respectively. It is clear that the “shift” in the GEV prediction is much smaller than in case of the Gumbel model, although the two models differ.

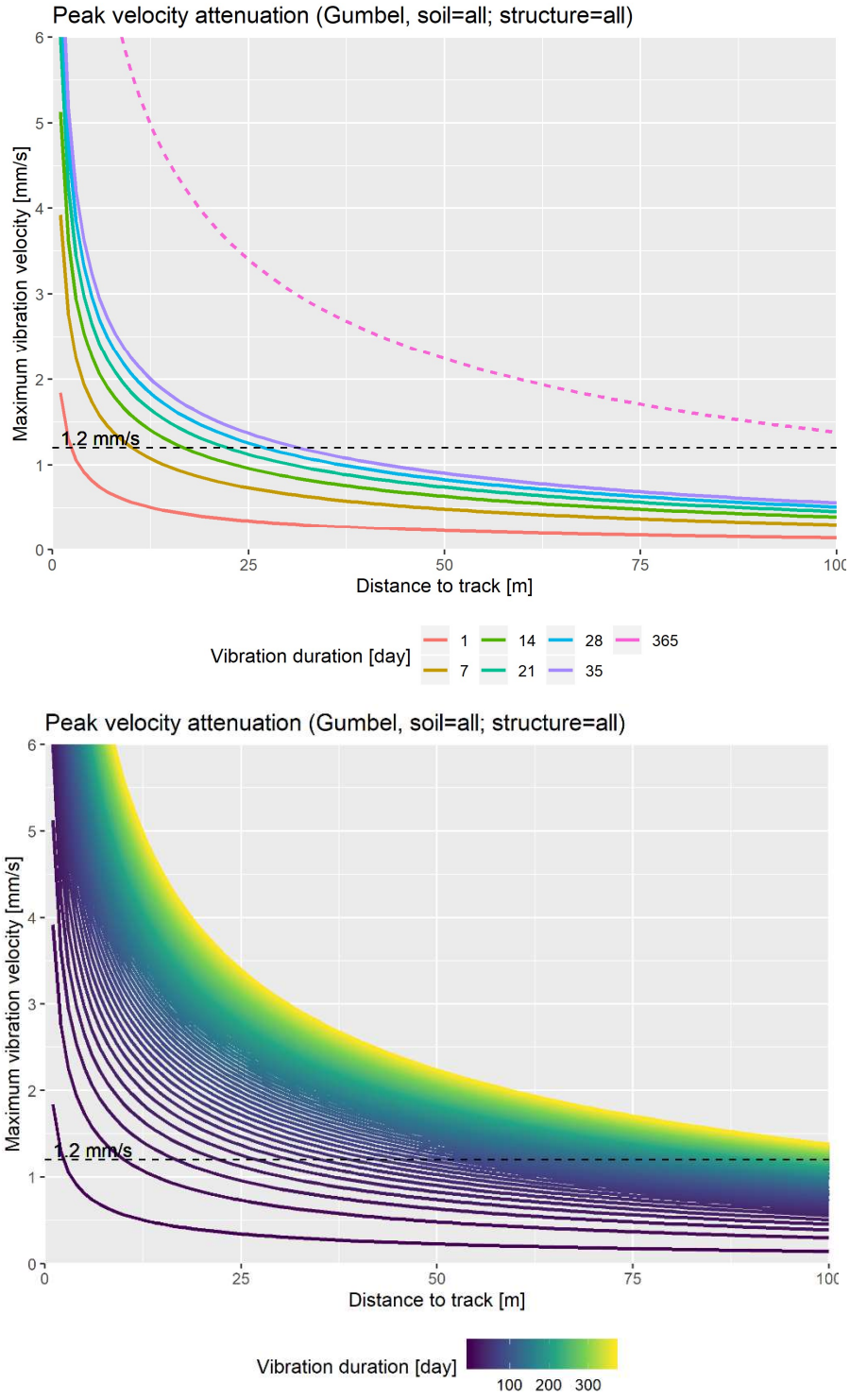


Figure E-1: Illustration of the time dependency of the mode of the fitted peak velocity attenuation function for all soil types and all structure types. Top: for a few selected exposure periods; bottom: for more exposure periods in order to better illustrate the “shift” of a longer exposure period. Data: N-M-SM.

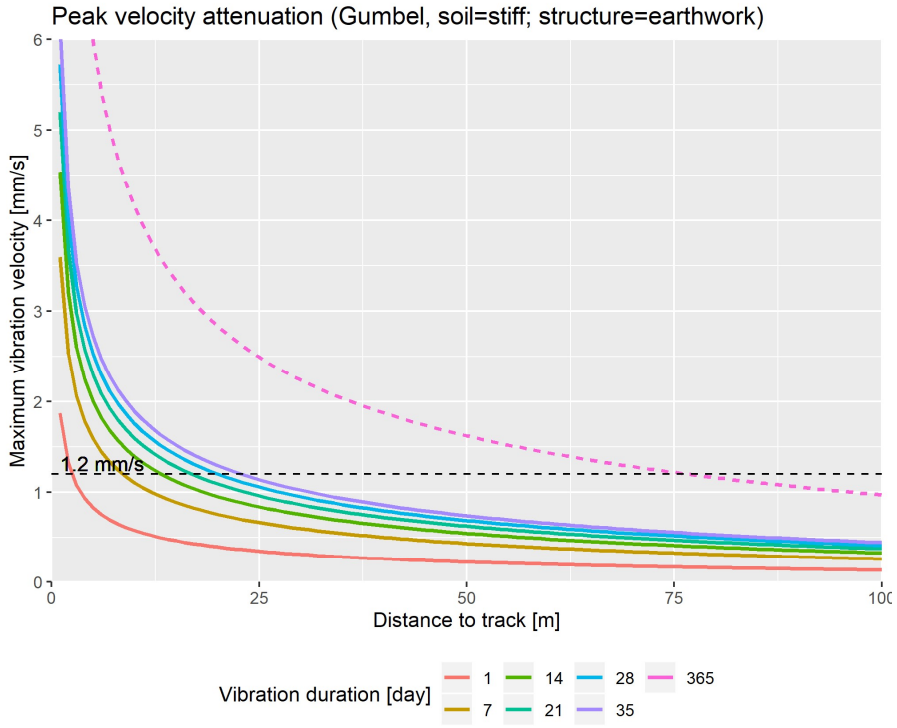


Figure E-2: Illustration of the time dependency of the mode of the fitted peak velocity attenuation function using the Gumbel model. Compare this plot with Figure E-3 and Figure E-4. Data: N-M-SM.

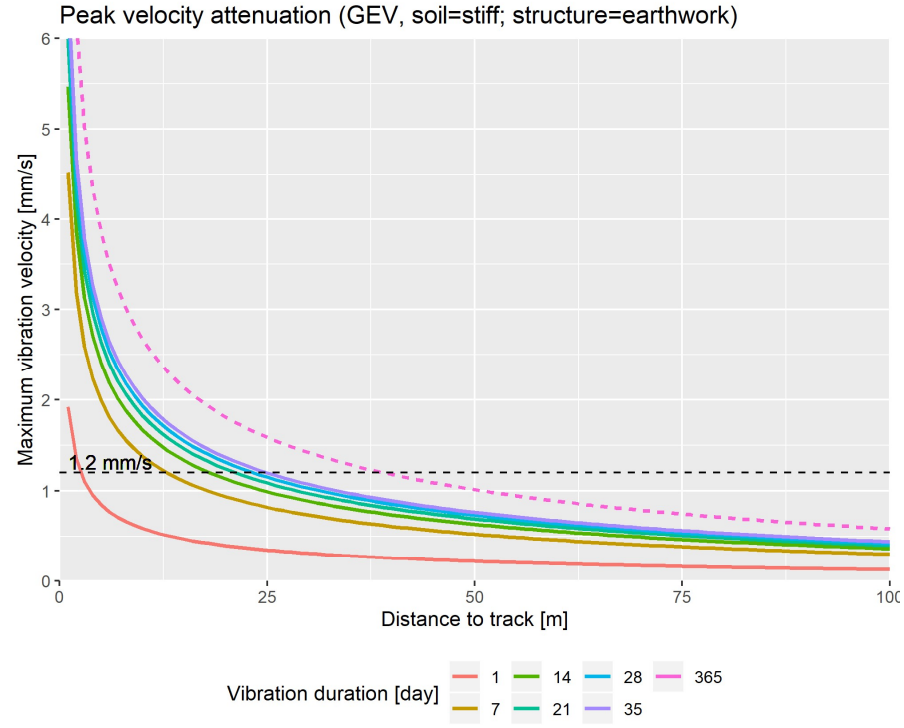


Figure E-3: Illustration of the time dependency of the mode of the fitted peak velocity attenuation function using the GEV model. Compare this plot with Figure E-2 and Figure E-4. Data: N-M-SM.

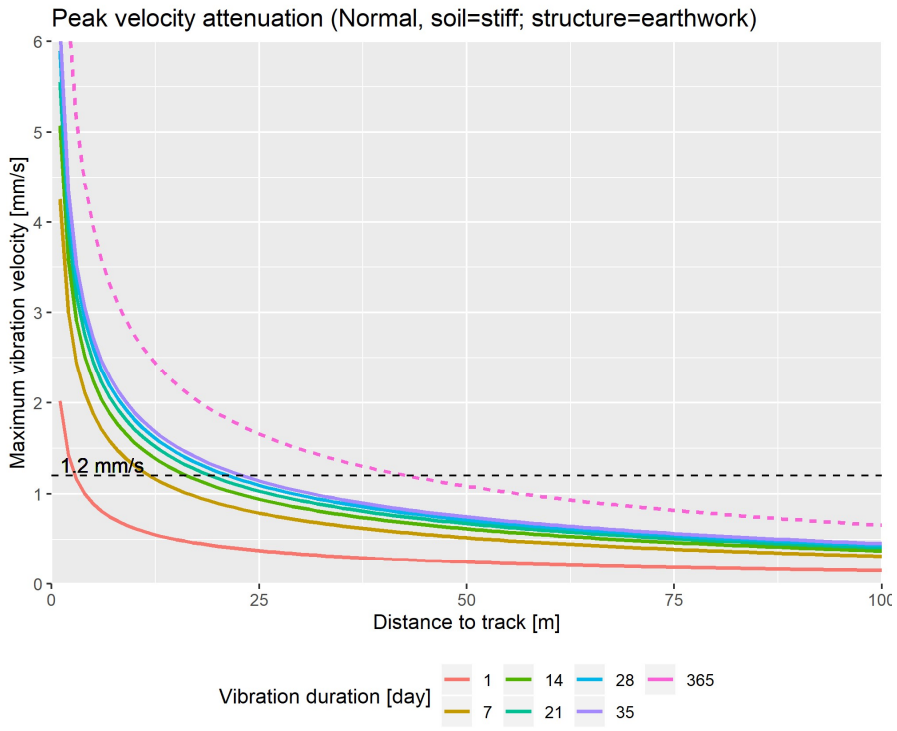


Figure E-4: Illustration of the time dependency of the mode of the fitted peak velocity attenuation function using the Normal distribution. Compare this plot with Figure E-2 and Figure E-3. Data: N-M-SM.

F Time dependence analysis for N-P-SM

F.1 Disclaimer

After the completion of this analysis we learned that the N-P-SM dataset is incomplete, i.e. some train passages are not included; therefore the results presented in this annex should be treated with caution.

F.2 Autocorrelation analysis

F.2.1 Independently of train types

To further investigate any potential periodicity in the data and to examine if the $v_{\text{top,daily}}$ values are dependent, time series and autocorrelation plots are prepared and shown in Figure F-1, Figure F-2, and Figure F-3 for Dorst, Oisterwijk, and Rijen respectively. For each location the time series reveal marked peaks and similarly to the heatmap plots they do not indicate any discernable pattern. The autocorrelation plot of Dorst has a very low correlation (<0.2) even for events a few days apart. For this location the $v_{\text{top,daily}}$ values can be reasonably assumed to be independent. In case of Oisterwijk, with the exception of the 1-day lag, the autocorrelation plot shows a similar pattern to Dorst. The one day lag correlation coefficient is about 0.4. Location Rijen shows higher but still relatively low autocorrelation values. Its 1-day lag correlation coefficient is about 0.5 and it decreases to about 0.25 for the 7-day lag.

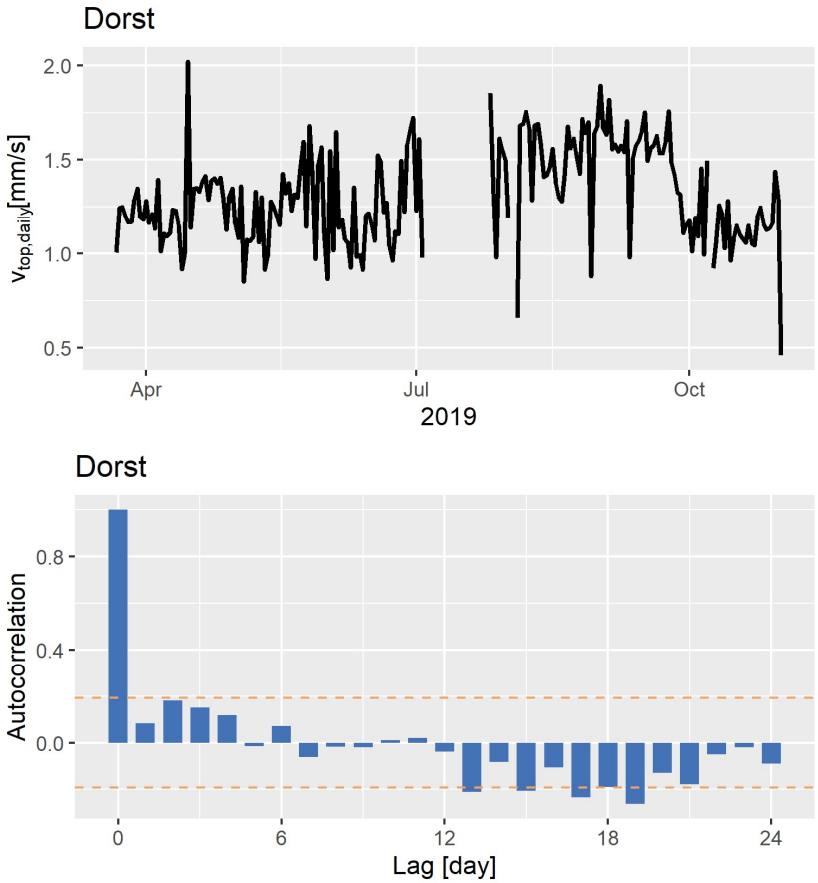


Figure F-1: Daily vibration velocity maxima for Dorst. Time series plot (top) and autocorrelation plot (bottom). The dashed orange lines indicate the endpoints of an approximate 95% confidence interval for a white noise process of the same sample size. Data: N-P-SM.

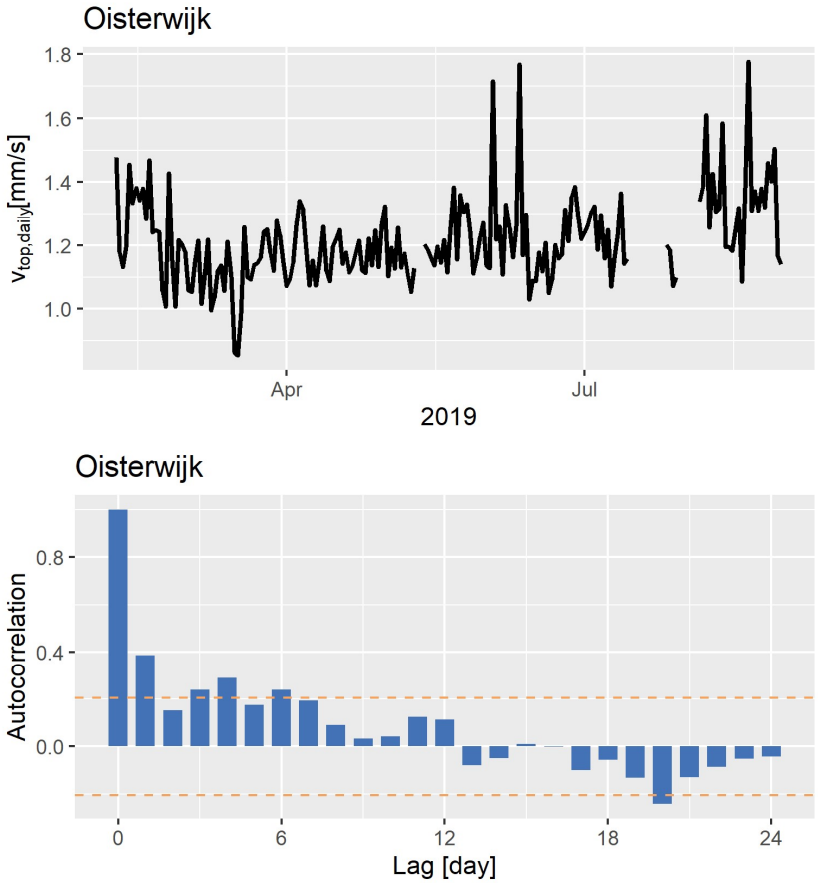


Figure F-2: Daily vibration velocity maxima for Oisterwijk. Time series plot (top) and autocorrelation plot (bottom). The dashed orange lines indicate the endpoints of an approximate 95% confidence interval for a white noise process of the same sample size. Data: N-P-SM.

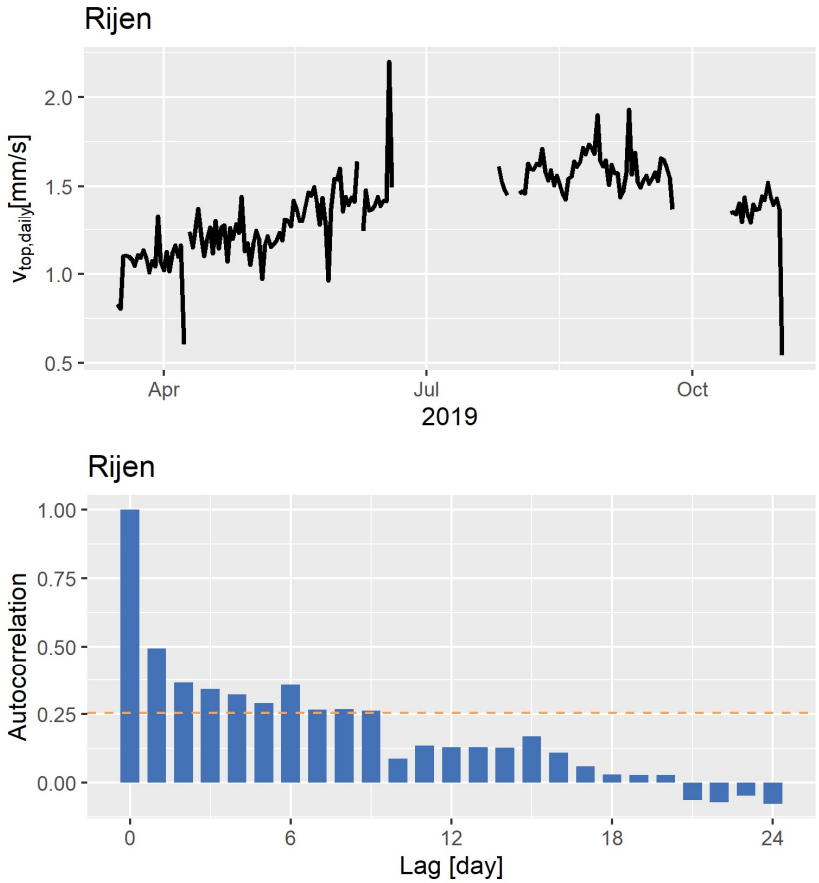


Figure F-3: Daily vibration velocity maxima for Rijen. Time series plot (top) and autocorrelation plot (bottom). The dashed orange lines indicate the endpoints of an approximate 95% confidence interval for a white noise process of the same sample size. Data: N-P-SM.

F.2.2 Accounting for train types

Since the observed low autocorrelation is surprising, we decided to check if higher autocorrelation exists between trains within the same train category. For this purpose, the analyses presented in the previous section is repeated for each location-train type pair. A few selected descriptive statistics are summarized in Table F-1 that reveals considerable differences between locations and between train types. The results in terms of autocorrelation plots are presented in Figure F-4.

It is visible that the autocorrelations are even weaker than for the data grouped only per location. During the completion of this work we learned that the dataset (N-P-SM) does not contain all train passages so that might be a reason of the observed weak autocorrelation.

To further explore the influence of train types and passage speed on the induced vibration velocity v_{top} and corresponding train speed pairs are plotted in Figure F-5 and Figure F-6 considering $v_{top,daily}$ and $v_{top,passage}$ respectively. The plots show little to no dependence of v_{top} on train speed. Moreover, they show that the train can have a considerable effect on v_{top}, for example see Figure F-5, location Oisterwijk and compare SPR to cargo, and IC to ICRmh+TRAXX. The members in each pair have comparable speed while the induced vibrations are markedly different.

Table F-1: Summary of selected descriptive statistics of the passage maxima database per location and train type. ICRmh+TRAXX: ICRmh Intercity passenger wagons and TRAXX engine; IC: InterCity; SPR: Sprinter; ICE: Intercity Express. Data: N-P-SM.

Location		Cargo	ICRmh+TRAXX	ICE	IC	SPR
Dorst	mean($v_{top,passage}$) [mm/s]	1.89	1.85	NA	1.28	2.02
Oisterwijk		1.78	1.58	0.684	1.27	1.03
Rijen		1.50	1.93	0.478	2.20	1.34
Dorst	mean(daily train passages)	27.8	15.3	NA	8.78	6.92
Oisterwijk		40.1	51.4	1.5	6.18	25.7
Rijen		40.7	26.8	1	34.9	3.08
Dorst	total number of train passages	5562	3036	0	1729	1364
Oisterwijk		7302	9403	3	1019	4609
Rijen		6960	4479	3	5904	360

NA: not applicable.

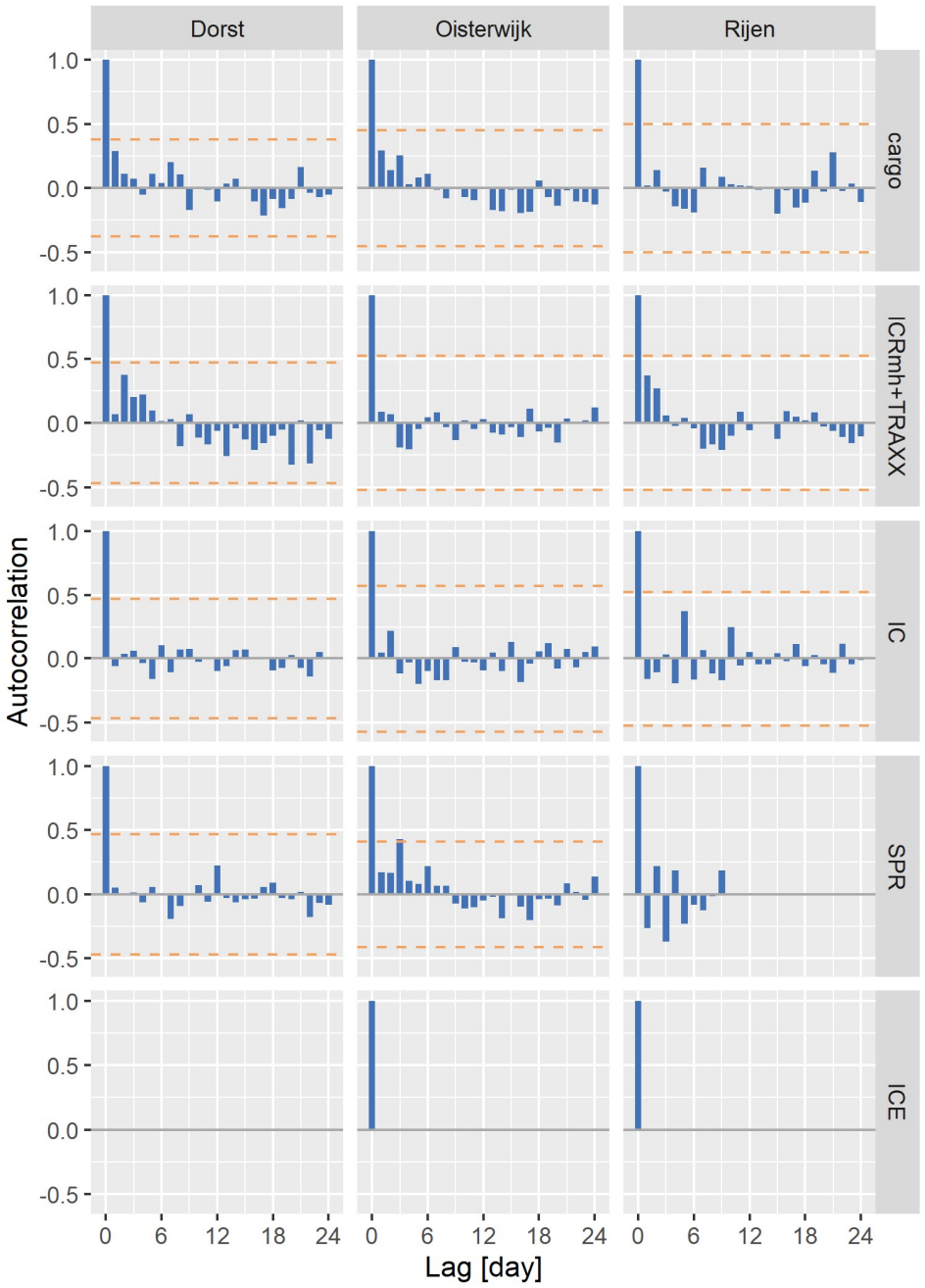


Figure F-4: Daily vibration velocity maxima autocorrelation plots for each location-train type pair. The dashed orange lines indicate the endpoints of an approximate 95% confidence interval for a white noise process of the same sample size. ICRmh+TRAXX: ICRmh Intercity passenger wagons and TRAXX engine; IC: InterCity; SPR: Sprinter; ICE: Intercity Express. Data: N-P-SM.

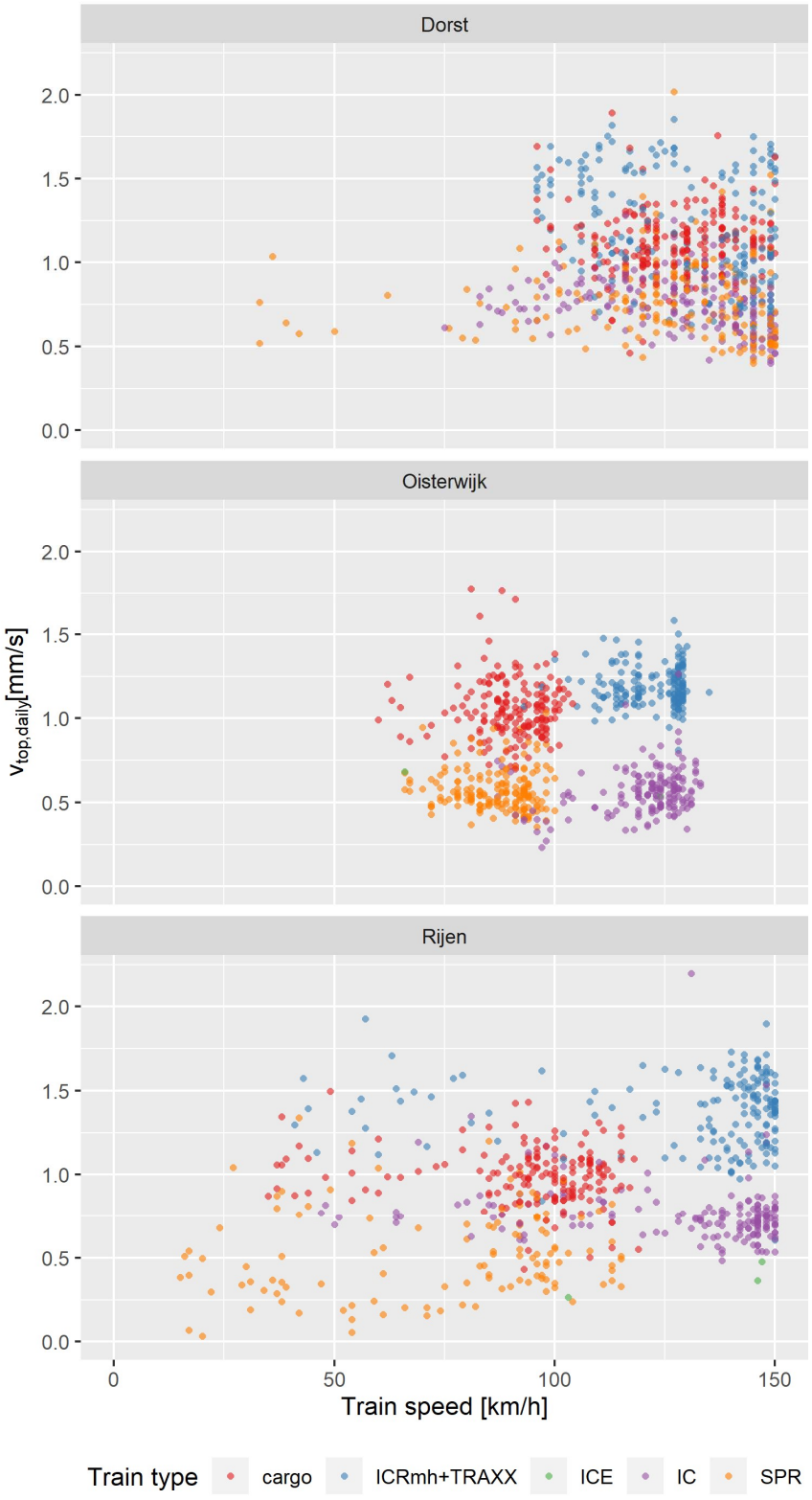


Figure F-5: Daily vibration velocity maxima and corresponding train speed per location and train type, i.e. each data point represents the maximum daily value for a particular location and particular train type. ICRmh+TRAXX: ICRmh Intercity passenger wagons and TRAXX engine; IC: InterCity; SPR: Sprinter; ICE: Intercity Express. Data: N-P-SM.

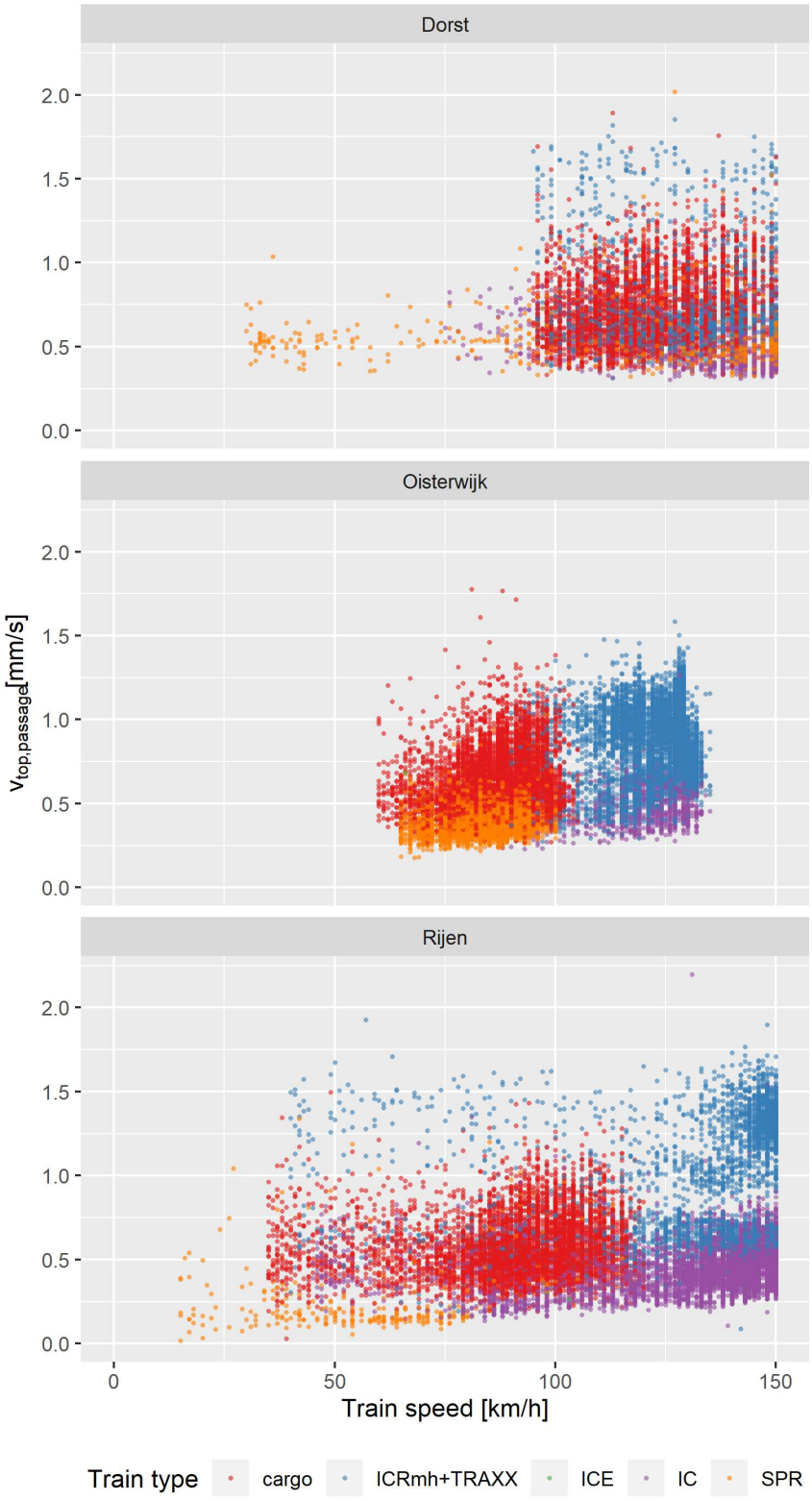


Figure F-6: Passage vibration velocity maxima and corresponding train speed per location and train type, i.e. each data point represents the maximum daily value for a particular location and particular train type. ICRmh+TRAXX: ICRmh Intercity passenger wagons and TRAXX engine; IC: InterCity; SPR: Sprinter; ICE: Intercity Express. Data: N-P-SM.

F.3 The impact of exposure period on peak velocity

In this section the impact of the vibration exposure period on the peak velocity is investigated. The N-P-SM database is large enough to investigate such an impact without too many modelling assumptions, i.e. let the data speak. For Dorst, Oisterwijk, and Rijen the total number of days with vibration measurements are 200, 185, and 171, respectively.
Assuming that there is no or little dependency between the daily maxima, the following approach is used:

1. First the days with no measurement data (see for example Figure F-1) are removed and the days with measurements are treated as consecutive days.
2. The measurement period is partitioned into equal length disjoint intervals (Figure F-7): block size (vibration exposure period). The intervals start from the left end of the total measurement period.
3. The maximum of each block is calculated and they form a sample. Statistics of the samples are calculated:
    a. Mean of block maxima and the estimation of its sampling uncertainty, the standard error (*se*) of the mean is calculated using: $se = s/\sqrt{n}$, where *s* is the sample standard deviation and *n* is the sample size. See Figure F-8 for the results.
    b. 1-P_{th} fractile ($v_{top,th,block}$) is calculated after fitting a generalized extreme value distribution (GEV) to the block maxima (Coles, 2001). Maximum likelihood method is used for the distribution fitting and the confidence interval of the fractile is estimated using the delta method (Dorfman, 1938). See Figure F-10 for the results.

Note that no long term trends were removed from the time series although there might be a seasonal trend present.

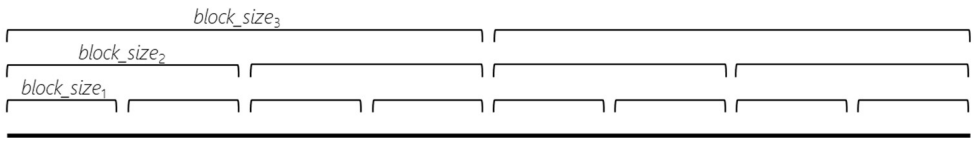


Figure F-7: Illustration of the partitioning of the measurement period into equal length disjoint intervals. The maximum is taken from each block.

The results presented in Figure F-8 show a clear increase mean peak velocity as the block size increase. The widening confidence interval is due to the decreasing sample size with increasing block size, i.e. smaller number of large blocks fit into the same measurement period. The jaggedness of the plots for larger block sizes is also caused by the small sample size (sampling variability). Moving from 1-day to 50-day maxima we observe an abut 30-40% increase in mean $v_{top,block}$ that indicates a substantial effect of the vibration exposure period. Note that this analysis does not make any modelling assumptions besides that the block maxima are assumed to be independent.

Since for threshold distances not the mean but the 1-P_{th} fractile (P_{th}=0.176) of $v_{top,block}$ is used, we also assessed how that changes with increasing the block size (vibration exposure period). The fractile can be estimated from sample only for

small block sizes when there is sufficient point to get an empirical estimate. In order to get an estimate for larger block sizes as well we fitted a GEV distribution to each block sample and calculated its 1-P_{th} fractile. The distribution fitting is done only when there are at least five data points in the sample. The results are summarized in Figure F-10. A similar trend can be observed as for the $v_{top,block}$ mean: the increase in $v_{top,th,block}$ is substantial as the block size (exposure period) is increasing.

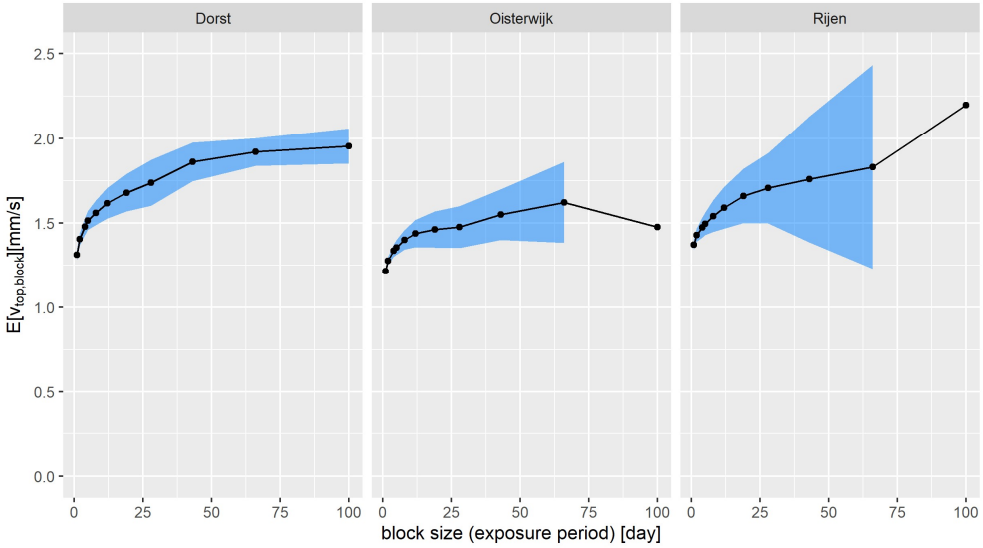


Figure F-8: The mean of block maxima for different block sizes (vibration exposure periods). The black line shows the estimated mean and the blue band shows an approximate 90% confidence interval. The lack of confidence intervals for some block sizes is due to the small sample size that does not allow for its estimation. Data: N-P-SM.

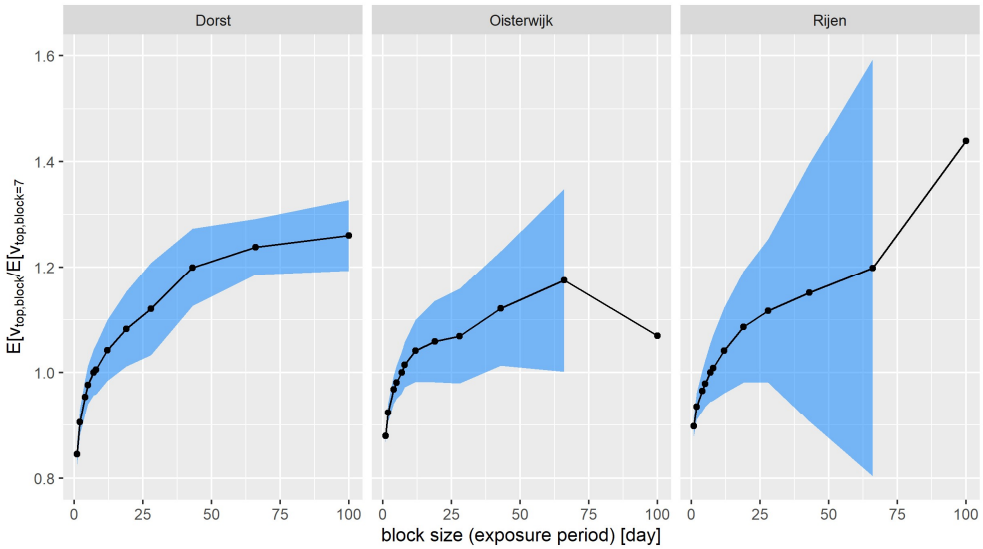


Figure F-9: The normalized mean of block maxima for different block sizes (vibration exposure periods). All values are normalized using the mean of 7-day block maxima for each location. The black line shows the estimated mean and the blue band shows an approximate 90% confidence interval. The lack of confidence intervals for some block

sizes is due to the small sample size that does not allow for its estimation. Data: N-P-SM.

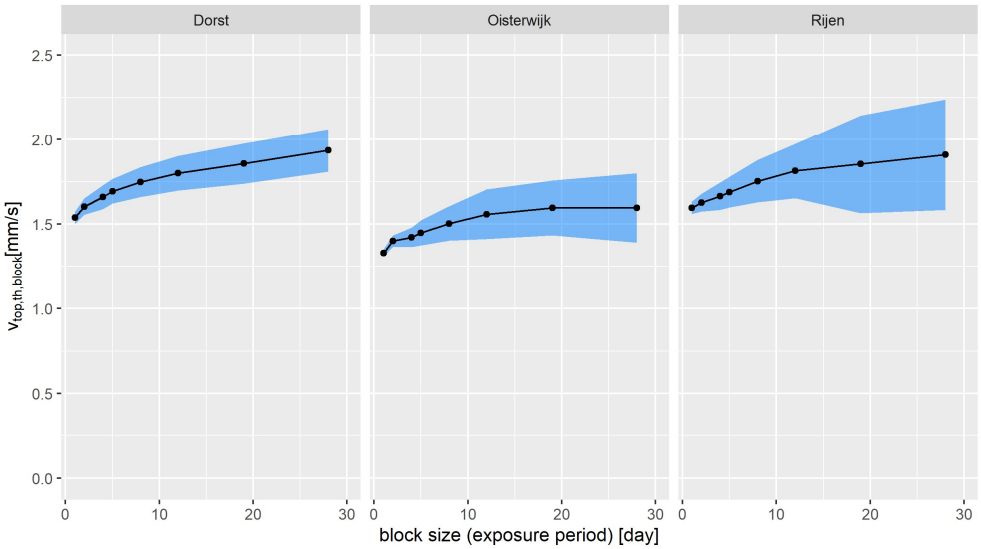


Figure F-10: 1-P_{th} fractile ($v_{top,th,block}$) for different block sizes (vibration exposure periods).). All values are normalized using $v_{top,th,block}$ of 7-day block maxima for each location. The black line shows the maximum likelihood point estimate and the blue band shows an approximate 90% confidence interval, both obtained from fitting a GEV distribution to the sample. We have shorter block sizes as in Figure F-8 as we fitted a GEV distribution if there were at least five data points. Data: N-P-SM.

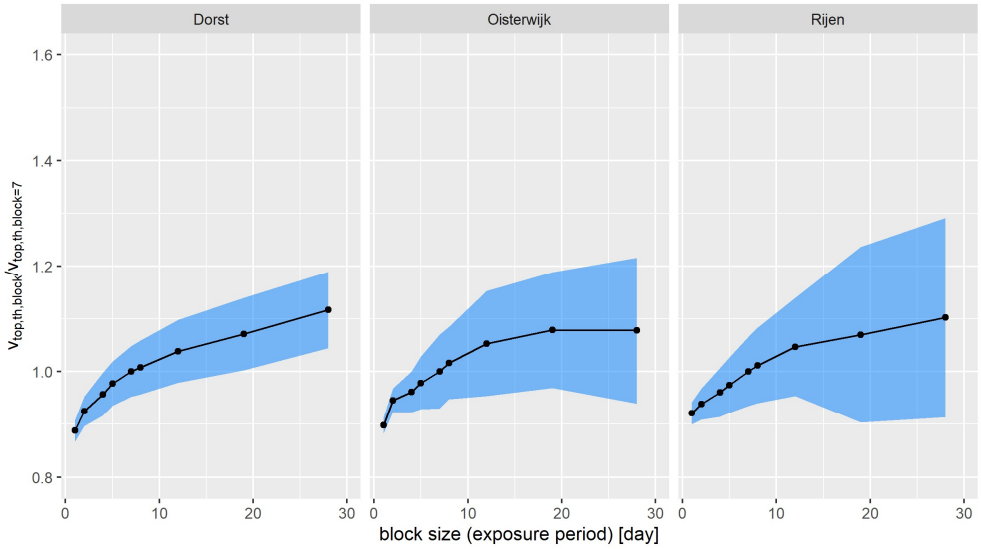


Figure F-11: Normalized 1-P_{th} fractile ($v_{top,th,block}$) for different block sizes (vibration exposure periods). The black line shows the maximum likelihood point estimate and the blue band shows an approximate 90% confidence interval, both obtained from fitting a GEV distribution to the sample. We have shorter block sizes as in Figure F-8 as we fitted a GEV distribution if there were at least five data points. Data: N-P-SM.

If we assume independent block maxima a single distribution fitted to the a sample of block maxima can already give information about the influence of the vibration

exposure period. GEV distribution is fitted to the 1-day, 3-day, and 7-day block maxima and plotted in Figure F-12, Figure F-13, and Figure F-14, respectively. Maximum likelihood estimate is used for the parameter estimation and the confidence interval of the fractiles estimated using the delta method; moreover, the observations and the models are presented in the Gumbel space typically used for visualization in extreme value analysis (Coles, 2001; Rózsás, 2016). The results show a similar trend as observed in Figure F-8 and Figure F-10.
Among the analysis presented in this section this ones relies the most on modelling assumptions Fitting a GEV to the logarithm of $v_{top,block}$ is results in very similar plots hence they are nor provided here.

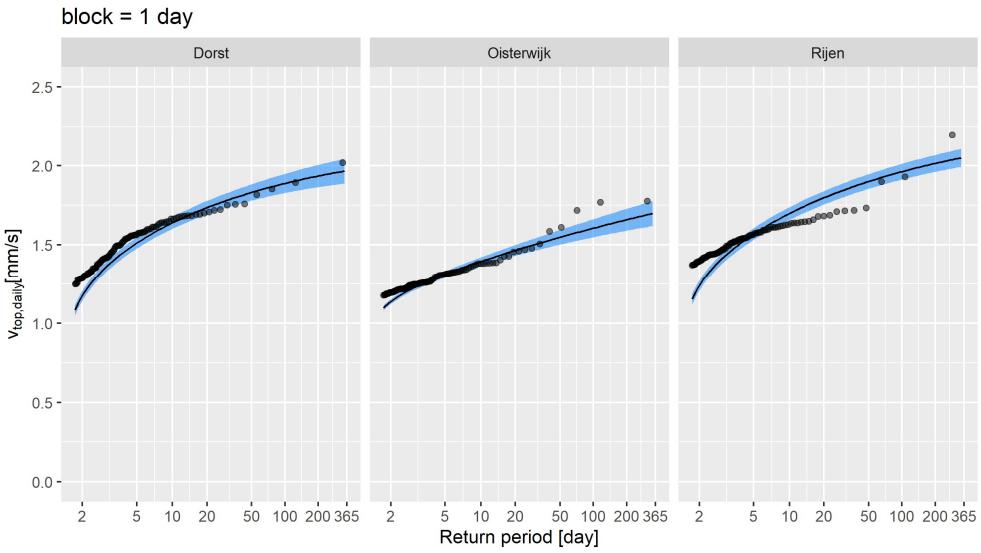


Figure F-12: Daily vibration maxima (circles) and fitted GEV distribution in Gumbel space. The black line indicates the maximum likelihood point estimate while the blue band corresponds to the 90% confidence interval (estimated using the delta method).Compare it with Figure F-13 and Figure F-14 Data: N-P-SM.

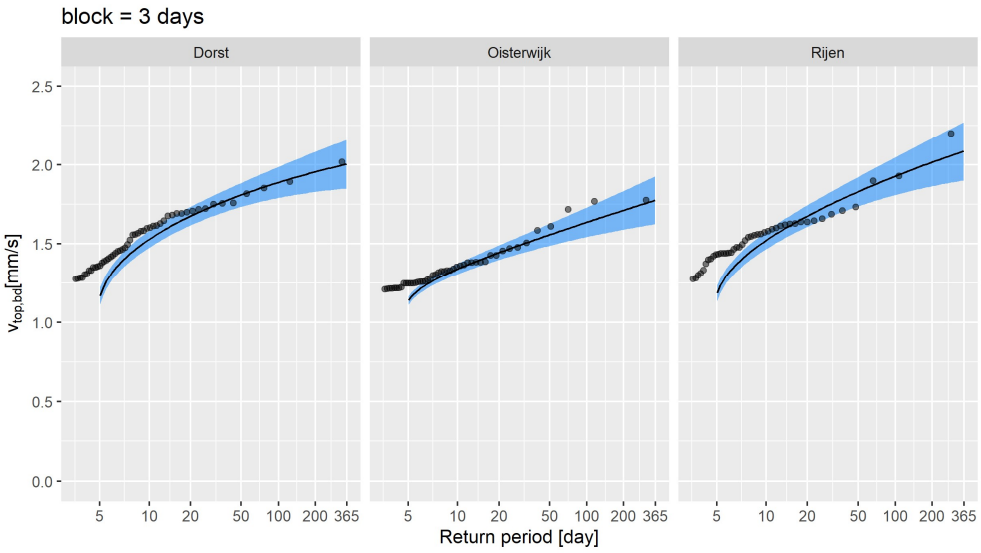


Figure F-13: 3-day vibration maxima (circles) and fitted GEV distribution in Gumbel space. The black line indicates the maximum likelihood point estimate while the blue band

corresponds to the 90% confidence interval (estimated using the delta method). Compare it with Figure F-12 and Figure F-14. Data: N-P-SM.

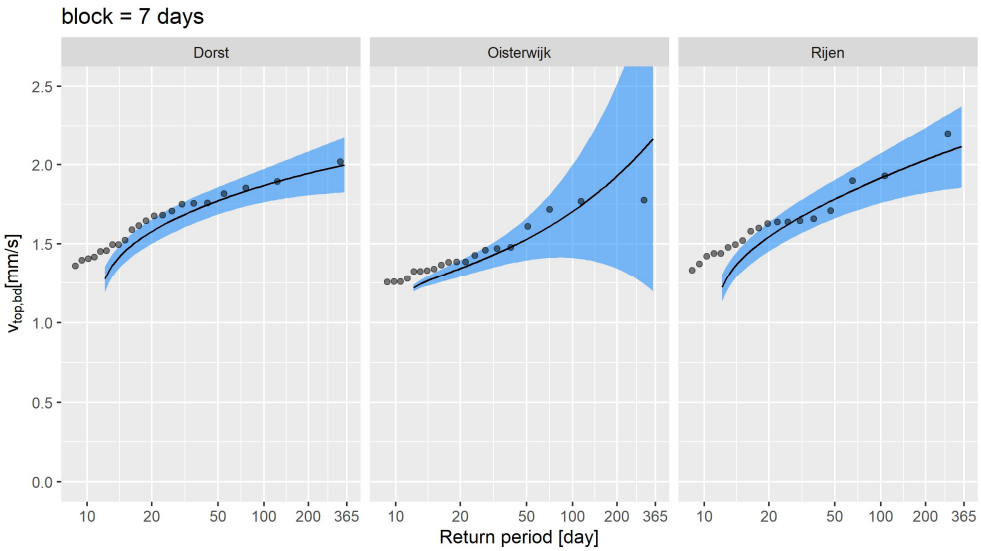


Figure F-14: 7-day vibration maxima (circles) and fitted GEV distribution in Gumbel space. The black line indicates the maximum likelihood point estimate while the blue band corresponds to the 90% confidence interval (estimated using the delta method). Compare it with Figure F-12 and Figure F-13. Data: N-P-SM.

Table F-2: Summary statistics of the daily vibration velocity maxima for three locations. Data: N-P-SM.

[mm/s] if applicable	Dorst	Oisterwijk	Rijen
Number of days with measurements	200	185	171
Mean of daily maxima	1.31	1.21	1.37
Standard deviation of daily maxima	0.255	0.137	0.232

G Peak(top) velocity attenuation functions

Some of the fitted models with their central values and uncertainty bands are presented in Figure G-1 and Figure G-2. The figures also show the measurement data used for the model fitting and indicate the goodness of the fit. The plots of the fitted models for all considered situations can be found in digital annex G.

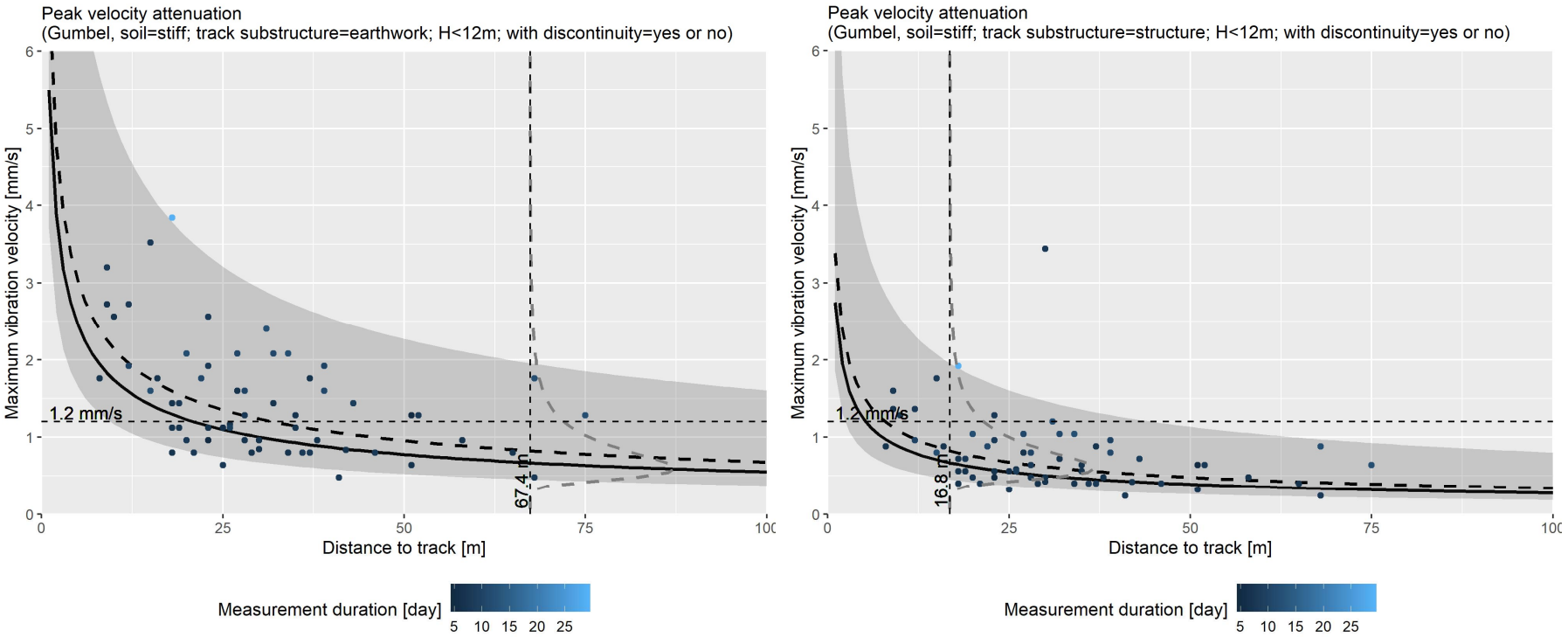


Figure G-1: Fitted 7-day peak(top) velocity attenuation curves and measurements for stiff soils. Solid black line: model mode; dashed black line: model mean; gray band: model 90% confidence interval; dashed gray line: probability density of the model prediction at the threshold distance; black dashed lines: threshold peak(top) velocity (v_{th}) and calculated threshold distance (R_{th}). Mind the inconsistency in the time duration of the attenuation model and the measured maxima which can distort the visual comparison. Data: N-M-SM.

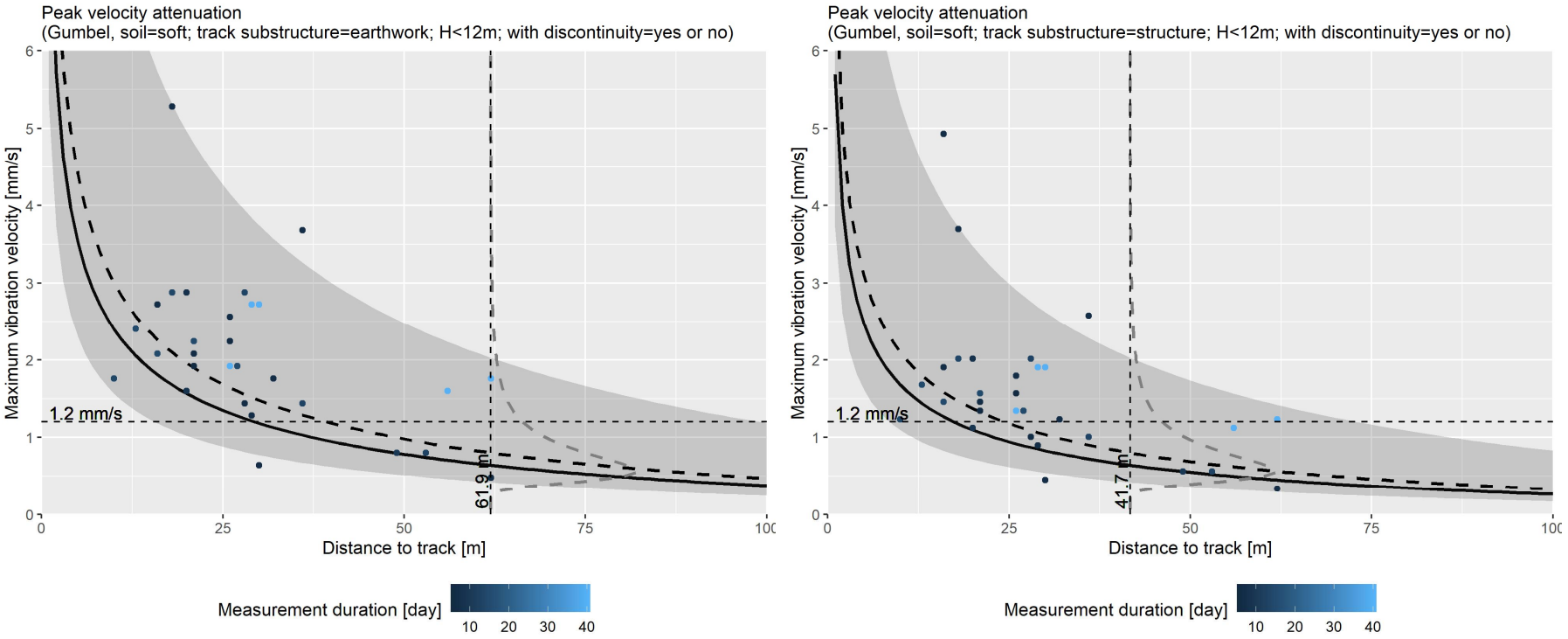


Figure G-2: Fitted 7-day peak velocity attenuation curves and measurements for soft soils. Solid black line: model mode; dashed black line: model mean; gray band: model 90% confidence interval; dashed gray line: probability density of the model prediction at the threshold distance; black dashed lines: threshold peak(top) velocity (v_{th}) and calculated threshold distance (R_{th}). Mind the inconsistency in the time duration of the attenuation model and the measured maxima which can distort the visual comparison. Data: N-M-SM.

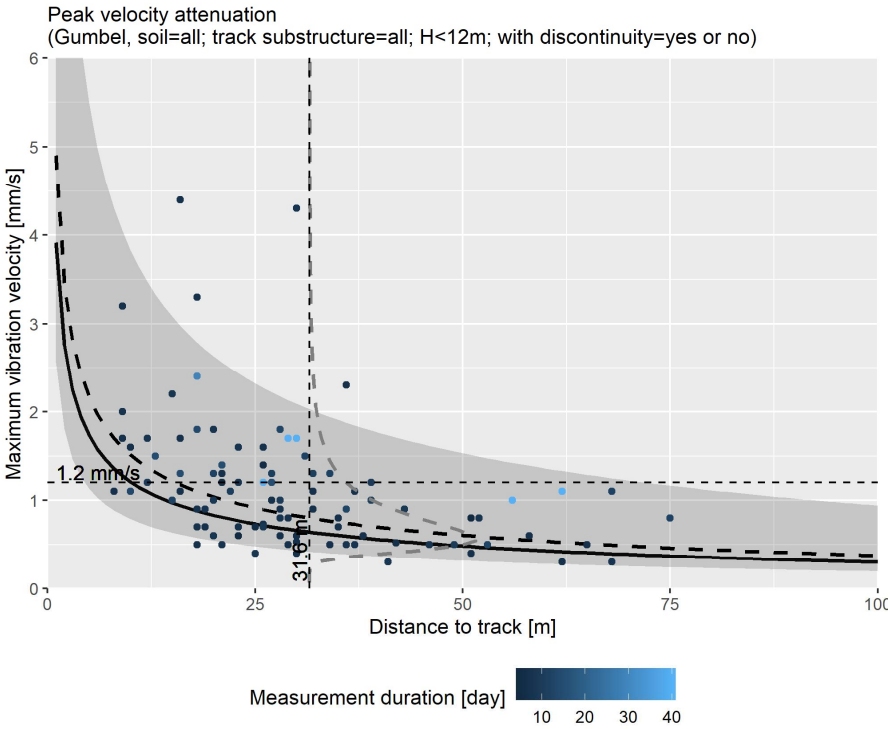


Figure G-3: Fitted 7-day peak(top) velocity attenuation curves and measurements for all soils and all structures. Solid black line: model mode; dashed black line: model mean; gray band: model 90% confidence interval; dashed gray line: probability density of the model prediction at the threshold distance; black dashed lines: threshold peak(top) velocity (v_{th}) and calculated threshold distance (R_{th}). Mind the inconsistency in the time duration of the attenuation model and the measured maxima which can distort the visual comparison. Data: N-M-SM.