1. Algorithm Flow Chart

**Figure DR1:** Flow chart of the algorithm to further clarify the calculation scheme.
2. Reconstruction of the Pre-Glacial Alpine Topography

![Simplified tectonic map of the Alps from Schmid et al. (2004)](image)

**Figure DR2:** Simplified tectonic map of the Alps from Schmid et al. (2004)

2.1. Experiment Setup

We used re-sampled SRTM data (Farr et al. 2007) to calculate the pre-glacial topography. The best trade-off between resolution and computational time was found for horizontal resolutions of ~540 m and ~780 m in the longitudinal and latitudinal directions, respectively. Concavity indices for glacial and non-glacial catchments for the present-day topography are statistically indistinguishable, averaging 0.73 and 0.71, respectively (Norton et al. 2010). We expect pre-glacial concavity to be lower than the modern value and set the parameters $m$ and $n$ equal to 0.5 and 1, respectively (i.e. reference $\theta$ equal to 0.5). Hillslope processes limit channel gradients upstream of a critical drainage area, $A_c$, and Flint’s law is only valid below this point (Stock and Dietrich 2003). 70% to 90% of topographic relief, however, is represented by fluvial channels where Flint’s law is valid (Whipple 2004), and when the spatial resolution is in the order of hundreds of meters, as the case of this work, debris flow and channel nodes are indistinguishable (e.g. Beaumont et al. 1992; Chase 1992). The base level (Figures DR3, DR4) is traced along the perimeter of the mountain range and follows the boundaries between the Alpine tectonic units and the peri-alpine sedimentary cover. Where this boundary cannot be identified (e.g., between the Alpine-Dinarids and Alpine-Apenninic domains) the base level was simply defined to form a closed polygon following, as far as possible, the boundaries between other tectonic units. The ordering of the nodes is based on the CASCADE algorithm (Braun and Sambridge 1997), and the top-of-the-stack nodes were chosen as channel head nodes (Figure DR3). We only select channel heads lying above the mean topography (calculated through a ~20x20 km sliding window) plus 1σ standard deviation, to avoid taking into account channel head nodes that were most likely to be eroded by glacial activity (Figure DR4). The selected channel heads (a total of 54500 channel head nodes over 439353 nodes in the full DEM) match closely with crest lines and reconstructions of Last Glacial Maximum (LGM) nunataks (Kelly et al. 2004; Schlüchter et al. 2009), which gives support to the assumption that glacial erosion did not impact these elevations (Figure DR3). The isostatically driven rock uplift induced by enhanced erosion during the glacial period was calculated using spatially uniform crust and mantle densities (2.7 and 3.3, respectively). Poisson’s ratio was set to 0.25
(Christensen 1996) and the elastic thickness of the crust was set to 20 km (Stewart and Watts 1997).

Figure DR3: Location of the base level (white line) and channel head nodes (black dots). The inset shows the Zermat-Saas catchment and surrounding areas during the LGM as mapped by Schlüchter et al. (2009, reproduced by permission of swisstopo-JA100120) and the channel head nodes used. Note that these match crest line and nunataks which were independently determined.

Figure DR4: Illustration of selection process of the channel head nodes. The present-day topography, the profile locations (white straight lines) and the base level (white outline) are shown in the map. In the profiles, the solid and dashed lines represent the present-day and mean topography, respectively. Dotted lines are one standard deviation away from the mean topography. Gray circles represent the selected channel heads nodes. The mean topography is calculated using a ~20 by 20 km sliding window.

2.2. Optimization to Data Fit

Several random seeds were tested to ensure that the final solution is not dependent on the intrinsic function of Fortran 90 used to generate pseudorandom numbers.
The residual misfit between present-day and modeled channel head elevations increases with increased length scale (Figure DR5a). This is a classic trade-off curve reflecting our ability to fit all the data at low values of the length scale and smoothing the solution at larger length scales. Because the data contain errors, we prefer some smoothing of the solution and so select a value near the corner of the curve (e.g., Soldati et al. 2006). We chose ~30 km as the preferred length scale for $k_s$ in the Alps. The results of the experiment with a length scale of ~30 km are shown in Figures DR5b-DR5c.

**Figure DR5:** a) The residual misfit of different experiments is plotted against length scale. b) First part of the calculation procedure for the experiment with a length scale of approximately 30 km, where the spatially uniform $k_s$ minimizing the residual misfit between modeled and present-day channel head elevations is sought. c) Second part of the calculation procedure for the experiment with a length scale of 30 km where, at each iteration, the best fitting spatially uniform $k_s$ is perturbed to further reduce the residual misfit at the channel heads. $NN$ is the number of nodes in the considered grid.

Finally, to establish how well each modeled parameter (i.e., pre-glacial $k_s$ values) is resolved with respect to data points (i.e., channel head nodes), we show in Figure DR6 the normalized standard deviation ($1\sigma$) of the variation of residual misfit produced by a number of perturbations (within ±100 m) of the calculated pre-glacial $k_s$ at each node.

**Figure DR6:** Resolution test of the preferred model. Each calculated pre-glacial $k_s$ value is perturbed within a range of ±100 m; the normalized standard deviation ($1\sigma$) of the variation of residual misfit produced by these modulations defines how well each calculated pre-glacial $k_s$ is resolved with respect to the channel head nodes. Black dots within the base level represent the selected channel head nodes.
3. Exploring the Sensitivity of the Reconstructed Pre-Glacial $K_s$ to Potential Lowering of Channel Head Elevations by Glacial Erosion and Reference Concavity Index, $\Theta$

Figure DR7: Reconstructed pre-glacial $k_s$ map obtained assuming an arbitrary glacially-induced decrease of channel head elevations equal to a) 25% and b) 50% of the current height. Other parameters are the same as those of the preferred model. The magnitude of pre-glacial $k_s$ appears to be sensitive to potential lowering of the channel heads by glacial erosion, but the pattern is not (see also Figure 1b).

Figure DR8: Reconstructed pre-glacial $k_s$ with a reference concavity index, $\Theta$, equal to 0.4 (a) and 0.7 (b). Other parameters are the same as those of the preferred model. While the magnitude appears to be highly sensitive to the reference concavity index, the overall pattern remains similar.

REFERENCES CITED


