Automated Detection of Long-Period Disturbances in Seismic Records; MouseTrap Code

by Jiří Vackář, Jan Burjánek, and Jiří Zahradník

INTRODUCTION

Sudden disturbances in strong-motion acceleration records, referred to as baseline offsets, are well known (Boore *et al.*, 2002; Javelaud *et al.*, 2011). The physical nature of this phenomenon remains unclear, although some cases are well explained by a permanent ground tilt (Graizer, 2010; Javelaud *et al.*, 2012). The ground tilt may be caused by near-field source effects or a local tilt at a station produced by seismic vibrations in highly heterogeneous substrata (i.e., strong gradients of material properties at small scales). Tilts are frequently observed in volcanic areas (Wielandt and Forbriger, 1999; Wiens *et al.*, 2005). Delorey *et al.* (2008) performed an experiment showing that tilting of the instrument produces a long-period disturbance of the characteristic shape.

In these cases, the disturbances predominate on horizontal components. The disturbances can also have a purely instrumental origin (Iwan *et al.*, 1985; Shakal and Petersen, 2001; Boore, 2003). The latter is a more suitable explanation in cases where the disturbances are also strong on the vertical component.

Similar artifacts have also been reported on different types of broadband seismometers worldwide (Zahradník and Plešinger, 2005; Pillet and Virieux, 2007; Delorey *et al.*, 2008). A strong disturbance is frequently characterized by a one-sided pulse in raw output velocity. A weaker disturbance is often masked by high-frequency content of the velocity record. However, it can be easily visible in the integrated output of a broadband instrument (raw displacement) in which it looks like a baseline step, the duration of which is equal to the seismometer corner period (Fig. 1a).

The disturbances may easily be overlooked in band-pass filtered records. Using such disturbed records can lead to wrong results in many seismic applications, for example, the moment tensor inversion (Zahradník and Plešinger, 2005; Zahradník *et al.*, 2008). Similar problems might arise in the spectral estimation of the moment magnitude; the disturbed records may erroneously pass through a routine data quality control, for example the signal-to-noise ratio (SNR) control (Sokos and Zahradník, 2013), because the disturbances yield apparently very large values of the SNR. As seen in Figure 1b,

the disturbances cause significant spurious increase of the spectrum at low frequencies. The signal spectrum is contaminated up to the frequencies much higher than the low-frequency corner of the instrument.

The main objective of this study was to develop a code for an automated detection of such disturbances aimed at removing them from the routine seismic data processing. Such a tool would also allow for systematic studies of the disturbances to identify and explain their origins. As an application, we analyzed the disturbances in 18 years of recordings in the Swiss Digital Seismic Network.

The developed code is available as an ObsPy module (under the GNU Lesser General Public License [LGPL]).

## MODELING THE DISTURBANCES: FORWARD PROBLEM

Many of the observed disturbances can be modeled as a seismometer response to an acceleration step on the input (Zahradník and Plešinger, 2005) (Fig. 1a). Thus, a characteristic one-sided pulse appears in the raw velocity output, which results in a baseline offset in the raw displacement (Fig. 2). The step visible in the raw displacement is not a permanent displacement but a permanent acceleration (see the acceleration scale at the right side of Fig. 2d). It is because the integrated broadband output is proportional to displacement only at frequencies above the low-frequency corner of the instrument, whereas it is proportional to acceleration at the low-frequency limit. Consequently, a simple integration of the instrumentally uncorrected record in its full frequency band will reveal the characteristic shape of the disturbance (smoothed ramp, see Fig. 1a).

Because of its shape, we name the disturbance an "artificial fling step" (for brevity also "fling step" in the following text). Such disturbances are well known to users of moment tensor inversion software ISOLA (Sokos and Zahradník, 2008, 2013), because removal of disturbed records is necessary. Informally, in the community of ISOLA users, such disturbances are referred to as "mouse." We use this informal name in the developed code.

The time series $m_d(t)$ of the raw displacement disturbance caused by a unit acceleration step with zero onset time is described by

$$m_d(t) = \int_{-\infty}^{t} \int_{-\infty}^{+\infty} s(\tau)\tau h(t - \tau)\mathrm{d}\tau\mathrm{d}t, \qquad (1)$$

in which $h(t)$ is an impulse response of the instrument to input velocity and $s(t)$ is the input unit acceleration step, that is the
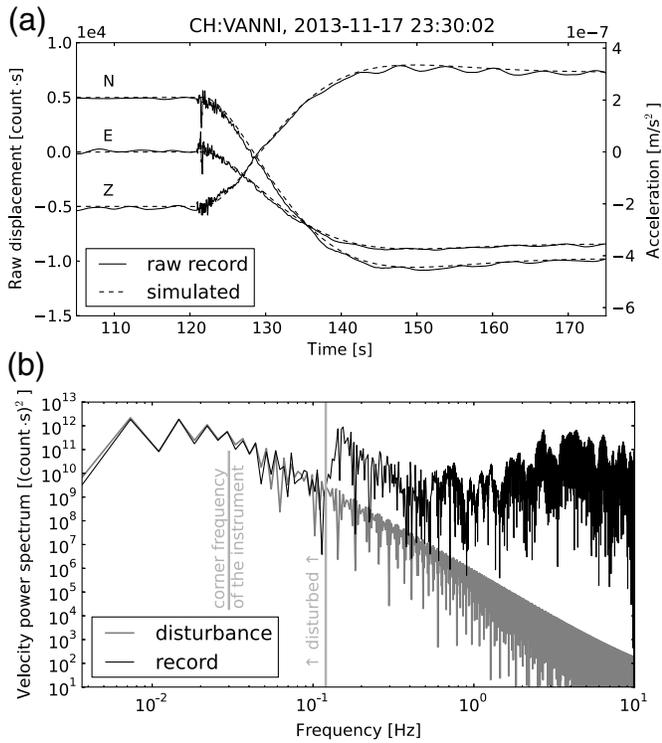
▲ Figure 1. (a) Example of a steplike disturbance in the integrated output (raw displacement) of the Nanometrics Trillium T40 seismometer. The disturbance is well fitted by the simulated instrument response to an acceleration step of amplitude $A = 8.7 \times 10^{-7}$ m·s⁻², azimuth $\phi = 209.7°$, and inclination $\theta = 35.9°$. The recorded earthquake has magnitude $M_{Lh} = 1.7$ (Swiss Seismological Service), and its epicentral distance is 1.9 km. (b) Power spectrum of raw velocity (Z component) and of the disturbance. The disturbance dominates in the spectrum up to frequency ~0.12 Hz.

▲ Figure 2. Modeling a disturbance. The (a) input acceleration step is equivalent to the (b) velocity ramp. After applying the instrumental response (here for Nanometrics Trillium T40), we get a characteristic one-sided pulse in the (c) output raw velocity, and a steplike disturbance in the (d) raw displacement. The right scale of (d) corresponds to the low-frequency limit of integrated broadband output, which is proportional to acceleration.

Heaviside function ($s(t) = 1$ for $t \geq 1$ and $s(t) = 0$ for $t < 0$). The inner integral is a time convolution of the input velocity ramp $s(\tau)\tau$ and instrument response $h(\tau)$. The outer integral transforms the velocity to displacement.
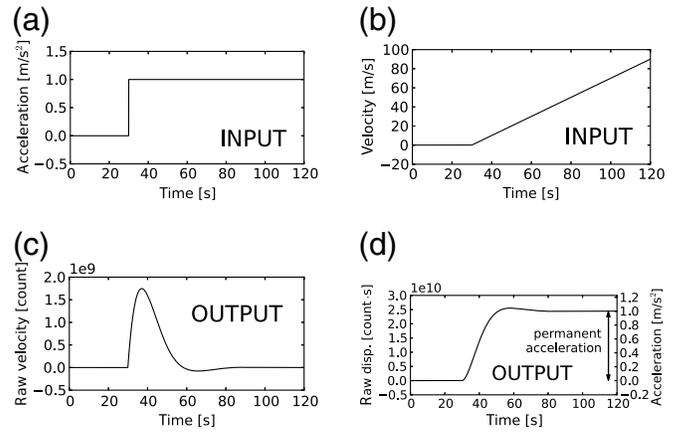
The disturbances often occur simultaneously on three-component recordings; in another words, the causative acceleration step has three components as well. Thus, we use four parameters for the description of a real fling step: time t_0 of the onset of the input acceleration step, amplitude A of the acceleration step and two spatial angles, horizontal azimuth ϕ, and its inclination θ from the horizontal plane. The east (E), north (N), and vertical (Z) components of the raw displacement disturbance recording are therefore

$$m_E(t) = m_d(t - t_0)A \sin \phi \cos \theta, \qquad (2)$$

$$m_N(t) = m_d(t - t_0)A \cos \phi \cos \theta, \qquad (3)$$

and

$$m_Z(t) = m_d(t - t_0)A \sin \theta. \qquad (4)$$

FITTING THE DISTURBANCES: INVERSE PROBLEM

Explaining a recorded disturbance by the synthetic fling step formally means solving an inverse problem with four parameters: t_0, A, ϕ, and θ. To have a unique solution, A must be positive and the angles ϕ and θ are in the intervals 0°–360° and −90° to +90°, respectively. The inverse problem is solved by the least-squares method (LSQ) to minimize the L2-norm difference between the observed record and synthetic fling step. The agreement is quantified by variance reduction (VR). Analytical expressions of the partial derivatives with respect to A, ϕ, and θ are used in the LSQ fitting (see Appendix for details). The values of t_0 are selected from a grid of values, and the inverse problem is solved for the other three parameters in each grid point; then the corresponding parameters for t_0 with the highest VR are selected.

Besides VR, we also need to define a criterion—whether the fling step is present or absent. To this goal, we evaluate an existence criterion

$$mp = VR - \frac{|t_0 - t_S|}{\text{const}}, \qquad (5)$$

in which t_S is the S-wave arrival. The term $|t_0 - t_S|/\text{const}$ serves to penalize steplike disturbances not related to the earthquake to exclude false detection due to noise. The penalty term was motivated by the observation that most fling steps occur shortly after the S-wave arrival ($t_0 \geq t_S$). The term "const" is a numerical constant chosen to be much larger than $t_0 - t_S$ in case of earthquake-induced fling step (const = 50 s in application of this article, chosen *ad hoc* after few preliminary tests). If the const is too low, the earthquake-induced fling steps might be excluded; if it is too high, many cases in which

the long-period noise is fitted by synthetic fling step, such as in Figure 3d, are taken into account. After visually inspecting dozens of disturbances with various mp values, we set up an empirical criterion that the cases with $mp > 0.7$ are fling steps, $mp < 0.2$ are not fling steps, and cases $0.2 < mp < 0.7$ require visual inspection to distinguish whether the fling step is present or not. There are examples of records and fitted synthetic disturbances in Figure 3. The record of station SIMPL (Fig. 3a, $mp = 0.75$) is a typical example of a strong disturbance. Record of MRGE (Fig. 3b, $mp = 0.63$) is unclear: when we look at the record in detail, the baseline trend starts seconds before the earthquake and its shape differs from the synthetic disturbance. At station BERNI (Fig. 3c, $mp = 0.28$), the disturbance might be present, although it is weak and the shape is not fitted perfectly. In record of ZUR (Fig. 3d, $mp = 0.12$), there is a relatively well-fitted long-period noise but with relation neither to an earthquake nor to the studied disturbance.

MOUSETRAP CODE

We implemented the described approach into an ObsPy module called MouseTrap. The module consists of the following functions:

mouse.create. Calculates the instrument output for a given transfer function and a unit acceleration step on the input.

PrepareRecord. Removes the before-event mean value, integrates record into displacement, and analyzes the SNR of the record (ratio of before-event maximum to record maximum). Testing SNR is necessary to prevent explaining strong long-period noise as a disturbance.

mouse.fit_3D. Fits a given three-component record with a synthetic fling step.

mouse.exist. Distinguishes whether the fling step is present according to the fit value, synthetic fling step amplitude, and its ratio to record amplitude. The output can be optionally the binary value (0, 1) or the mp value.

mouse.params. Returns the parameters t_0, A, ϕ, and θ of the detected fling step.

mouse.plot. Plots the comparison between observed record and synthetic fling step.

The module is published under GNU/GPL license and is available at http://geo.mff.cuni.cz/~vackar/mouse (last accessed January 2015). The detailed documentation and a few examples are also on the website. Running the code requires Python interpreter (version 2.7.x) with ObsPy (Beyreuther *et al.*, 2010) and Matplotlib (Hunter, 2007) libraries installed. The module can be used for two tasks: the fling step detection in a single record or the detection of fling steps in a set of records of a seismic network. The latter requires earthquake catalog in machine-readable or database form.
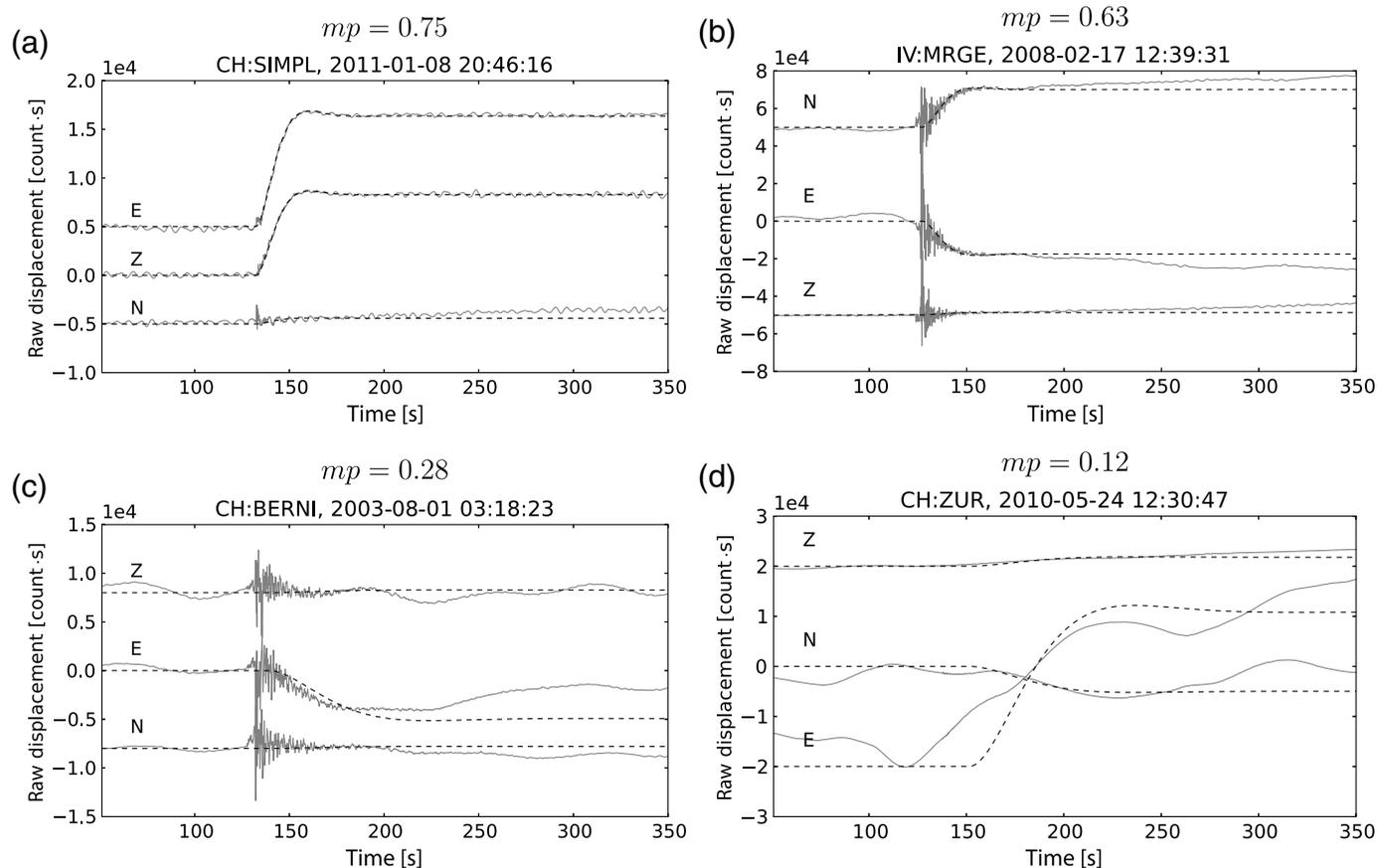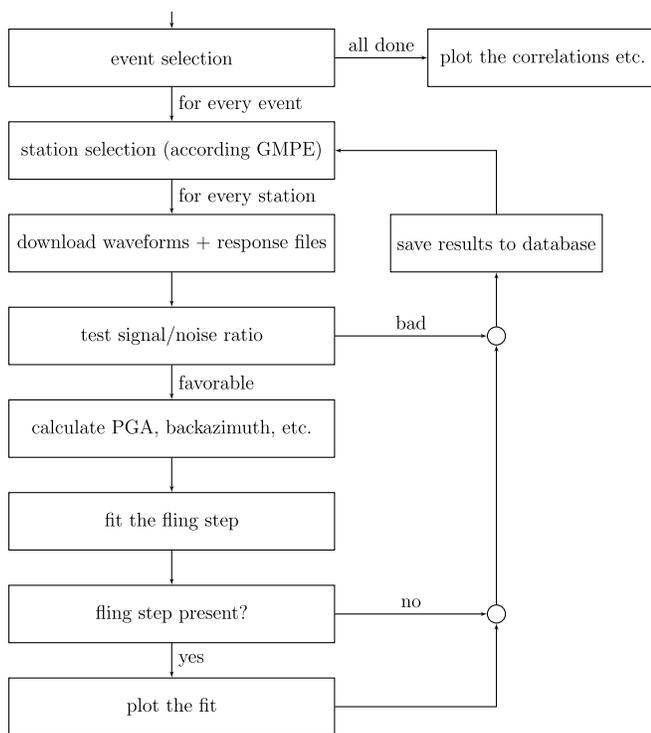

▲ **Figure 3.** Agreement between observed records (solid line) and fitted synthetic disturbances (dashed line) for different values of the existence criterion mp.

AUTOMATED DETECTION OF DISTURBANCES IN A SEISMIC NETWORK

Here, we demonstrate the code capabilities through a systematic detection of fling steps using a database of seismic records. The source code of this procedure (SwissMouse) can be found at the website mentioned above, and the results of its application to SWISS DIGITAL SEISMIC NETWORK are described in the following section. The algorithm flowchart is shown in Figure 4.

The procedure starts with event selection from the database. For each earthquake, a set of stations is selected in the distance range according to magnitude and the Swiss ground-motion prediction equation (GMPE) (Edwards and Fäh, 2013; Cauzzi et al., 2014). We take into account the stations in which the exceedance of a fixed acceleration value is expected. As *a posteriori* verification, we checked that only a few detected disturbances lie next to the cutoff line of the criterion. Waveform data and poles and zeros are downloaded from the ArcLink server. The SNR (ratio of before-event maximum to record maximum) is tested, and the unfavorable records are skipped. Simple parameters of the records, such as epicentral distance and azimuth, as well as the PGA, PGV, and PGD values are evaluated. In addition to standard PGA and PGV values, these values were also calculated in different frequency bands, that is, the record was band-pass filtered and then maximal values determined. Next, the fling step modeling is applied using the MouseTrap code, and the inverted parameters of the fling step are stored in an SQL database. The entire process is fully automated, but the comparison between observed and synthetic waveform is also plotted for possible visual inspection. Last but not least, correlation plots and histograms, which are discussed in the following chapter, are automatically plotted using predefined SQL queries.

CASE STUDY: DISTURBANCE DETECTION IN 18 YEARS OF RECORD OF SWISS DIGITAL SEISMIC NETWORK

We examined 18 years of records of Swiss Digital Seismic Network using the developed MouseTrap and SwissMouse codes. The purpose was to test the detection algorithm and obtain a set of fling steps for a rough evaluation of their properties.

The magnitude–distance criteria were met by 6240 broadband records, but only 1126 records were passed to further analysis (see Table 1 for details). From the analyzed records, in 83 cases a fling step was detected with no doubt ($mp > 0.7$). Most of the detected disturbances are in records in which the seismic signal is significantly below the saturation level. Another 129 cases were unclear ($0.7 > mp > 0.2$), and a visual inspection was necessary for decision about the fling step existence.

The main properties of fling steps in our dataset can be briefly summarized as follows: Some directions of the fling step are more frequent. Namely, many of fling steps were observed in azimuth ϕ equal to 30°, 270°, and 150°, or the opposite directions (Fig. 5a). These azimuths coincide with directions of the three inclined pendulums in Galperin's design seismographs (like Streckeisen STS-2 and Nanometrics Trillium T40, which are the most frequent broadband instruments in the Swiss Digital Seismic Network). Most of the disturbances are horizontal ($\theta = 0$), but a significant number have the inclination θ close to $\pm 35°$ (Fig. 5b). The inclination of the pendulums in the Galperin's seismographs is 35.3°. Therefore, these observations might indicate the instrumental origin of the disturbances. Looking at individual stations, there are stations where most of the fling steps are horizontal but correlated neither with the azimuth of the pendulum nor with the back azimuth of the event, and there are some other stations where

▲ **Figure 4.** Flowchart of the systematic detection of fling steps with database of records using SwissMouse code.

Table 1
Statistics of the Analyzed Records

Records total	6240
Skipped records	5114
Problems with download	1087
Unfavorable signal-to-noise ratio	3965
Gap in data	42
No poles-and-zeros file	20
Analyzed records	1126
No fling step detected	914
Fling step is unclear (visual inspection necessary)	129
Fling step detected with no doubt	83

The records were analyzed by code SwissMouse while it was applied on 18-year set of Swiss data.
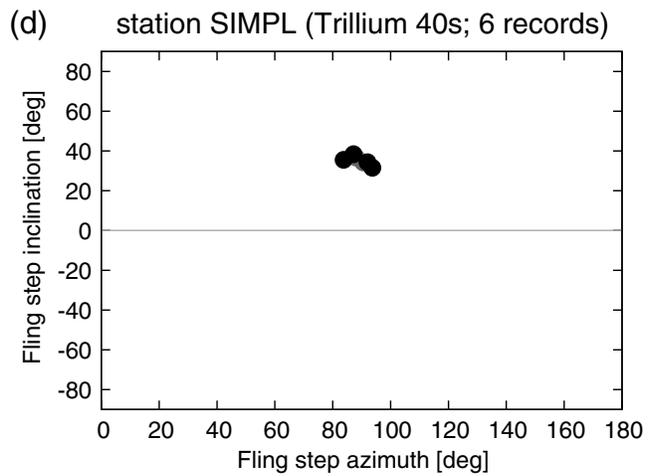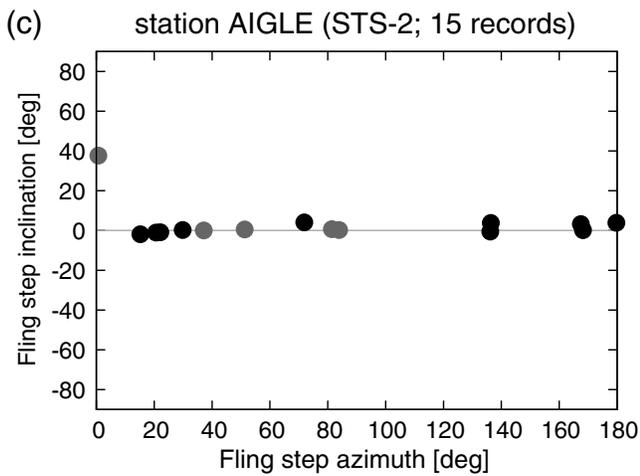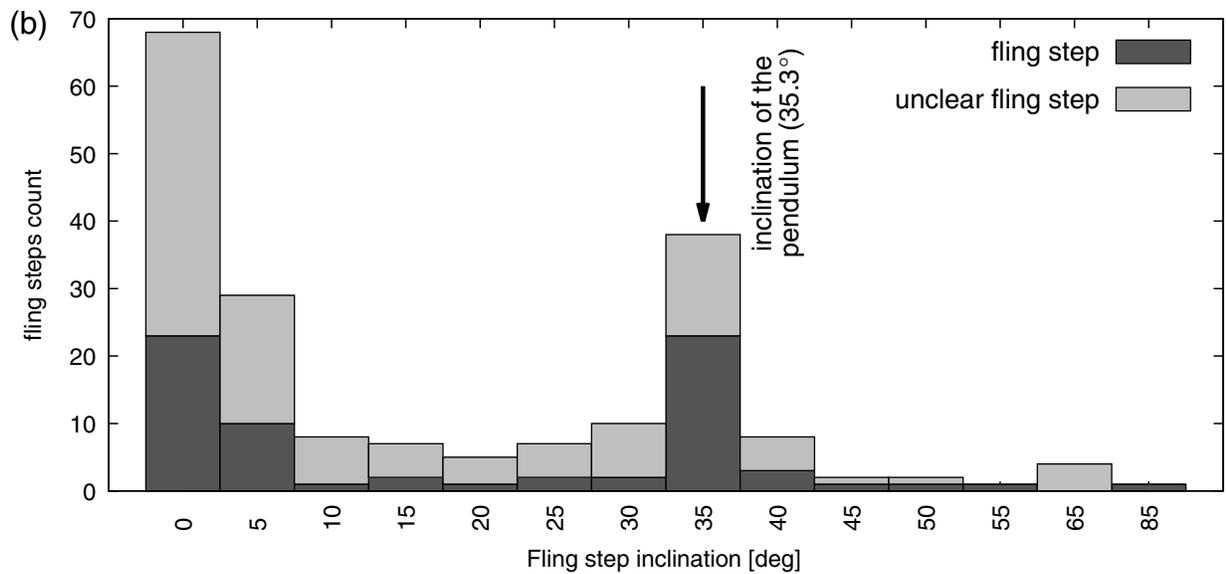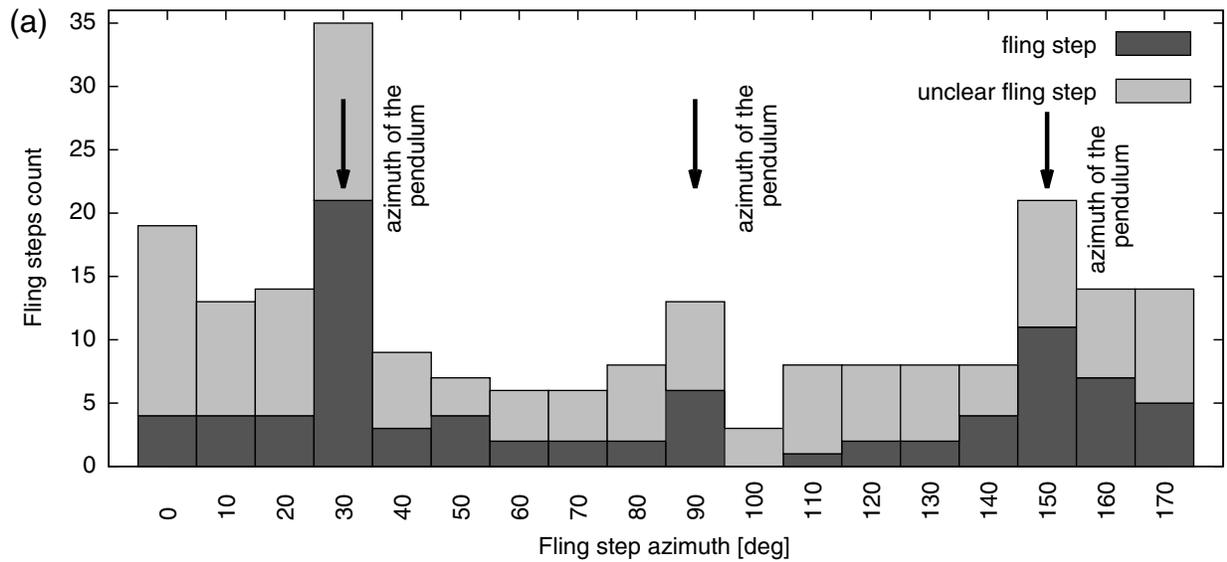
▲ **Figure 5.** (a) Azimuths and (b) inclination of the steplike disturbances in the 18-year dataset of disturbances detected in the Swiss Seismic Network. For better visibility, the azimuth is restricted to the interval 0°–180° and inclination to 0°–90° because some of disturbances are in the direction of the pendulum and some in the opposite direction. There are (c) some stations where the fling steps are nearly horizontal and in different azimuths, and (d) some other stations where all the fling steps have azimuth and inclination of one preferred pendulum.

most of the fling steps are in the azimuth and inclination of one preferred pendulum (Fig. 5c,d).

Fling steps are more common near the epicenter and at higher magnitudes (Fig. 6a). This is probably related to stronger ground shaking but does not correlate well with PGA, PGV, or PGD. Beside these parameters, the fling step/no fling step ratio is also dependent on epicentral distance (Fig. 6b).

We also evaluated the fling step/no fling step ratio at different seismograph types (Fig. 7a). The ratio is similar for all three types of the broadband instruments used in Swiss Digital Seismic Network. There are minor differences between broadband instruments, but this result is partly influenced by the presence or absence of events near to the investigated stations.

When we looked at fling step count at the individual investigated stations (Fig. 7b), a large variability was found, ranging from stations very rich in fling steps to stations with no fling step at all. Some of the stations where fling steps are common are situated very close to some earthquakes, so fling steps are mostly caused by near events (e.g., stations VANNI,

FUORN). But there are also stations (e.g., SIMPL, AIGLE), where the fling step occurrence is highly independent of the epicentral distance of the event (Fig. 7c,d).

Finally, the disturbance occurrences in Swiss network have no simple rules and can be hardly predicted. There are stations, where the disturbance occurrence is much more probable, which might indicate some instrumental problem. Nevertheless, we demonstrated their systematic existence especially at stations within short epicentral distances. Such stations are crucial in moment tensor inversions of weak events. The tool developed here is capable to detect and characterize such disturbances, so the disturbed records could be easily removed from processing without considering their origin.

CONCLUSION AND DISCUSSION

We developed a code MouseTrap for automated detection of fling step disturbances in seismic records. The program is available under free license at website http://geo.mff.cuni.cz/~vackar/mouse (last accessed January 2015). The spurious input acceleration step and the instrument response explain many of the observed disturbances very well. Fitting the synthetic disturbance into real records provides four parameters of the input acceleration step, namely its onset time t_0, amplitude A, azimuth ϕ, and inclination θ. The code can be applied either to an individual record or to a set of records with metadata in a database.

We expect the code to be useful in many applications for automatic data processing (e.g., waveform inversion and SNR evaluation) in which detection and removal of contaminated records is a must.

We also developed the code SwissMouse for automated analysis of fling step existence in broadband records of Swiss Digital Seismic Network over the last 18 years, at stations close to located events. Fling steps are present at all types of studied broadband instruments, at many different stations. We observed a higher percentage of fling steps at some stations. Azimuths of fling steps remain the same at some (but not all) stations. In particular, there is a higher occurrence of fling steps with azimuths 30°, 90°, and 150°. Most of the fling steps are horizontal, but a significant number have inclination ~35°. These preferred directions are very likely related to pendulums in Galperin's design seismometers, so these cases are probably of instrumental origin. There is a good reason (supported by limited authors' experience) to expect, that in instruments operating with pendulums in N, E, Z setup, such as Güralp CMG-3T, the disturbances might sometimes be preferentially related to a single component. The fling steps are more common at records with higher PGA and PGV, near the source, and at higher magnitudes (M_w 1–4 mostly examined). Besides this article, we also observed such disturbances at short-period instruments, where they are naturally characterized by much shorter durations than at broadband seismometers.

There are also similar disturbances that can be described as instrument response to a spurious step in the input velocity (Zahradník and Plešinger, 2010), which can be explained as caused by saturation in the force-balance system. These might
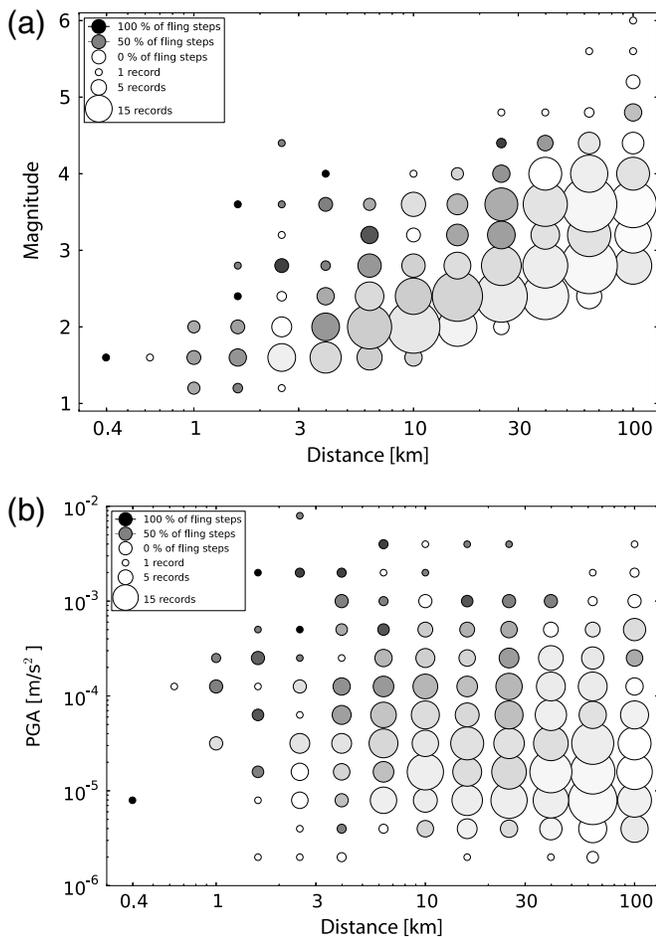
▲ **Figure 6.** (a) Fling steps are more common near the epicenter and at higher magnitudes. (b) Fling steps are more common at higher PGA. The magnitude (a) comes from Swiss Seismological Service earthquake catalog, moment magnitude M_L, and local magnitudes M_L and M_{Lh} are used for different events.
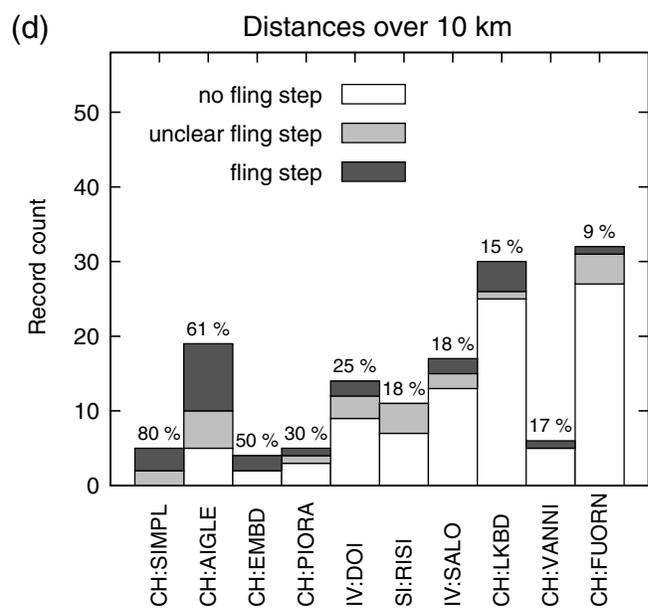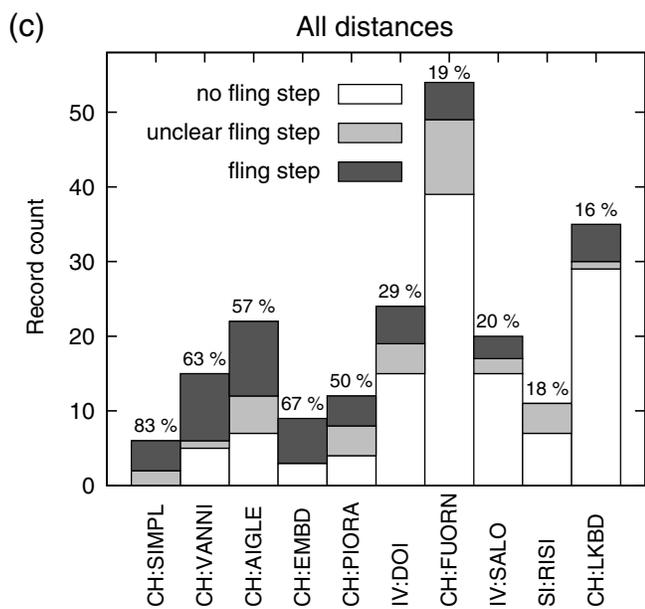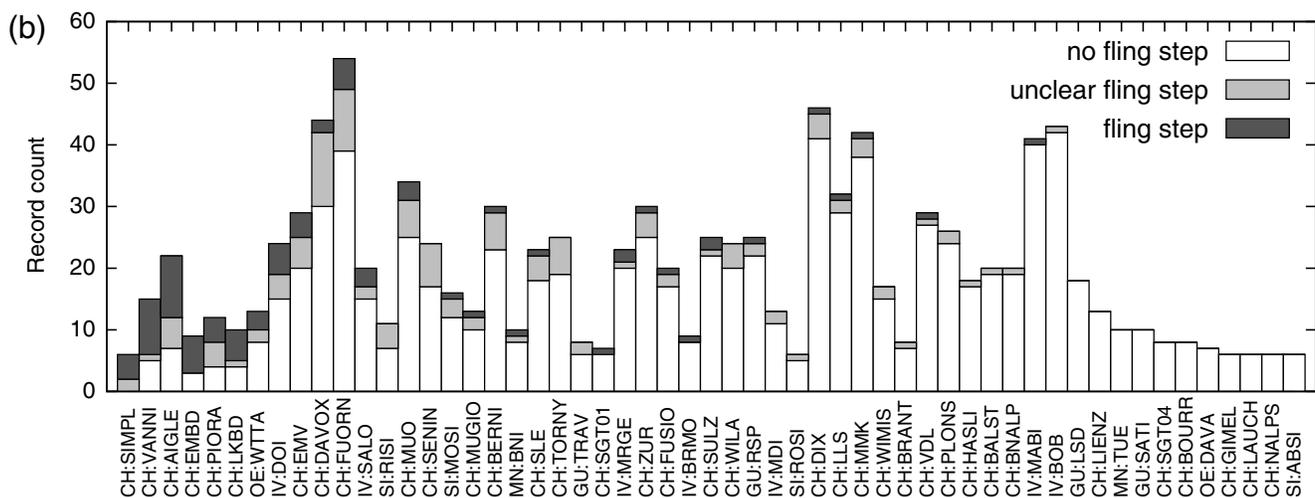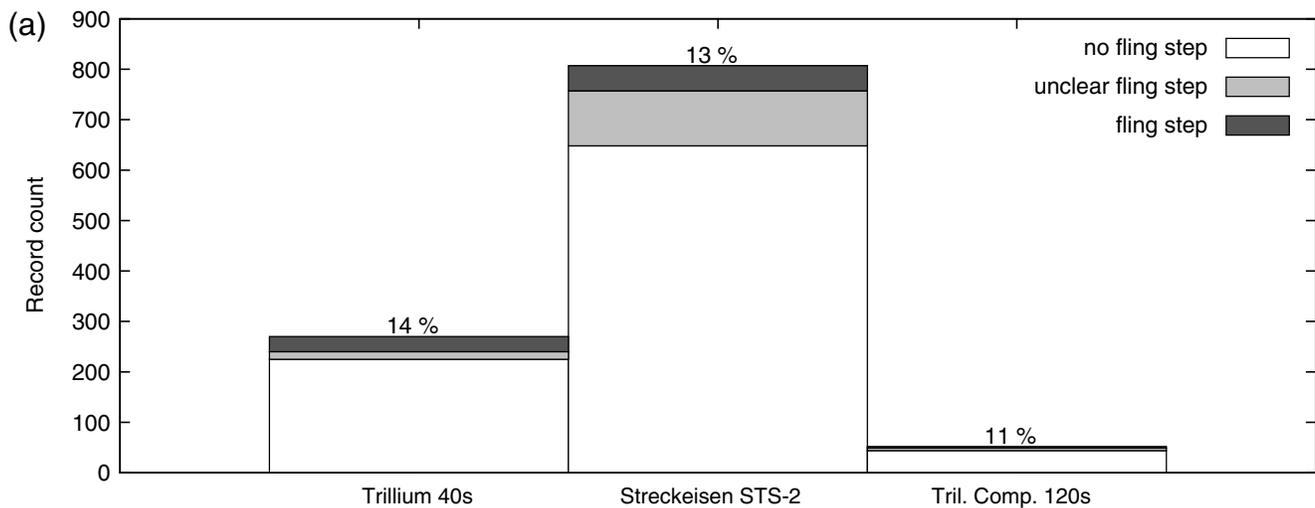
▲ **Figure 7.** (a) Fling steps count at studied broadband sensors. (b) Fling steps counts at the studied stations with more than five records. (c,d) Fling step count at selected stations for events in all epicentral distances and for epicentral distances over 10 km.

be implemented in future updates of the MouseTrap code, including diagnostics of a joint occurrence of both kinds of disturbances. Another challenging issue is to develop the fling step detection in continuous records. ⌇

ACKNOWLEDGMENTS

REFERENCES

Beyreuther, M., R. Barsch, L. Krischer, T. Megies, Y. Behr, and J. Wassermann (2010). ObsPy: A Python toolbox for seismology, *Seismol. Res. Lett.* **81,** 530–533.

Boore, D. M. (2003). Analog-to-digital conversion as a source of drifts in displacements derived from digital recordings of ground acceleration, *Bull. Seismol. Soc. Am.* **93,** 2017–2024.

Boore, D. M., Ch. D. Stephens, and W. B. Joyner (2002). Comments on baseline correction of digital strong-motion data: Examples from the 1999 Hector Mine, California, earthquake, *Bull. Seismol. Soc. Am.* **92,** 1543–1560.

Cauzzi, C., B. Edwards, D. Fäh, J. Clinton, S. Wiemer, P. Kästli, G. Cua, and D. Giardini (2014). New predictive equations and site amplification estimates for the next-generation swiss shakemaps, *Geophys. J. Int.* **200,** no. 1, 421–438.

Delorey, A. A., J. Vidale, J. Steim, and P. Bodin (2008). Broadband sensor nonlinearity during moderate shaking, *Bull. Seismol. Soc. Am.* **98,** 1595–1601.

Edwards, B., and D. Fäh (2013). A stochastic ground-motion model for Switzerland, *Bull. Seismol. Soc. Am.* **103,** 78–98.

Graizer, V. (2010). Strong motion recordings and residual displacements: What are we actually recording in strong motion seismology? *Seismol. Res. Lett.* **81,** 635–639.

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment, *Comput. Sci. Eng.* **9,** 90–95.

Iwan, W. D., M. A. Moser, and C. Y. Peng (1985). Some observations on strong-motion earthquake measurement using a digital accelerograph, *Bull. Seismol. Soc. Am.* **75,** 1225–1246.

Javelaud, E. H., T. Ohmachi, and S. Inoue (2011). A quantitative approach for estimating coseismic displacements in the near field from strong-motion accelerographs, *Bull. Seismol. Soc. Am.* **101,** 1182–1198.

Javelaud, E. H., T. Ohmachi, and S. Inoue (2012). Estimating small permanent rotation from strong-motion records: What is comparison with external measurements telling us? *Bull. Seismol. Soc. Am.* **102,** 2257–2263.

Pillet, R., and J. Virieux (2007). The effects of seismic rotations on inertial sensors, *Geophys. J. Int.* **171,** 1314–1323.

Shakal, A. F., and C. D. Petersen (2001). Acceleration offsets in some FBA's during earthquake shaking (abstract), *Seismol. Res. Lett.* **72,** 233.

Sokos, E. N., and J. Zahradník (2008). ISOLA a Fortran code and a Matlab GUI to perform multiple-point source inversion of seismic data, *Comput. Geosci.* **34,** 967–977.

Sokos, E. N., and J. Zahradník (2013). Evaluating centroid-moment-tensor uncertainty in the new version of ISOLA software, *Seismol. Res. Lett.* **84,** 656–665.

Wielandt, E., and T. Forbriger (1999). Near-field seismic displacement and tilt associated with the explosive activity of Stromboli, *Ann. Geofisc.* **42,** 407–416.

Wiens, D. A., S. H. Pozgay, P. J. Shore, A. W. Sauter, and R. A. White (2005). Tilt recorded by a portable broadband seismograph: The 2003 eruption of Anatahan Volcano, Mariana Islands, *Geophys. Res. Lett.* **32,** L18305, doi: 10.1029/2005GL023369.

Zahradník, J., and A. Plešinger (2005). Long-period pulses in broadband records of near earthquakes, *Bull. Seismol. Soc. Am.* **95,** 1928–1939.

Zahradník, J., and A. Plešinger (2010). Toward understanding subtle instrumentation effects associated with weak seismic events in the near field, *Bull. Seismol. Soc. Am.* **100,** 59–73.

Zahradník, J., J. Janský, and V. Plicka (2008). Detailed waveform inversion for moment tensors of M ∼ 4 events; examples from the Corinth Gulf, Greece, *Bull. Seismol. Soc. Am.* **98,** 2756–2771.

APPENDIX

Analytical expressions of partial derivatives are used during least-squares fitting of observed record by synthetic fling step.

Let us have a three-component record with north–south, east–west, and vertical components s_i^N, s_i^E, and s_i^Z, respectively, in which i indexes time samples. We want to minimize the difference between the record and a synthetic disturbance m_i by finding proper amplitude A, azimuth ϕ, and inclination θ of the disturbance. The difference in the L2 norm

$$\sum_i (s_i^N - A m_i \cos\phi \cos\theta)^2 + \sum_i (s_i^E - A m_i \sin\phi \cos\theta)^2$$
$$+ \sum_i (s_i^Z - A m_i \sin\theta)^2 \tag{A1}$$

should be minimal, so its partial derivatives should be zero.

From $(\partial/\partial\phi) = 0$ we get

$$\frac{\sum_i m_i s_i^E}{\sum_i m_i s_i^N} = \frac{\sin\phi}{\cos\phi} = \tan\phi. \tag{A2}$$

From $(\partial/\partial\theta) = 0$ we get

$$\frac{\sum_i m_i s_i^Z}{\sum_i m_i (s_i^N \cos\phi + s_i^E \cos\phi)} = \frac{\sin\theta}{\cos\theta} = \tan\theta. \tag{A3}$$

From $(\partial/\partial A) = 0$ we get

$$\frac{\sum_i m_i (s_i^N \cos\phi \cos\theta + s_i^E \sin\phi \cos\theta + s_i^Z \sin\theta)}{\sum_i m_i^2} = A. \tag{A4}$$

Jiří Vackář
Jiří Zahradník
Charles University in Prague
Faculty of Mathematics and Physics
V Holešovičkách 2
180 00 Praha 8, Czech Republic
vackar@geo.mff.cuni.cz
jz@karel.troja.mff.cuni.cz

Jan Burjánek
Swiss Seismological Service
ETH Zurich
Sonneggstrasse 5
8092 Zürich, Switzerland
burjanek@sed.ethz.ch