# Supplementary materials

Aerial photographs dataset

In order to figure out fault surface ruptures and deformations registered by geomorphic markers, an original historical dataset of aerial photographs has been used. These photographs were taken by a Russian mission, a few months after the 1957 Gobi-Altai earthquake, during the first part of 1958 (Florensov and Solonenko, 1965). Only little information about these photographs is available: date and hour of each photographs, the relative horizontality (bubble level, not actually significant in a flying plane), and an approximate value of the focus length of 100.0 ± 0.1 mm. Pictures are printed version, square shaped, 17.9 cm wide, and have been developed a long time ago (at least 40 years). Orientations of the photographs are not known, and checking the bubble level it is assumed that all pictures are near-nadir. Aerial photographs are all referenced, and split in two series. In the first one, named S10000 regarding to the approximate scale of 1:10000, we find low altitude aerial photographs following the main scarps. Ground configuration is therefore linear, and their scales could be quite variable, due to the reliefs of Bogd mountain ranges. The coverage of the surface rupture by S10000 is incomplete, especially along the Baga Bogd Mountain (Eastern third of the Eastern Bogd fault) where the fault rupture is distributed. Almost two thirds of the whole main surface rupture are imaged by these photographs. The second series, named S25000, which covers the integrality of the Bogd mountain ranges, has been taken at variable flight altitudes. These are higher altitude aerial photographs than S10000, with scales ranging from 1:20000 to 1:25000. S25000 photos are used in this study whenever S10000 are not available, or at places where the fault is highly distributed, with part of the surface rupture out of the S10000 coverage. All the useful photographs selected for their interest in studying surface rupture have been scanned at 1600dpi with a professional flatbed scanner. Considering the scale, the picture size and the scanning resolution, resulting orthophotographs resolutions are comprised between 15 to 20 cm (theoretically 18 cm) for orthophotographs generated from S10000, and from 40 to 60 cm for orthophotographs produced from S25000. These document are used together with freely available high-resolution satellite images (e.g. Google Earth, Bing Maps), and SRTM30 DEM dataset (Slater et al., 2006) for further analysis of the Eastern Bogd fault.

Orthophotographs generation

* Parametrization

Due to unknown internal and external parameters of the pictures, it is not possible to process the data using a classical photogrammetric process. Thereafter we use a structure from motion approach (Koenderink and Doorn, 1991), allowing to reconstruct internal parameters (camera locations and orientation). The orthophotographs generation follows a standard structure from motion process using the professionnal *Agisoft® Photoscan* *Pro* software. Due to the linear disposition of the S10000 photographs and in order to avoid distortion propagation in the photogrammetric treatment, photogrammetric computing projects are composed of 4 to 10 aerial photographs. The spatial coverage of these projects must be located within the same high resolution satellite image, in order to avoid georeferencing disconnection. The 1600 dpi scanned pictures are then loaded with the fiducial marks as well as complementary information (picture reference, level, and clock) manually masked in order to preserve a maximum of the visible surface. The camera calibration is parametrized with a frame camera type and a 100 mm focal length, according to meta-information relative to the photographs. GCPs are mainly natural features, such as isolated rocks, trees and sharp topographic heights, but could also be anthropic like unchanged livestock paddock walls, or archeological features. The GCPs have to be widely distributed on each picture both in planimetry and in altitude to constrain in 3D the correlation model. Because GCP positions within one project are all extracted from a unique satellite image, even if the georeferencing of that image is incorrect, relative distances within the image are preserved. All aerial photographs are of the same dimension in pixel, allowing the focal length approximation by the structure from motion process (James and Robson, 2014). Absolute georeferencing errors are preserved without affecting the orthorectification process regarding the satellite image. Altitude data are specified from the SRTM30 DEM, with expectedly high uncertainties resulting in a low accuracy in altitude of our own 3D models. A minimum of 5 GCPs per picture are specified in UTM coordinates (UTM 47N) with 0.3 m precision and 1 m relative uncertainty.

These values have been used both to maintain the georeferencing accuracy in spite of large uncertainties in altitude inherent to the SRTM DEM, and to allow a certain amount of liberty for the computation in order to avoid incoherencies (e.g. GCPs out of the point cloud or important noise in the point cloud) which might lead to artifacts.

* Processing

The first step in the photogrammetric processing is to align aerial photographs. This step generates a sparse point cloud by computing Tie Points (TPs) between photographs. Next step is to build the geometry (relative position of cameras) from that sparse point cloud. Basically, sparse point cloud average spatial density from the S10000 1958 aerial photography is 2000 to 9000 TP per photograph, not evenly distributed. A first alignment and geometry-building computing allows efficient selection of the GCPs and give estimation of their accuracy. This marker/model error has to be minimized both in pixels (relative to the on-picture position) and in meters (relative to GCPs’ coordinates). This error improvement could be done by optimizing the photographs alignment, by adding GCPs, correcting inaccurate GCPs, or by changing the size of photogrammetric projects (the addition of more pictures might foster correlation, or instead discarding some low-neatness picture might optimize the correlation). Once the GCPs location errors optimized and the geometry built, a dense point cloud may be computed. Obtained dense-point clouds are 1 to 10 pts.m-², depending on the accuracy requested for the computing. Based on a trade-off in point density and correlation errors, there are four levels of quality for the TPs detection in *Agisoft® Photoscan* (low, medium, high and very high), generating different TPs density and distribution regarding to the images. In order to optimize orthophotographs quality and due to the age of aerial photographs, several qualities have been computed for each computing project. The last computation steps consist of building a 3D mesh from the dense cloud, and to generate texture elements for each facet. Finally, DEM and orthophotographs are exported in GeoTIFF format. DEMs resolutions are 40 to 50cm for S10000 series, and 1 to 2 m for S25000. Produced DEM exhibit two limitations: i) computational noise due to the poor quality of original photographs; ii) vertical offsets between photogrammetric projects, due to inaccuracy of the GCPs coordinates.

Overall, our DEM are adequate to perform ortho-rectification but do not allow high-quality infra-metric measurements across the fault scarps. Orthophotographs are exported at resolution of 15 cm to 20 cm for the orthophotographs generated from S10000 and 50cm for the orthophotographs produced from S25000. Figure 2.a. illustrates the S10000 orthophotographs rendering compared to Bing image, and their geomorphic interpretations.

* Coverage

The coverage of the fault surface rupture by S10000 is not exhaustive. The linear configuration offers a 1 to 2 km wide band, following the main scarp along the western 2/3 of its total length. However, some images are missing and two projects have been done from S25000 in order to complete the coverage. Locally, where 1958 data were not available to cover the full surface rupture, modern high resolution optical satellite images have been used. The easternmost 80 km of the Gobi-Altai earthquake surface rupture, corresponding to the Baga Bogd area, present a highly distributed fault pattern, and only a few S10000 pictures display identifiable faulting. These images have been computed to produce orthophotographs, but their coverage is very weak compared to the distributed surface rupture area. Total budget of fault covered by 1958 aerial photos is 150km of localized deformation along the Eastern part of the Bogd fault (Bhaar Uul, Noyon Uul, Ikh Bogd and Dullan Bogd), with 120km covered by S10000 and 30km by S25000.

* Orthophotographs dataset

The orthophotographs have been published in the *figshare* platform, in *kml* format in order to minimize their weight. Titles of the orthophotographs correspond to the names of the site they cover. Associated DOI is 10.6084, and the data can be found at the following URL: <https://doi.org/10.6084/m9.figshare.5773134.v1>.

* Limits

Such photogrammetric modeling from historical aerial photographs has its limits. As previously exposed, the aim is to be as accurate as possible, staying aware of unavoidable inaccuracies. Main identified limits of this approach are: georeferencing, orthorectification, accuracy and partial coverage. Main limiting factors are low resolution altitude data from SRTM, unknown camera position and orientation, lens distortion, photographs preservation, uncertainties of GCPs locations, and S10000 linear configuration. Actually, linear distribution pattern is not favorable for structure from motion photogrammetric process (James and Robson, 2014).

Due to the old age of the original aerial photographs, some of them could present some defects. For instance all of the original photographs have bug-holes (sometime in the center-part of the pictures); some of them are fuzzy from inadequate printing, others are folded, torn, or have stains and pen strokes. These defects will partially figurate in the final orthophotographs, reducing their readability and accuracy. Moreover, the previously cited inaccuracies in the georeferencing data and some missing information like the camera lens distortion parameters or the camera locations could generate other artefacts during the photogrammetric process. Such artefacts might be miss-georeferencing, texture artefact, or ghosting within orthophotographs. Artefacts presence will further be integrated in the evaluation of each measurement done on the orthophotographs.

* Validation

In order to validate this orthorectification approach, a validation test has been performed, comparing 100 length measurements done on 5 different satellite images used for the georeferencing process, and the generated orthophotographs. The dataset is divided into 4 classes: [0 - 5 m[, [5 – 15 m[, [15 - 50 m[ and [50 – 150 m].

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Nb. meas. | Err av. (m) | Err av. (%) | 1σ (m) |
| [0-5m] | 25 | 0,15 | 4,16 | 0,17 |
| ]5-15m] | 25 | 0,31 | 3,11 | 0,38 |
| ]15-50m] | 25 | 0,86 | 2,66 | 1,09 |
| ]50-150m] | 25 | 0,87 | 1,23 | 1,17 |

Table: Ortho-rectification validation test results. Nb. meas: number of measurement; Err av.: average absolute error in meters (m) and in percent (%); 1σ: one standard deviation of the error distribution.

Results in table show less than 5 % error for all measurements, and accounting for the average error of about 30 cm for measurement below 15 m and about 90 cm for measurements above 15 m, we conclude that this test validates the accuracy of produced orthophotographs for offset measurements.

Mapping

Detailed mapping is given in 4 kml files. Three are the main strike slip surface rupture: mains scarp, cracks and secondary faults as detailed in the text. One more file is for secondary fault rupture south and south-east from Baga Bogd.

Offset measurements dataset

Hereafter is a table listing offset measurements. Seg.: segment or relay reference; Lon.: longitude UTM, Lat.: latitude UTM, Marker res.: resolution of the marker (/2), encompassing both sinuosity (/1) and straightness (/1); Nb of splay: number of fault splay affecting the marker; Scarp 2nd: number of secondary scarp out of the measurement area; Offset Pref: preferred offset value (in m); Offset min/max: minimum/maximum offset value (in m); Δ+/-: absolute difference between preferred and maximum/minimum offset values (in m); Image type: type of image used for the measurement (mostly orthophotographs from 1958 aerial photographs, something Bing on Google Map images); Image res.: image resolution (in m).