DATA PAPER
https://doi.org/10.1590/2317-4889202420240006

BJGEO
Brazilian Journal of Geology

Landslides of the 2023 summer event of São Sebastião, southeastern Brazil: spatial dataset

Rebeca Durço Coelho[1,4], Camila Duelis Viana[1,4], Vivian Cristina Dias[2,4], Carlos Henrique Grohmann[3,4*]

Abstract

In February 2023, anomalously heavy rainfall caused widespread landslides in the coastal city of São Sebastião (Southeastern Brazil). This report describes the first version of a landslide inventory dataset for this event. The inventory is primarily based on the analysis of aerial images with 10 cm spatial resolution acquired immediately after the event, as well as archive images from Google Earth and PlanetScope. Delimitation of the landslides relied on a comparison of the images along with the area's Digital Surface Model (DSM) and hydrography. The spatial vector dataset (shapefile and geopackage) contains points representing the landslide's crowns and polygons indicating the affected area and is openly available in Zenodo.

KEYWORDS: landslide dataset; mass movement; digital terrain analysis; GIS.

INTRODUCTION

Between February 18 and 19, 2023, over 680 mm of rainfall poured down over the coastal city of São Sebastião (Southeastern Brazil) (CEMADEN 2023, Marengo *et al.* 2024), triggering hundreds of landslides (G1 2023). The event caused significant damage, including road blockages and mudslides that affected water and power supplies, resulting in 63 deaths, 40 missing persons, and approximately 1,730 homeless individuals (Agência Brasil 2023). In response to this crisis, the federal government declared a state of public emergency (Brasil 2023).

The city of São Sebastião is located on the northern shore of the São Paulo State (Fig. 1). The geomorphology is marked by the contrast between the coastal plains and the high relief of the Brazilian Coastal Range ("*Serra do Mar*"), a 1,500-km-long mountain range composed mainly of igneous and metamorphic rocks with deep valleys and altitudes reaching 800 m.a.s.l. near the city of São Sebastião (Almeida and Carneiro 1998, Vieira and Gramani 2015).

The Coastal Range region has a high potential for landslides and debris flows, the most common mass movements are shallow landslides in areas of medium to high slopes (Brollo *et al.* 2015, Cruz 1974, De Ploey and Cruz 1979, Fúlfaro *et al.* 1976, Gramani 2001, Coelho Netto *et al.* 2013, Wolle and Hachich 1989). The advancement of urbanization toward the mountains often exacerbates these events. A recent example of a disaster related to mass movements in the region is the 1967 event of Caraguatatuba, which took around 450 lives and left over 3,000 people homeless (Cunha *et al.* 2022).

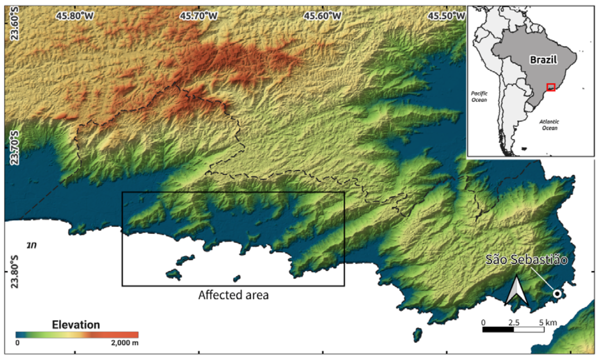

Figure 1. São Sebastião location and geomorphology.

This study presents the first version of a spatial dataset documenting the landslides that occurred in São Sebastião during the summer of 2023. The main rationale for providing the dataset as spatial vector files, openly available in Zenodo (see Conclusion and Data availability), is to provide an example of how openly sharing research data can be beneficial not only to the Geoscientific Brazilian community but also to all researchers interested in the phenomenon, as well as policy- and decision-makers.

The "*FAIR Principles for scientific data management and stewardship*" (Wilkinson *et al.* 2016) stand for Findability, Accessibility, Interoperability, and Reuse of digital assets. The fourth principle (Reuse) is of particular interest to those studying mass movements because it is still common for different research groups to recreate landslide inventories for a

[1]Universidade de São Paulo, Instituto de Geociências, São Paulo, SP, Brasil. E-mails: rebeca.coelho@alumni.usp.br, camila.viana@usp.br

[2]Universidade de São Paulo, Instituto de Energia e Ambiente, São Paulo, SP, Brasil. E-mail: vivian.cristina.dias@alumni.usp.br

[3]Universidade de São Paulo, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, São Paulo, SP, Brasil. E-mail: guano@usp.br

[4]Universidade de São Paulo, Spatial Analysis and Modelling Lab (SPAMLab), São Paulo, SP, Brasil.

*Corresponding author.

particular event based on free orbital imagery (e.g., Google Earth) or on small-scale maps from scientific publications, such as the one presented by Fúlfaro *et al.* (1976). An open dataset, based on very-high-resolution imagery, improves the reproducibility of risk or susceptibility analysis and allows for unbiased comparisons of different methods and algorithms.

METHODS

The classification of landslides was established according to definitions provided in the literature. Expanding on Varnes' (1978) definition, Cruden (1991) provided a more comprehensive explanation of landslides and defined them as the downward movement of rock, earth, or debris masses on a slope. It is important to note that landslides can take various forms beyond simple sliding failures, including falls, flows, topples, and spreads.

The dataset primarily consists of translational landslides (Varnes 1978, Cruden and Varnes 1996). When identifying translational landslides in the area, the evaluation considered the following criteria:

- Hillslopes with no direct human intervention;
- Hillslopes showing evidence of recent mobilization of vegetation cover that allows identification of the extension of landslide debris directly downslope.

The landslides were identified through the visual interpretation of aerial photographs with a spatial resolution of 10 cm provided by IDE-SP (2023). The aerial survey was carried out on February 25, 2023, immediately after the event, to provide support for on-site personnel (civil protection, police, firemen, paramedics, etc.) involved in the search and rescue efforts.

QGIS version 3.28 was used to manually digitize the landslides and organize the dataset files. The aerial mosaic image was divided into 1 × 1 km quadrants to provide sufficient detail during mapping. Relict landslides (triggered before the 2023 event) or other geomorphologic processes such as gullies and ravines were not considered in the mapping process.

It is important to consider the limitations of the landslide dataset in the São Sebastião region when using and analyzing it. One of the main limitations of such inventories is the visual identification of landslides in areas with dense vegetation or complex relief.

Landslides might be of different types and sizes and may not always be easily identifiable through the visual analysis of aerial images alone. In some cases, there may even be hidden landslides that require a field inspection to be detected (Fig. 2). Therefore, it is important to exercise caution when interpreting aerial images and to conduct thorough fieldwork inspections to ensure accurate landslide identification.

Another limitation is the lack of historical landslide data in the region. Comparison with previous data can help quantify when landslides occurred and identify trends. However, as historical landslide data for the São Sebastião region are scarce, this comparison is currently impossible.

During the fieldwork, a debris flow was identified (Fig. 3). The extent of the debris flow was determined based on the presence of visible erosion along stream banks, the presence of deposition evidence, and the terrain slope, which was determined using DEM extract contour lines with a resolution of 20 m. The total zone of the debris flow was delimited through visual interpretation using the high-resolution image of IDE (10 cm).

DATASET DESCRIPTION

The dataset is divided into points and polygons (Fig. 4). Points correspond to the crown of the landslide, which is

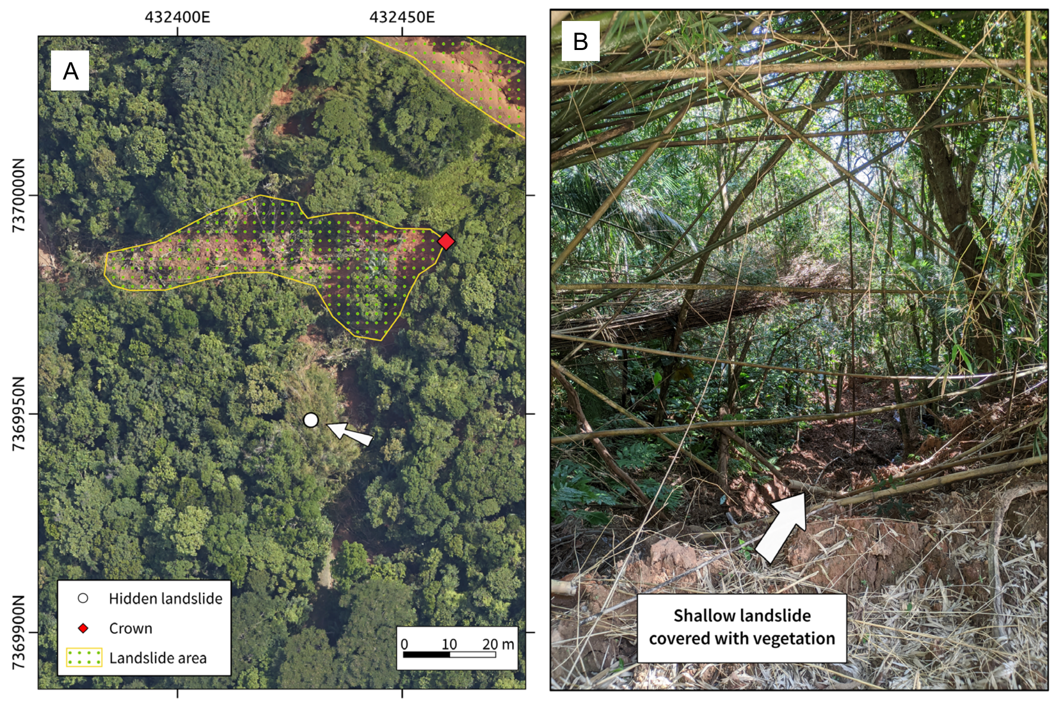

Figure 2. (A) Shallow landslide covered with vegetation. B) Verification of the complete extent of the landslide is only possible through fieldwork (the white arrow indicates the direction of the mobilized material).

defined as the area that indicates the material that remains in place, is adjacent to the highest parts of the main scarp, and is essentially undisplaced (Cruden and Varnes 1996, Highland and Bobrowsky 2008).

The polygons represent the regions where the material was displaced during the 2023 event, outlining the total affected area that is visible for vectorization. This mapping excludes drainage or other morphodynamic processes, such as gullies.

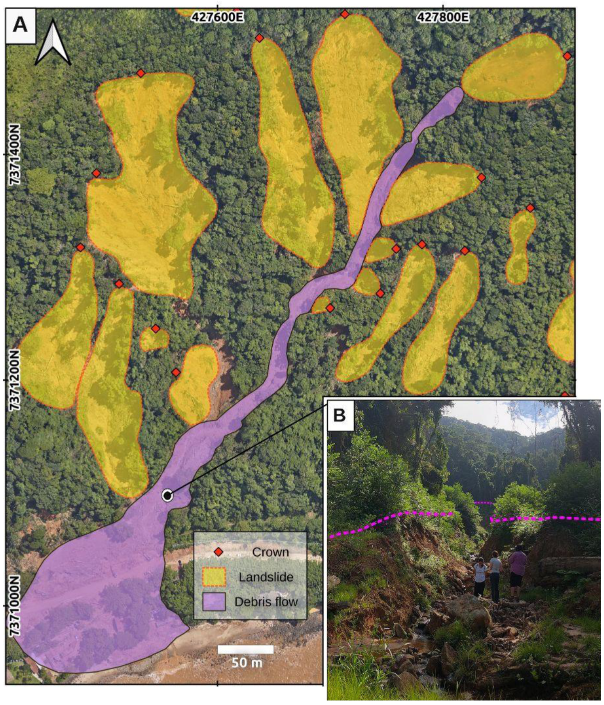

Figure 3. A debris flow was identified in fieldwork. (A) Debris flow runout extension. (B) Field investigation and data collection; purple dots indicate the debris flow zone.

The dataset contains 983 polygons and 1,070 points. The aerial images' high resolution (10 cm) allowed for the identification of small/hidden landslides, indicated with a single point. These represent locations where it is possible to identify soil mobilization, but not to clearly define its boundaries due to vegetation cover or shadows (see Fig. 2).

To characterize the dataset's basic geomorphometric statistics, we used a TanDEM-X Digital Elevation Model with a resolution of 12.5 m (European Union 2022, Grohmann 2018). The geomorphometric variables slope, aspect, and profile curvature were calculated in QGIS using WhiteboxTools (Lindsay 2016), and their values were sampled at the crown's point locations.

Geometrical attributes for polygons (area, perimeter) were calculated with the QGIS tool "Add geometry attributes." The azimuth and length of the main axis for each scar were calculated with the "Oriented minimum boundary box" tool. While aspect angle is given in the 0–360° range, the azimuth of the main axis is produced in the 0–180° range. Descriptive statistics (Tables 1 and 2) were determined using the QGIS "Show statistical summary" tool, except for the circular mean and circular standard deviation of aspect and main axis azimuth, which were computed with the Python library Astropy (Astropy Collaboration *et al.* 2013).

Histograms of elevation, slope, profile curvature, and area are presented in Fig. 5. Rose diagrams for aspect and main

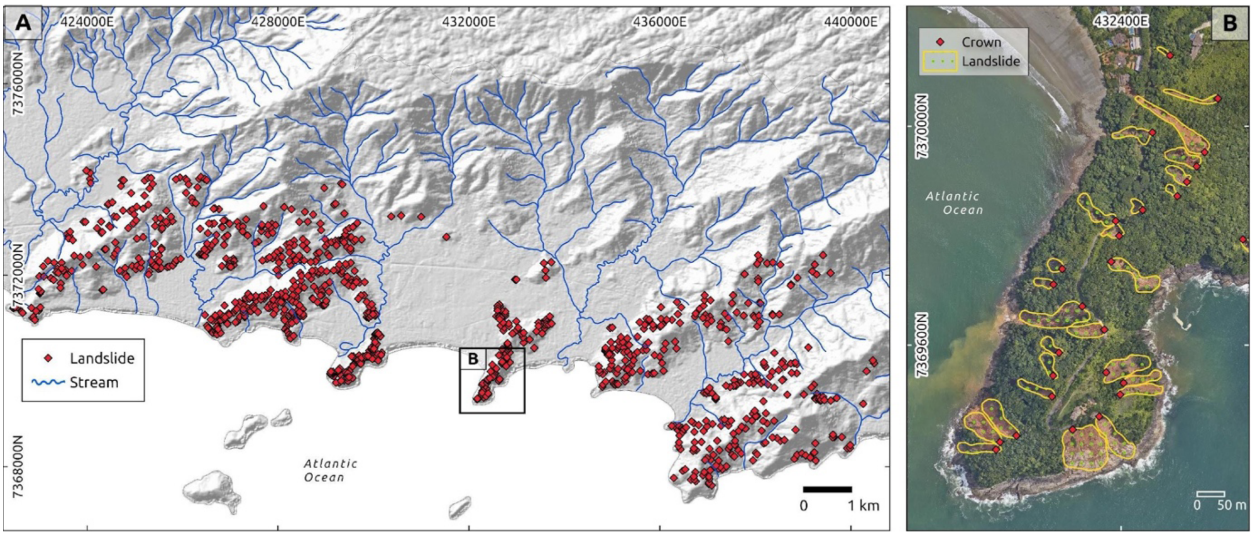

Figure 4. (A) Spatial distribution of shallow landslides (crowns). (B) Example of points (crown) and polygons (landslide's affected area) of the landslide dataset.

Table 1. Descriptive statistics for geomorphometric parameters sampled at landslides' crowns (points).

Parameter	Elevation (m)	Slope (°)	Slope (%)	Aspect (°)	Prof. Curvature (m^{-1})
Minimum	3.00	0.82	1.44	0.23	-0.02632
Mean	136.14	27.21	53.17	175.45	0.00250
Std. dev.	79.53	9.17	21.28	1.30	0.00600
Maximum	458.1	63.08	196.92	358.77	0.03534

Table 2. Descriptive statistics for geometrical attributes of landslides' scars (polygons).

Parameter	Area (m^2)	Perimeter (m)	Axis azimuth (°)	Axis length (m)
Minimum	18.04	17.37	0.00	5.88
Mean	2,510.04	215.00	100.62	84.10
Std. dev.	3,840.77	179.26	0.84	68.29
Maximum	39,076.20	1,360.05	180.00	516.76

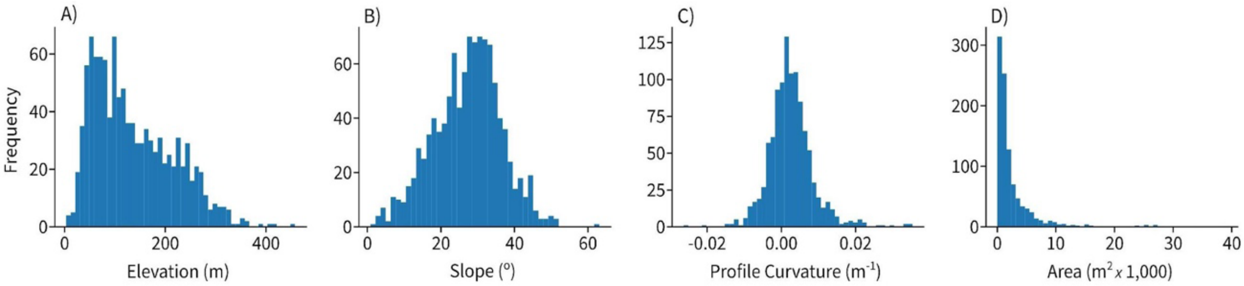

Figure 5. Histograms of geomorphometric parameter values. (A) Elevation. (B) Slope (degrees). (C) Profile curvature. (D) Area.

axis azimuth were created in OpenStereo (Grohmann and Campanha 2010), projecting the azimuths into the northern half of the diagram (Fig. 5).

Landslides identified in this study occur in all geomorphological settings, with crowns at low elevations (3 m) and near-flat slopes to elevations beyond 450 m and steep slopes over 60°. The distribution of elevation for crowns is asymmetric, with the majority of occurrences between 50 and 100 m (Fig. 5A). Slope values for crowns show a clear peak around 30° (Fig. 5B). The distribution of profile curvature, although symmetrical, is not centered around zero but shifted toward positive values (Fig. 5C), indicative of flow acceleration in convex profiles. The histogram of landslide scar area is strongly positively skewed, with the majority of scars having an area under 5,000 m^2 (Fig. 5D). Rose diagrams show the main axis of scars oriented mostly at NW-SE and E-W (Fig. 6).

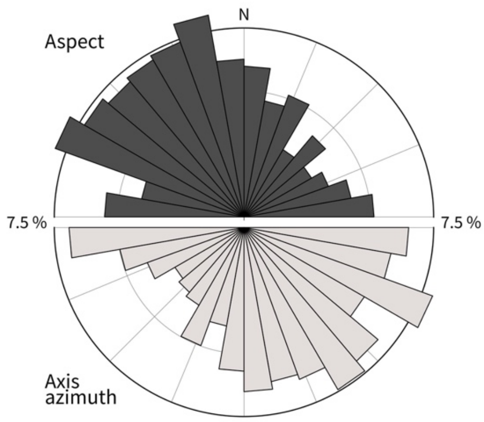

Figure 6. Rose diagram of aspect at crowns' locations (upper plot) and azimuth of scars' main axis (lower plot).

CONCLUSION AND DATA AVAILABILITY

This data paper presents a dataset of landslide crowns (983 points) and scars (1,070 polygons) for the February 2023 disaster of São Sebastião, manually interpreted over aerial images with 10 cm spatial resolution surveyed immediately after the event.

The dataset is openly available on Zenodo (https://doi.org/10.5281/zenodo.11120078), in ESRI shapefile and Geopackage formats, and is in UTM projection, zone 23, southern hemisphere, WGS84 datum.

The authors expect to update the dataset when new high-resolution elevation data based on airborne laser altimetry (LiDAR) become available.

ACKNOWLEDGMENTS

The authors acknowledge the financial support provided by CAPES (Finance code 001), FAPESP (grants #2019/26568-0, #2022/04233-9, and #2023/11197-1), and CNPq (grant #311209/2021-1). TanDEM-X data were provided by the German Aerospace Centre (DLR, proposals DEM_HYDR3714 and DEM_GEOL0538). We would like to express our thanks to the Graduate Program in Mineral Resources and Hydrogeology at The Institute of Geosciences (IGc-USP). Acknowledgments are extended to the Associate Editor and the anonymous reviewers for their criticism and suggestions, which helped to improve this paper.

ARTICLE INFORMATION
Manuscript ID: e20240006. Received on: 29 JAN 2024. Approved on: 26 JUNE 2024.

How to cite: Coelho R.D., Viana C.D., Dias V.C., Grohmann C.H. 2024. Landslides of the 2023 summer event of São Sebastião, southeastern Brazil: spatial dataset. *Brazilian Journal of Geology*, 54(2):e20240006. https://doi.org/10.1590/2317-4889202420240006

R.D.C.: Data curation (Equal), Formal analysis (Equal), Investigation (Equal), Methodology (Equal), Software (Equal), Visualization (Equal), Writing—original draft (Equal), Writing—review and editing (Equal). C.D.V.: Conceptualization (Equal), Data curation (Equal), Investigation (Equal), Methodology (Equal), Validation (Equal), Writing—review and editing (Equal). V.C.D.: Data curation (Equal), Investigation (Equal), Methodology (Equal), Writing—review and editing (Equal). C.H.G.: Conceptualization (Equal), Data curation (Equal), Funding acquisition, Supervision, Writing—original draft (Equal), Writing—review and editing (Equal).

REFERENCES

Agência Brasil. 2023. Governo federal reconhece calamidade em seis municípios paulistas. *Agência Brasil*. Available at: https://agenciabrasil.ebc.com.br/geral/noticia/2023-02/governo-federal-reconhece-calamidade-em-seis-municipios-paulistas. Accessed in: June 2023.

Almeida F.F.M., Carneiro C.D.R. 1998. Origem e evolução da serra do mar. *Brazilian Journal of Geology*, **28**(2):135-150. https://doi.org/10.25249/0375-7536.1998135150

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., Greenfield, P., Droettboom, M., Bray, E., Aldcroft, T., Davis, M., Ginsburg, A., Price-Whelan, A. M., Kerzendorf, W. E., Conley, A., Crighton, N., Barbary, K., Muna, D., Ferguson, H., Grollier, F., Parikh, M. M., Nair, P. H., Unther, H. M., Deil, C., Woillez, J., Conseil, S., Kramer, R., Turner, J. E. H., Singer, L., Fox, R., Weaver, B. A., Zabalza, V., Edwards, Z. I., Azalee Bostroem, K., Burke, D. J., Casey, A. R., Crawford, S. M., Dencheva, N., Ely, J., Jenness, T., Labrie, K., Lim, P. L., Pierfederici, F., Pontzen, A., Ptak, A., Refsdal, B., Servillat, M., and Streicher, O. 2013. Astropy: A community Python package for astronomy. *Astronomy & Astrophysics*, **558**: A33. https://doi.org/10.1051/0004-6361/201322068

Brasil. 2023. Portaria nº 799, de 19 de fevereiro de 2023. Reconhece o estado de calamidade pública no município de São Sebastião/SP. *Diário Oficial da União*. Available at: https://www.in.gov.br/web/dou/-/portaria-n-799-de-19-de-fevereiro-de-2023-465407891. Accessed in: June 2023.

Brollo M., Santoro J., Penteado D., Fernandes da Silva P., Rodrigues R. 2015. Itaoca (SP): Histórico de acidentes e desastres relacionados a perigos geológicos. *Simpósio de Geologia do Sudeste*, 14, Campos do Jordão. Electronic proceedings. Available at: http://www.sbgeo.org.br/assets/arquivos/simposiodosudeste/2015.pdf. Accessed in: May 2023.

Centro Nacional de Monitoramento e Alertas de Desastres Naturais (CEMADEN). 2023. *Mapa interativo da rede observacional para monitoramento de risco de desastres naturais do CEMADEN*. CEMADEN. Available at: http://www2.cemaden.gov.br/mapainterativo/#. Accessed in: May 2023.

Coelho Netto A.L., Sato A.M., de Souza Avelar A., Vianna L.G.G., Araújo I.S., Ferreira D.L.C., Lima P.H., Silva A.P.A., Silva R.P. 2013. January 2011: The Extreme Landslide Disaster in Brazil. In: Margottini C., Canuti P., Sassa K. (eds.). *Landslide Science and Practice*. Berlin, Heidelberg: Springer. p. 377-384. https://doi.org/10.1007/978-3-642-31319-6_5

Cruden D.M. 1991. A simple definition of a landslide. *Bulletin of the International Association of Engineering Geology*, **43**(1):27-29. https://doi.org/10.1007/BF02590167

Cruden D.M., Varnes D.J. 1996. Landslide types and processes. In: Turner A.K., Schuster R.L. (eds.). *Landslides investigation and mitigation*. Special Report 247. Washington, D.C.: Transportation Research Board, US National Research Council. p. 36-75.

Cruz O. 1974. *A serra do mar e o litoral na área de Caraguatatuba: contribuição à geomorfologia tropical litorânea*. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo.

Cunha M.A., Paula M.S., Iyomasa W.S., Gramani M.F., Massad F. 2022. *Debris flow na Serra do Mar – O caso de Caraguatatuba 1967*. São Paulo: Oficina de Textos.

De Ploey J., Cruz O. 1979. Landslides in the serra do mar, Brazil. *Catena*, **6**(2):111-122. https://doi.org/10.1016/0341-8162(79)90001-8

European Union. 2022. *TanDEM-X Digital Elevation Model (DEM) - Global, 12m*. European Union. Available at: http://data.europa.eu/88u/dataset/5eecdf4c-de57-4624-99e9-60086b032aea. Accessed on: Dec. 30, 2023.

Fúlfaro V.J., Ponçano W.L., Bistrichi C.A., Stein D.P. 1976. Escorregamentos de Caraguatatuba: Expressão atual, e registro na coluna sedimentar da planície costeira adjacente. *Congresso Brasileiro de Geologia de Engenharia*, **2**:341-350.

G1. 2023. Temporal causa 40 mortes, deixa desabrigados e fecha estradas no litoral norte de SP. *G1*. Available at: https://g1.globo.com/sp/vale-do-paraiba-regiao/noticia/2023/02/19/chuva-bloqueia-rodovias-cancela-carnaval-e-provoca-alagamentos-em-cidades-do-litoral-de-sao-paulo.ghtm. Accessed on: June 2023.

Gramani M.F. 2001. *Caracterização geológica-geotécnica das corridas de detritos ("debris flows") no Brasil e comparação com alguns casos internacionais*. Unpublished.

Grohmann C.H. 2018. Evaluation of TanDEM-X DEMs on selected Brazilian sites: Comparison with SRTM, ASTER GDEM and ALOS AW3D30. *Remote Sensing of Environment*, **212**:121-133. https://doi.org/10.1016/j.rse.2018.04.043

Grohmann C.H., Campanha G.A.C. 2010. OpenStereo: Open Source, Cross-Platform Software for Structural Geology Analysis. *Fall Meeting*. San Francisco: American Geophysical Union,

Highland L.M., Bobrowsky P. 2008. *The Landslide Handbook: A Guide to Understanding Landslides*. Reston: U.S. Geological Survey Circular 1325.

Infraestrutura de Dados Espaciais do Estado de São Paulo (IDE-SP). 2023. *Vôo 2023 São Sebastião 10 cm Vila Sahy*. São Paulo: IDE-SP. Available at: http://www.metadados.idesp.sp.gov.br/catalogo/srv/eng/catalog.search#/metadata/ b72b2043-e767-400f-b012-73d03afbcb7f. Accessed on: May 2023.

Lindsay J. (2016). Whitebox gat: A case study in geomorphometric analysis. *Computers & Geosciences*, **95**:75-84. https://doi.org/10.1016/j.cageo.2016.07.003

Marengo J.A., Cunha A.P., Seluchi M.E., Camarinha P.I., Dolif G., Sperling V.B., Alcântara E.H., Ramos A.M., Andrade M.M., Stabile R.A., Mantovani J., Park E., Alvala R.C., Moraes O.L., Nobre C.A., Gonçalves D. 2024. Heavy rains and hydrogeological disasters on February 18th–19th, 2023, in the city of São Sebastião, São Paulo, Brazil: from meteorological causes to early warnings. *Natural Hazards*, 120:7997-8024. https://doi.org/10.1007/s11069-024-06558-5

Varnes D.J. 1978. Slope movement types and processes. In: Schuster R.L., Krizek R.J., eds. *Landslides: Analysis and Control*. Washington, D.C.: TRB, National Research Council. p. 11-33.

Vieira B.C., Gramani M.F. 2015. Serra do Mar: The Most "Tormented" Relief in Brazil. In: Vieira B., Salgado A., Santos L. (eds.). *Landscapes and Landforms of Brazil. World Geomorphological Landscapes*. Dordrecht: Springer. https://doi.org/10.1007/978-94-017-8023-0_26

Wilkinson M.D., Dumontier M., Aalbersberg I.J., Appleton G., Axton M., Baak A., Blomberg N., Boiten J.-W., da Silva Santos L.B., Bourne P.E., Bouwman J., Brookes A.J., Clark T., Crosas M., Dillo I., Dumon O., Edmunds S., Evelo C.T., Finkers R., Gonzalez-Beltran A., Gray A.J., Groth P., Goble C., Grethe J.S., Heringa J., 't Hoen P.A., Hooft R., Kuhn T., Kok R., Kok J., Lusher S.J., Martone M.E., Mons A., Packer A.L., Persson B., Rocca-Serra P., Roos M., van Schaik R., Sansone S.-A., Schultes E., Sengstag T., Slater T., Strawn G., Swertz M.A., Thompson M., van der Lei J., van Mulligen E., Velterop J., Waagmeester A., Wittenburg P., Olstencroft K., Zhao J., Mons B. 2016. The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, **3**(1):160018. https://doi.org/10.1038/sdata.2016.18

Wolle C.M., Hachich W.C. (1989). Rain-induced landslides in southeastern Brazil. In *Proceedings*. Rotterdam: Balkema. Available at: https://www.issmge.org/publications/publication/rain-induced-landslides-in-southeastern-brazil. Accessed in: May 2023.