Tutorial for Improved Uncertainty Estimating Method

Hanshaw's method and modification

We modified Hanshaw's (2014) equation that have been used to calculate lake-area mapping uncertainty. Lake perimeter and displacement error are widely used to estimate the uncertainty of glacier and lake mapping from satellite observation. Hanshaw and Bookhagen (2014) proposed an equation to calculate the error of area measurement by the number of edge pixels of the lake boundary multiplied by half of a single pixel area. The number of edge pixels is simply calculated by the perimeter divided by the grid size. The equation is expressed as below:

$$Error(1\sigma) = \frac{P}{G} \times 0.6872 \times \frac{G^2}{2} \tag{1}$$

$$D = \frac{Error(1\sigma)}{A} \times 100\% \tag{2}$$

Where G is the cell size of the remote sensing imagery (10 m for Sentinel-2 image and 30 m for Landsat image). P is the perimeter of individual glacial lake (m), and the revised coefficient of 0.6872 was chosen assuming that area measurement errors follow a Gaussian distribution. Relative error (D) was calculated by equation 3, in which A is the area of an individual glacial lake.

In the original equation 2, the number of edge pixels varies by the shape of lake and is indicated by $\frac{P}{G}$. However, the pixels in the corner are double counted. The total number of repeatedly calculated edge pixels equals the number of inner nodes. Therefore, we adjusted the calculation of the actual number of edge pixels as the maximum of edge pixels ($\frac{P}{G}$) subtracting the number of inner nodes. Accordingly, the equation of uncertainty estimation for lake mapping is modified as below:

$$Error(1\sigma) = (\frac{P}{G} - N_{inner}) \times 0.6872 \times \frac{G^2}{2} \tag{3}$$

Where N_{inner} is the number of inner nodes (inflection points) of each lake, which was automatically calculated based on the lake polygon in the ArcGIS software. The modified equation is also suitable for lakes with islands.

The Hanshaw's equation was originally proposed for pixelated polygons (such as a polygon directly extracted from a remote sensing image), and performed more robustly than manually digitized polygons (where vertices do not necessarily follow the pixel edges). Our improved method also performs better for pixelated polygons.

Procedure of uncertainty estimating method (using ArcGIS for example)

1. Removing redundant nodes (optional)

We found that a small proportion (~1%) of the pixelated lake polygons (directly extracted from satellite images) have redundant nodes, which affects the value of inner nodes. If no

redundant nodes exist, this step can be skipped. Or, we recommend using the "Simplify Polygon" tool in ArcGIS to remove those nodes (**Figure S1**).

In the Simplify Polygon panel
- Input your dataset.
- Set the output path and output file name.
- Choose the simplification algorithm. We recommended "POINT_REMOVE".
- Set the tolerance of simplification algorithm. In this step, we need to ensure that the polygon boundaries remain unchanged after deleting redundant nodes. Generally, a tolerance of 1 meter will suffice, or you can adjust the threshold until your satisfaction.
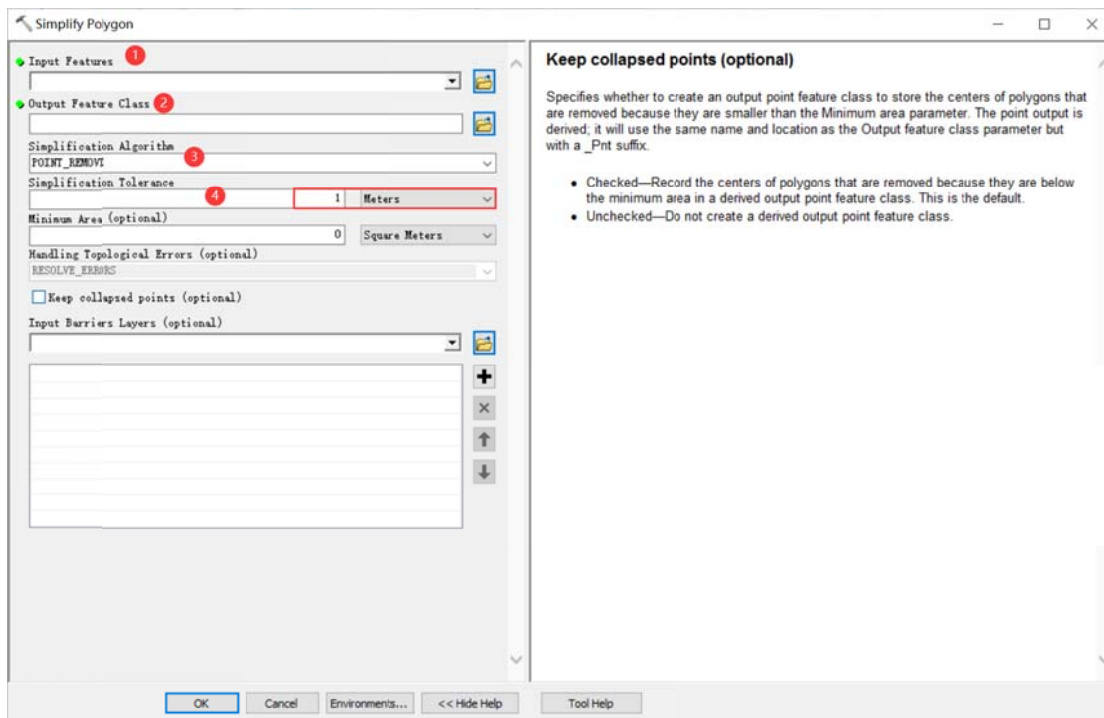
Figure S1. Input and option for Simplify Polygon in ArcGIS.

2. Calculating the total number of nodes using ArcGIS (**Figure S2**):
- Add a new field in the attribute table of dataset.
- Open Field Calculator.
- Switch the parser to python mode, and enter the following code "!shape.pointcount!" in the blue box to calculate the total number of nodes for each glacial lake boundary.
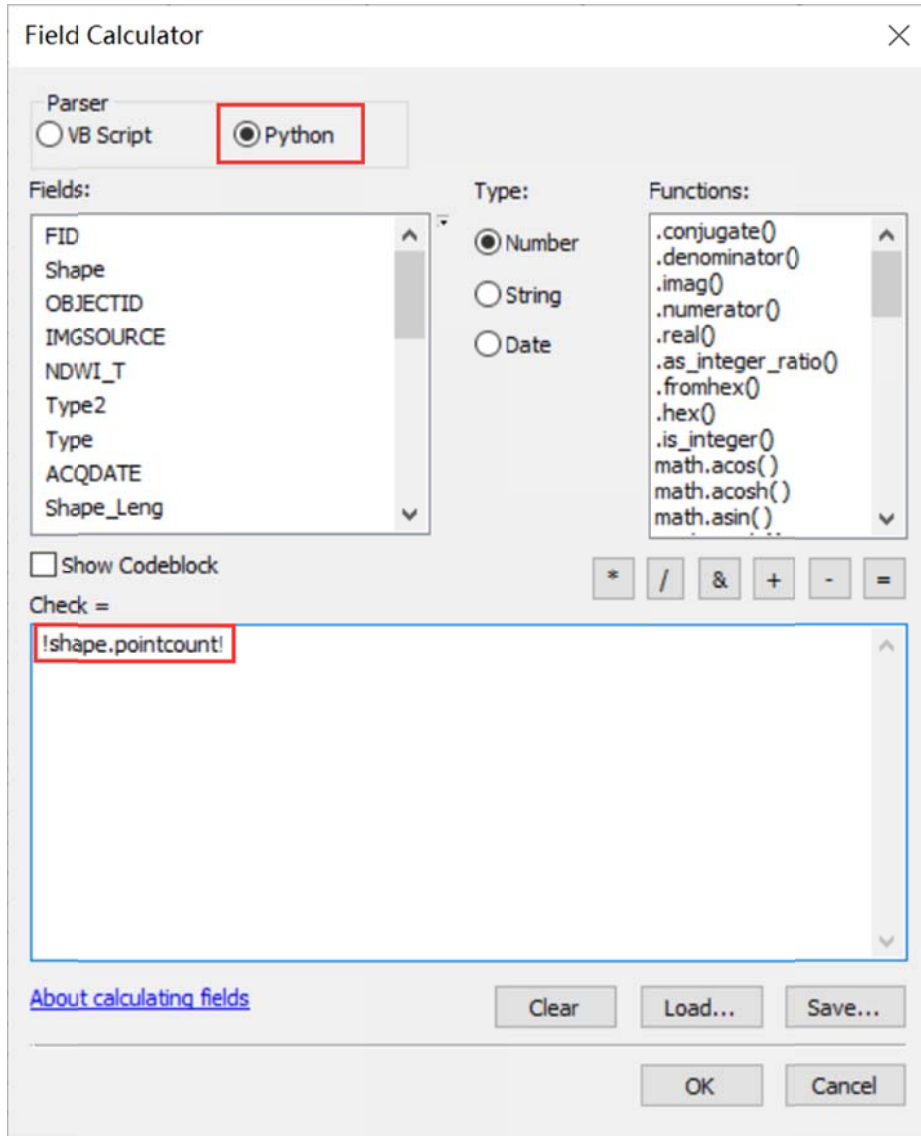
Figure S2. Total node calculation in ArcGIS.

3. Calculating the number of inner nodes:

For polygons without islands (Figure S3), use the following equation:

$$N_{inner} = (\frac{N_{Total}-4-1}{2})$$ (4)

N_{inner} is the total number of inner nodes. N_{Total} is the total number of nodes, including both the outer and inner. An inner node is a polygon vertex where the interior angle surrounding it is greater than 180 degrees. An outer node is the opposite of the inner node, where the interior angle is less than 180 degrees. We found that the outer nodes are usually four more than the inner nodes in our glacial lake dataset. The total nodes in ArcGIS contain one overlapping node to close the polygon, meaning the endpoint is also the startpoint. This extra count was deleted in the calculation (equation 4).
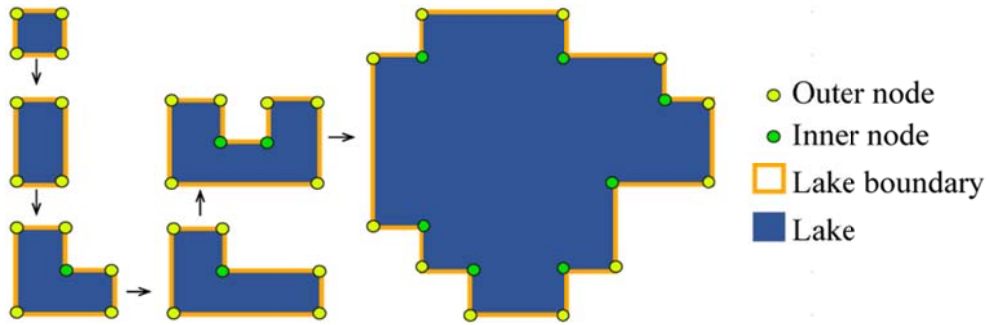
Figure S3. Sketch of outer and inner nodes of various glacial lakes without island.

For polygons with island (**Figure S4**) use the following equation:

$$N_{inner} = (\frac{N_{Total} - (N_{island} + 1) \times 5}{2})$$ (5)

N_{inner} is the number of inner nodes, N_{Total} is the total number of both nodes, and N_{island} is the number of islands within the polygon.
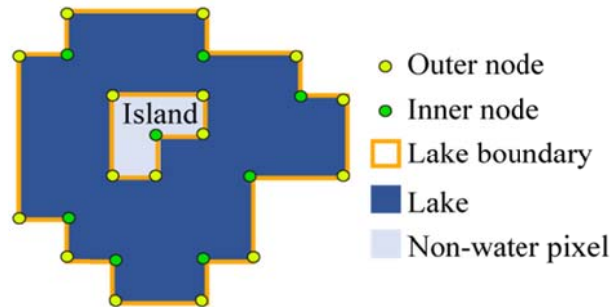
Figure S4. Sketch of outer and inner nodes for glacial lake with island.

We further specify the steps below to help implement equation 5.

Sept 1: detect the number of islands within each polygon.
- Convert the initial lake polygon to polyline using the "Feature To Line" tool (**Figure S5**).
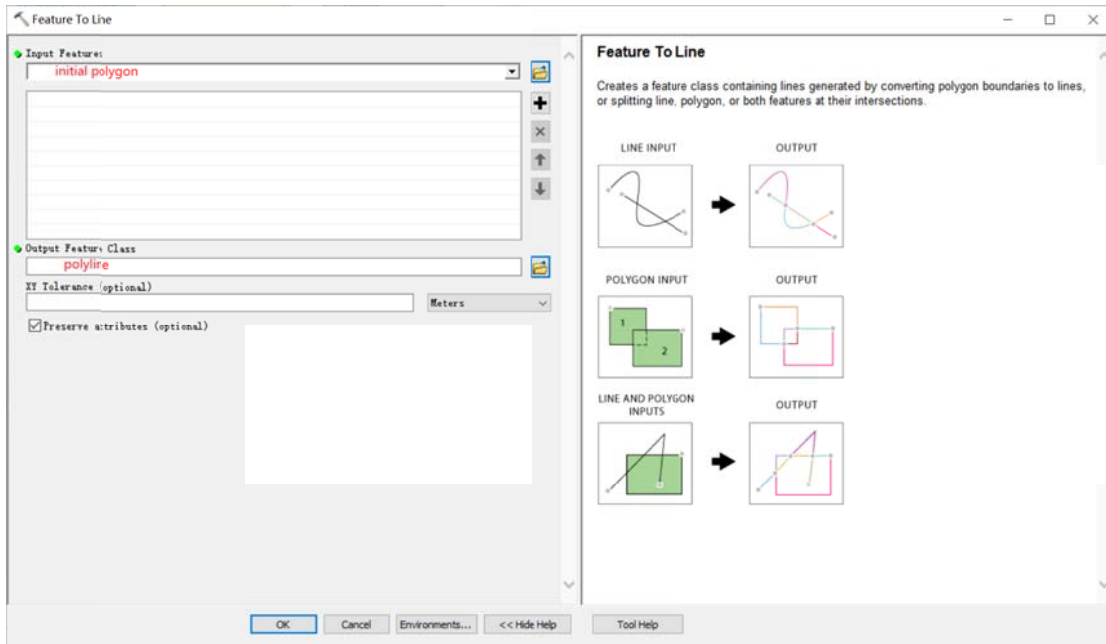
Figure S5. Feature To Line tool in ArcGIS

- Convert the polyline to generate a new polygon (**Figure S6**).
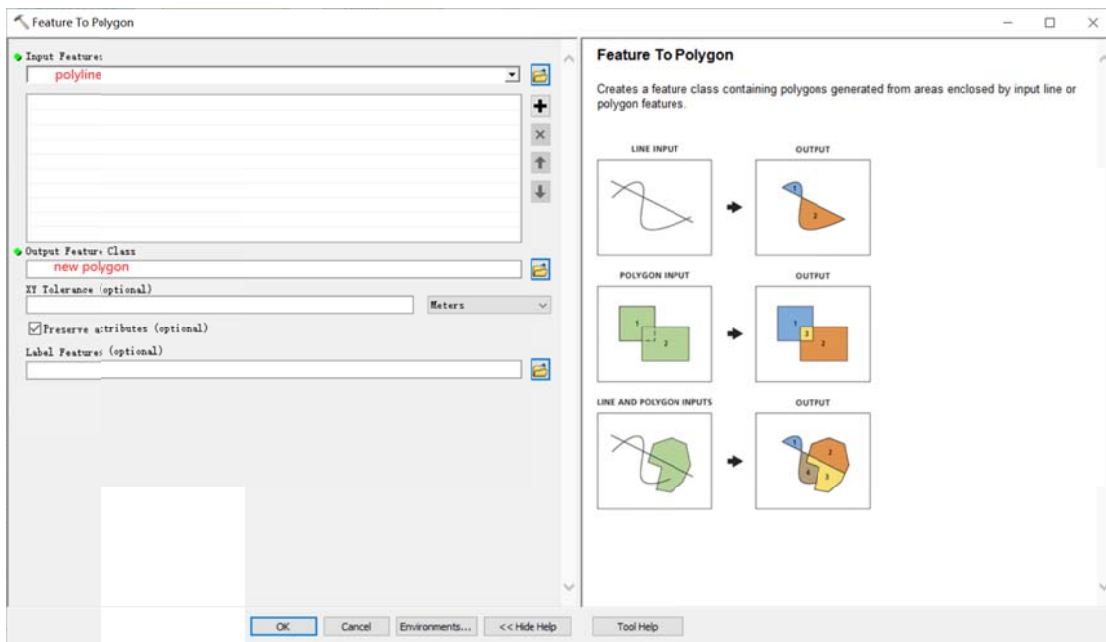
Figure S6. Feature To Polygon tool in ArcGIS

- Erase the new polygon by the initial polygon, which outputs the islands. Then we can count how many islands there are in each lake (**Figure S7**).
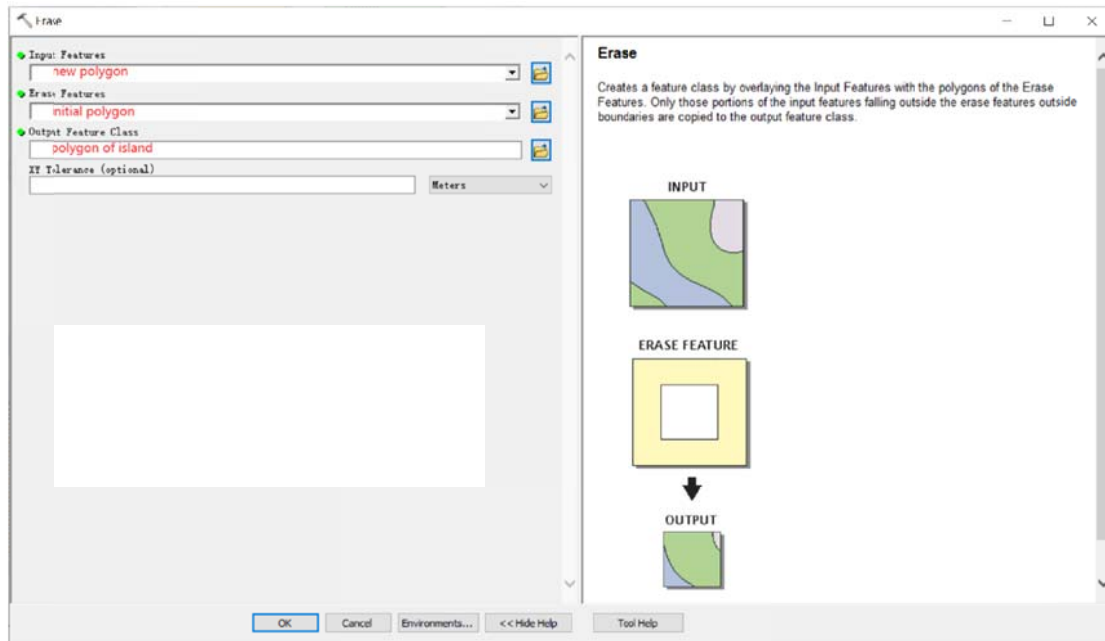
Figure S7. Erase tool in ArcGIS.

Step 2: calculate the number of inner nodes for each polygon with island using equation 5.

4. Calculating the uncertainty of lake mapping using equation 3.