USGS
science for a changing world

Session 85

Assessing Landslide Inventory Completeness Utilizing Benford's Law

by Maciej K. Obryk and Ben B. Mirus

THE GEOLOGICAL SOCIETY OF AMERICA®

Introduction

Statistical analysis of tabulated landslide inventories is the basis for hazard assessment (e.g., frequency of occurrence), and validation of landslide models and susceptibility maps. However, quantitatively evaluating landslide inventories is often difficult. Assessing their completeness is paramount for the effective hazard mitigation policies.

Contrarily, earthquake-triggered landslide inventories are often mapped completely (include all landslides produced by an earthquake), using high-resolution pre- and post-digital elevation models. Consequently, inventories of earthquake-triggered landslides can be used to evaluate new approaches for assessing completeness of non-earthquake-triggered landslide inventories.

Here we test Benford's law as a utility for assessing landslide inventory completeness.

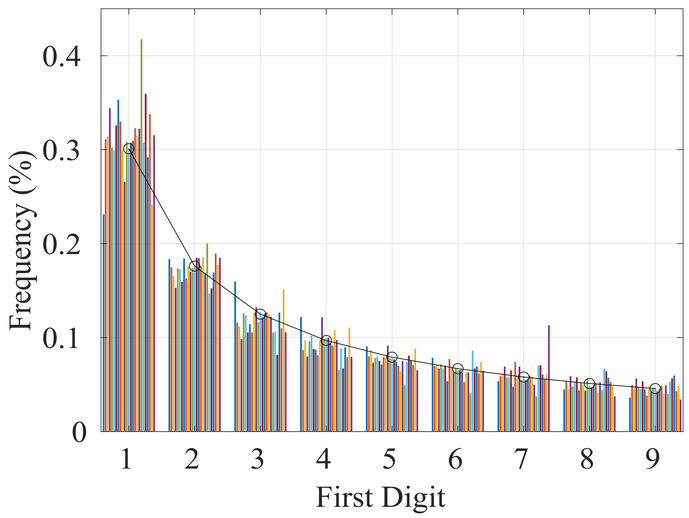

Figure 1. First digit frequency of surface areas from 25 seismically-triggered landslide inventories from China, Costa Rica, Guatemala, Haiti, Indonesia, Japan, Nepal, Pakistan, and the U.S. Benford's law distribution shown in open circles.

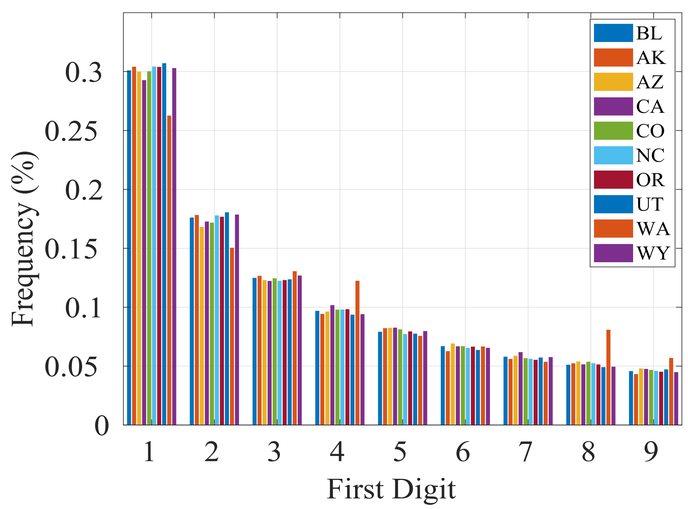

Figure 2. First digit frequency of surface areas from the U.S. landslide inventory. With the exception of the State of WA, mapped regions from shown states conform to Benford's law (BL - in blue).

Methods

Benford's law predicts the frequency of first digits in observational data sets and follows a rule:

$$P_{BL} = \log_{10}(1+1/D) \qquad (1)$$

where P_{BL} is the probability of the first digit D (D = 1, 2,…,9), disregarding leading zeros, decimal points, or the minus sign. Consequently, numbers starting with D = 1 are predicted to occur 30.1% of a time, numbers with D = 2 will occur 17.6% of time, and so on. Deviation from Benford's law is indicative of incomplete data sets (e.g., population not being represented by the sample, data errors, anomalies), rounding errors, or fraudulent data. Benford's law is both a scale- and base-invariant probability distribution. The degree of conformity was based on Mean Absolute Deviations[1]. Landslide inventories used in the analysis are available at www.sciencebase.gov.

Results

All analyzed seismically-triggered landslide inventories from around the world follow Benford's law distribution and 80% of those conform (Figure 1).

The U.S. landslide inventory is compiled from numerous sources, utilizing different classifications[2]. Consequently, each inventory is assigned a confidence level. Mapped regions within each analyzed state conform to Benford's law, except for WA (Figure 2). However, high confidence only inventories, derived using high resolution aerial photographs and LiDAR, conform to Benford's law (Figure 3).

These results suggest that Benford's law is a promising approach for evaluating landslide inventory completeness.

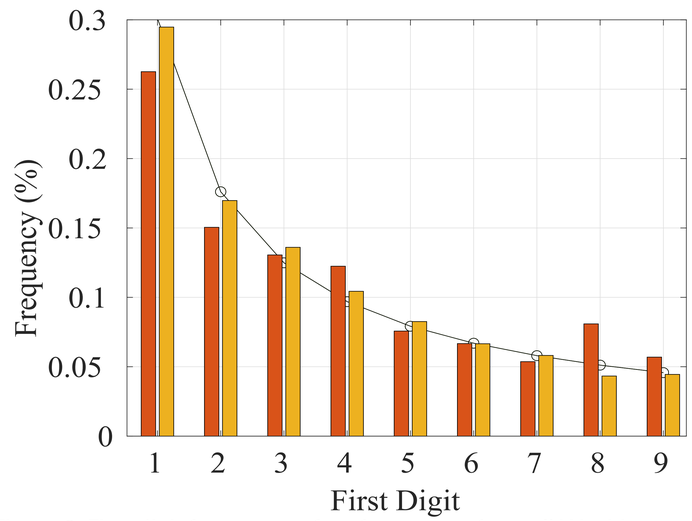

Figure 3. First digit frequency of surface areas from all Washington landslide inventories (orange - left bars) do not conform to Benford's law. However, high confidence landslides only (yellow - righ bars) conform.

References

[1]Nigrini, M.J., Wells, J.T. (2012). Benford's law: Applications for Forensic Accounting, Auditing, and Fraud Detection. John Wiley & Sons.

[2]Mirus, B. B., et al. (2020). Landslides across the USA: occurence, susceptibility, and data limitations. Landslides, 17(10). 10.1007/s10346-020-01424-4