Geosci. Model Dev. Discuss., 5, 1041–1076, 2012
www.geosci-model-dev-discuss.net/5/1041/2012/
doi:10.5194/gmdd-5-1041-2012

Geoscientific Model Development Discussions

TopoSUB: a tool for efficient large area numerical modelling in complex topography at sub-grid scales

J. Fiddes and S. Gruber

Glaciology, Geomorphodynamics & Geochronology, Department of Geography, University of Zurich, Switzerland

Received: 20 March 2012 – Accepted: 11 April 2012 – Published: 2 May 2012

Correspondence to: J. Fiddes (joel.fiddes@geo.uzh.ch)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Mountain regions are highly sensitive to global climate change. However, large scale assessments of mountain environments remain problematic due to the high resolution required of model grids to capture strong lateral variability. To alleviate this, tools are required to bridge the scale gap between gridded climate datasets (climate models and re-analyses) and unresolved (by coarse grids) sub-grid mountain topography. We address this problem with a sub-grid method. It relies on sampling the most important aspects of land surface heterogeneity through a lumped scheme, allowing for the application of numerical land surface models (LSM) over large areas in mountain regions. This is achieved by including the effect of mountain topography on these processes at the sub-grid scale using a multidimensional informed sampling procedure together with a 1-D lumped model that can be driven by gridded climate datasets. This paper provides a description of this sub-grid scheme, TopoSUB, as well as assessing its performance against a distributed model. We demonstrate the ability of TopoSUB to approximate results simulated by a distributed numerical LSM at around 10^4 less computations. These significant gains in computing resources allow for: (1) numerical modelling of processes at fine grid resolutions over large areas; (2) extremely efficient statistical descriptions of sub-grid behaviour; (3) a "sub-grid aware" aggregation of simulated variables to course grids; and (4) freeing of resources for treatment of uncertainty in the modelling process.

1 Introduction

Mountain regions extend over a large portion of the global land area and significantly influence climate as well as human livelihoods (Barnett et al., 2005; Gruber, 2012; Immerzeel et al., 2010). Complex topography in mountain regions causes high lateral variability of the surface-atmosphere boundary by: (a) altering the local energy and mass fluxes between the ground and the atmosphere (caused by e.g. air temperature,

GMDD
5, 1041–1076, 2012

shading, precipitation gradients); and (b) influencing subsurface properties (e.g. dry bedrock in steep slopes, fine sediments and abundant water in valleys). Mountain environments are currently undergoing rapid and significant change worldwide due to global changes in the earth's climate e.g. warming mountain permafrost (Harris et al., 2003; Isaksen et al., 2001); retreat of mountain glaciers (Paul et al., 2007; Zemp et al., 2006; Barry, 2006); and reduction of snow cover in many regions (Laternser and Schneebeli, 2003; Mote et al., 2005). In order to understand the impact of these changes, both now and under a future climate, tools are required to enable numerical modelling of physical processes that occur across a range of spatial scales.

Global climate models (GCM) and regional climate models (RCM) are able to generate continuous and physically consistent fields of climate variables for the observational period and for scenarios of future climatic conditions. However, the coarse grids such models operate on (~10–500 km) limit the ability to resolve the interactions and feedbacks between the land surface and climate systems in complex topography, which is characterised by strong variability and non-linear processes at the sub-grid scale (Giorgi and Avissar, 1997). A wealth of surface models exist which are capable of simulating processes in mountain regions on fine grids (~1–100 m) and can be driven by coarse grid data with suitable regionalisation techniques (e.g. Bartelt and Lehning, 2002; Gruber et al., 2004; Klok and Oerlemans, 2002; Paul and Kotlarski, 2010)), however the strong fine-scale variability of mountain systems (cf., Gubler et al., 2011; Riseborough et al., 2008) precludes the application of distributed numerical models over large areas. Therefore, the problem remains that despite the near-global availability of high-resolution digital elevation models, global climate datasets and suitable numerical methods, land surface processes in complex topography remain poorly quantified in many aspects (Fig. 1).

This problem of scale has been previously approached through various forms of sub-grid parameterization. This term can be defined as capturing the spatial variability of a modelled process at a suitable resolution, while reducing the demands for data and computation, by approximating its fine-scale distribution at a lower resolution (e.g.

GMDD
5, 1041–1076, 2012

Hebeler and Purves, 2008; Giorgi and Avissar, 1997). Previous sub-grid approaches can be broadly classified as either discrete mosaic types or continuous probability density function (PDF) schemes (Wood et al., 1988; Avissar, 1991; Giorgi and Avissar, 1997). This distinction can also be conceptualised as the modelling of sub-grid *inter*-patch and *intra*-patch heterogeneity (Giorgi and Avissar, 1997). In mosaic approaches, a number of homogeneous subregions ("tiles") are defined at the sub-grid scale, each with its own energy, momentum and water budget. The surface flux calculations are computed separately for each tile. Aggregation to the coarse grid is performed by averaging over the tiles which are weighted by their fractional cover (Avissar and Pielke, 1989; Koster and Suarez, 1992). Models differ on how these tiles are discretised. For example, Seth et al. (1994) and Dimri (2009) used a regularly spaced finescale sub-grid. Alternatively, a series of discrete classes based on surface vegetation type (Avissar and Pielke, 1989; Koster and Suarez, 1992) or topographical elevation (Leung and Ghan, 1995) have been used. Kotlarski (2007) developed a dynamic mountain glacier sub-grid parametrization for inclusion in RCM's which explicitly accounts for run off generation and adjusts glacier area (dynamic tile) based on accumulation/ablation conditions. PDF-based approaches attempt to describe the variability of sub-grid characteristics through analytical or empirical distribution functions (Avissar, 1991; Famiglietti and Wood, 1994; Liang et al., 2006). This is based on the assumption that surface characteristics as well as climatic forcings vary according to distributions that can be approximated by the given PDF. This approach then explicitly calculates gridbox average surface fluxes for a surface variable distribution using numerical or analytical integration over the appropriate PDF. Walland and Simmonds (1996) used sub-grid statistics (variance, kurtosis) of distributions of topographical parameters to improve the simulation of the snowpack in GCMs.

Mountain regions exhibit more relevant dimensions which control land surface processes than flat areas (e.g. elevation, aspect, slope, etc.). This means that a simple mosaicing of the land surface is not appropriate and a more sophisticated technique is needed to account for this heterogeneity. While, statistical models exist (e.g. Boeckli

GMDD
5, 1041–1076, 2012

et al., 2012) which provide good 2-D representation in complex topography, they usually do not resolve transient changes.

In this paper we describe a sub-grid method which samples the most important aspects of land surface heterogeneity based on input predictors (PREDs) which describe important dimensions of variability in complex topography. A lumped scheme then allows for the efficient application of numerical land surface models (LSM) over large areas. Aggregation of simulated target variables (TVs) to the coarse grid and spatialisation to the fine grid is achieved through a statistical technique based on membership functions. We use cumulative distribution functions (CDF) of TVs to provide rapid aggregated statistics of sub-grid behaviour over large areas. This Topographic SUBgrid tool (TopoSUB) allows for: (1) modelling of processes at fine grid resolutions, (2) efficient statistical descriptions of sub-grid behaviour, (3) a "sub-grid aware" aggregation of simulated TVs to coarse grids and (4) enables validation of results with fine-scale ground truth. The strength of the scheme is its ability to compute numerical simulations by several orders of magnitude faster than a distributed model and therefore free resources for spatially or temporally expensive simulations as well as for exploring uncertainties in input data (e.g. climate projections) or LSM parameters and physics. While we acknowledge a lumped approach compromises on 2-D representation (e.g. snow redistribution, surface runoff) it enables the application of sophisticated 1-D physics over large areas.

This paper provides a proof of concept of this tool by describing the method, providing guidance on parameter selection and performing validation experiments against baseline distributed model simulations. Whilst TopoSUB is designed as a tool for use in complex topography – it is thought the concept may also be of interest to communities outside of mountain environment related disciplines where alternative dimensions of variability are important.

GMDD
5, 1041–1076, 2012

2 Generic Methods

2.1 *K*-means clustering

Samples are formed using the *K*-means clustering algorithm of Hartigan and Wong (1979), an unsupervised learning algorithm which is a suitable for clustering multidimensional data. *K*-means aims to partition all points into *K* clusters such that the total sum of squares (or squared deviations) from individual points (pixels) to the assigned cluster centroids in multivariate attribute space is minimised. This then represents an optimal clustering of points in attribute space for prescribed number of samples. The algorithm is composed of the following steps:

1. Randomly place *K* points into the space represented by the objects that are being clustered. These points represent initial group centroids/seeds.
2. Assign each object to the group that has the closest centroid.
3. When all objects have been assigned, recalculate the positions of the *K* centroids.
4. Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into *K* clusters from which the metric to be minimized can be calculated.

Although it can be proven that the procedure will always terminate, the *K*-means algorithm is sensitive to the initial configuration of cluster seeds and does not necessarily find the optimal configuration (Kanungo et al., 2002), corresponding to the global objective function minimum (as opposed local minima). The *K*-means algorithm is run multiple times to reduce this effect.

The *K*-means algorithm has three important controlling parameters, *K* number of clusters, *iter.max* maximum allowed iterations of the algorithm and *nstart* number of random starts. Fixing *K* at 128 samples, a sensitivity analysis was performed on iter.max and nstart in order to define baseline parameter values to be used in TopoSUB. Stable performance, as measured by within sum of squares (WSS) was achieved

GMDD
5, 1041–1076, 2012

for iter.max = 20 and nstart = 10 (Fig. 2). Iter.max shows a much greater variation in WSS over its range as it is an intrinsic part of the K-means algorithm. K-means is significantly sensitive to initialisation i.e. location of initial cluster seeds. For this reason it is highly recommended to run the model with several random starts and average the results. A first run of K-means algorithm is performed on a sample of input data (10^5 pixels) with 10 random starts and maximum iterations set to 20 (as previously defined). The cluster centres defined by this training dataset are used to initialise K-means for the entire dataset (10^6 or more pixels). This allows for significant speed up (factor of 10, Table 1) while not compromising on quality of results.

2.2 Fuzzy membership

In contrast to crisp membership (yes/no), fuzzy methods allow for varying degrees of membership to multiple sets (Zadeh, 1965). A membership function is given valued in the interval 0–1. We apply fuzzy membership functions to allow for varying degrees of membership of pixels to multiple samples. Fuzzy methods of classification are usefully applied to multivariate classification problems when class overlap is required to represent continuous phenomena. The primary advantage of this method over crisp classification is that fuzzy methods allow for high resolution mapping of the TV, whilst accounting for topographic variability within each sample. Membership functions are calculated in two steps. First, the standardised squared distance (d^2) of the i-th pixel from the n-th sample centroid C of the j-th attribute is determined by:

$$d_{ni}^2 = \sum_{j=1}^{v} [(x_{ni} - C_{nj})/\mathrm{Sd}_{nj}]^2. \quad (1)$$

where Sd is class standard deviation (Burrough et al., 2001). Then we can derive the membership, μ of the i-th pixel to the n-th sample using the formula of Sokal and Sneath (1967):

GMDD
5, 1041–1076, 2012

$$\mu_{ni} = {d_{ni}^2}^{-1/(M-1)} / \sum_{n=1}^{K} {d_{ni}^2}^{-1/(M-1)}, \quad (2)$$

$$n = 1,2,...,K,$$

$$i = 1,2,...,$$

$$\mu_{ci} \in [0,1].$$

where M is the so called fuzzy exponent parameter. This parameter is a weighting exponent and it controls the degree of fuzziness of the membership grades. As M approaches 1, the clustering becomes crisper. As M becomes very large (i.e. $M > 100$), membership becomes almost constant so that clusters can no longer be distinguished. The scheme is sensitive to this value as it effectively changes the scheme from a crudely spatialised lumped model (all members of cluster receive same TV value) to a more intelligently spatialised scheme that accounts for topographic variability within each cluster.

The determination of an optimal value for M in a fuzzy classification process remains an open question (e.g. Okeke and Karnieli, 2006; Burrough et al., 2000)). Commonly cited values in the literature range from 1.3 to 3 depending upon application. Okeke and Karnieli (2006) propose a linear mixture model approach to optimise the value of M for any given dataset. We found an optimum value to exist between crisp ($M = 1$) and very fuzzy ($M > 2$) by iterating through values of M 1–2 in stages of 0.1. We identified 1.4 to give the most accurate results in majority TVs tested (compared with a distributed simulation). However, the optimal value of M is not the focus of the present study and will not be further discussed. A standard fuzzy membership algorithm will compute memberships for all pixels to all samples, resulting in a membership matrix of n samples by p pixels. Even with the modest large area simulation we give in this paper, this results in matrix of magnitude 10^8. In order to reduce storage demands we allow a reduced number of membership dimensions to be prescribed whilst preserving the functionality of the algorithm.

GMDD
5, 1041–1076, 2012

2.3 Statistical methods

Test distributions are compared with baseline distributions using a one-sample Kolmogorov-Smirnov test (KS-test) which is a non-parametric test for the equality of continuous, one-dimensional probability distributions that can be used to compare a sample with a reference probability distribution. If the sample comes from distribution $F(x)$, then the KS statistic, D converges to zero. Comparison of simulated fine grid data against baseline grids is performed using the normalised root mean squared error (NRMSE) to allow comparison across simulated TVs of different scale ranges. The RMSE is normalised against the standard deviation of the observations, being a robust statistic that is less influenced by outliers than the range. Correlation statistics stated are Pearson product-moment correlation coefficients (r-value). CDFs are computed according to the empirical cumulative distribution function in R statistical software.

3 TopoSUB methods

TopoSUB is designed to provide an effective approximation of a spatially distributed grid with a lumped model. A key design principle of the tool is to be generic, allowing for choice of driving inputs (station data, gridded climate data), numerical model and output (TVs, resolution) allowing for a wide range of possible applications. It has two main modules (Fig. 3): a pre-processor to run only once and a post-processor to run many times together with an LSM. Because it is intended for use in mountain areas it needs to be able to accommodate more dimensions of variability than is usual in mosaicing techniques used to partition the sub-grid. Besides differing surface and subsurface properties, the effects of elevation, slope exposition, slope angle and horizon elevation are likely to be of importance. To allow the scalable use of this scheme, i.e. its application over large mountain ranges, it should employ repeatable and robust methods to form sub-grid samples. Because the influence of predictor variables (dimensions of sub-grid variability that are accounted for) is not known a priori and may

GMDD
5, 1041–1076, 2012

change laterally, a method for informed sampling is required in which the importance of predictor variables on the simulated target variable(s) is evaluated. Model parameters are given in Table 2 together with default values.

3.1 Pre-processor

The pre-processor configures the sub-grid by creating samples for later simulation in the LSM and by determining the membership functions of original 2-D pixels to those samples. For a given experimental domain this module need only be run once. The scheme can use several PRED variables from which the sub-grid scheme is constructed and usually these include DEM-derived land surface parameters.

3.1.1 Training routine

The input PREDs are initially clustered using the K-means algorithm to form a predefined number of samples. Scaling of PREDs is important because it affects the relative number of samples formed along a given dimension. With the scaling of PREDs we thus influence how finely samples are resolved in which direction, a process that is important to optimize the number of samples. No scaling could result in e.g. differing results with elevation provided in units of meters or kilometres. At this initial stage, a simple scaling is thus applied which normalises all PREDs to a standard scale under the assumption that all are of equal importance with respect to simulated TV. Sample centroids define the topographic and environmental input to the LSM that is run for this initial set of samples in a training simulation.

3.1.2 Sample formation by informed clustering

Based on the training routine results, one linear regression model is constructed for each of n TVs using the whole set of i PREDs as regressors (Eq. 3) using generalised least squares (GLS). GLS is able to handle PREDs with non-normal distributions and/or which are partially correlated and therefore is more robust when implemented as an

GMDD
5, 1041–1076, 2012

automated method. GLS minimises the squared Mahalanobis distance as opposed the residual sum of squares, as in regression methods using ordinary least squares.

$$\mathrm{TV}_n = \mathrm{PRED}_1\beta_1 + \mathrm{PRED}_i\beta_i + ..., \quad \mathrm{TV} = 1...n, \quad \mathrm{PRED} = 1...i. \tag{3}$$

The resulting regression coefficients, β_i provide an informed scaling of the PRED_i for the clustering algorithm by transforming them into equivalents of the TV_n dimension and unit:

$$\mathrm{PRED}_{i,\mathrm{scaled}} = \mathrm{PRED}_i \cdot \beta_i. \tag{4}$$

Normalized to a sum of one, these coefficients can be averaged (β_{mean}) to accommodate more than one TV. Parameter W_{TV} is an optional weighting function for individual TVs if extra information exists to justify higher weighting of a TV in the averaging process:

$$\beta_{\mathrm{mean}} = \sum_{\mathrm{TV}} \left(\frac{\beta}{\sum_{\mathrm{PRED}} \beta_{\mathrm{PRED}}} \right) . W_{\mathrm{TV}}, \quad \mathrm{TV} = 1, ...n. \tag{5}$$

The set of PREDs, with informed scaling, are re-clustered to provide a predefined number of samples that are then more effectively distributed with respect to the desired TVs. The sample centroids now provide the required input to the LSM. Additionally, the regression model can be optimised by using disaggregated r^2 (Genizi, 1993) either manually or automatically, by removing PREDs that contribute less than a stated threshold to model significance.

GMDD
5, 1041–1076, 2012

3.1.3 Pre-processor output

A vector of sample weights is calculated according to total membership of pixels to each sample. This provides the means by which to (a) aggregate TVs to the coarse grid and, (b) provide a rapid statistical description of sub-grid behaviour using a CDF. A matrix of membership functions of individual pixels to samples provides a means of spatialising results to the fine grid. We employ crisp and fuzzy membership. Crisp membership implies that pixels may belong to only one sample (that which they are assigned in clustering). Fuzzy membership allows for varying degrees of membership to multiple samples. This then accounts to some extent for within sample variance (in terms of PREDs) that inevitably exists and provides an alternative and smoother means of spatialisation at reasonable computational cost.

The final output from the pre-processor is the sub-grid configuration (Fig. 4) as defined by: (A) a small matrix of sample characteristics. These are the environmental characteristics of each sample used to drive the LSM as well as the aggregated sample weights used for aggregation to grid level and statistical description of sub-grid behaviour. (B) Membership information for the spatialisation of results to the fine grid. For crisp membership this has the dimensions of the original fine grid (B1), for fuzzy membership the ID and weight for the s most important samples are stored for each original pixel (Table 2), increasing the dimensions of this information to 2*s times the original size (B2). Based on the sub-grid configuration, the LSM is run in 1-D mode for each sample.

3.2 Post-processor

Based on the sub-grid configuration, the LSM output is post-processed. The result of this can either be (a) summary statistics with respect to the coarse grid describing its sub-grid variability by a CDF or derived quantities; or (b) data spatialised to the original fine grid; or (c) data estimated for a list of discrete points to support validation studies using ground truth data. The coarse-grid summary statistics are computed according

GMDD
5, 1041–1076, 2012

to the aggregated membership functions of individual pixels to each sample. TVs are spatialised to fine grid resolution according to the membership functions (crisp or fuzzy) of each pixel to each significant cluster. This accounts for the sub-grid heterogeneity that exists between cluster centroids.

4 Simulation experiments

4.1 Data and tools

The high-resolution input into the scheme is a 25 m digital elevation model (DEM, obtained from Swisstopo). The DEM used in this study covers a test area of 25.6 × 25.6 km (~10^6 pixels) in the south-east of Switzerland (Fig. 5). The study area represents a good example of strongly variable mountain topography (elevation range 1556–4043 m a.s.l.) with which to test the performance of the scheme. Land surface parameters, slope (degrees), aspect (degrees) and sky view factor (fraction, 0–1) are derived from the DEM using SAGA-GIS. Driving meteorology is provided by a high elevation MeteoSwiss synoptic weather station (Corvatsch) at 3315 m a.s.l. We use an hourly time series of air temperature, relative humidity, global radiation, wind speed, wind direction and precipitation over the period 1 July 2009 to 1 July 2010. We employ the open-source LSM GEOtop (Endrizzi and Marsh, 2010; Dall'Amico et al., 2011; Rigon et al., 2006) which is a physically based model that simulates the coupled energy and water balance with phase change in soil, a multi-layer snow pack and surface energy fluxes in 1-D and distributed 2-D modes, which has been designed specifically for application in mountain regions. The model is run on an hourly timestep. We apply two years of spin up and then generate 1 yr of data. All tools were programmed using the open-source statistical software R.

GMDD
5, 1041–1076, 2012

4.2 Testing strategy

The primary aim of this evaluation is to test how well TopoSUB is able to reproduce results of a distributed model. This is done by comparing results obtained from TopoSUB (SUB) with results from a distributed LSM simulation on a regular 2-D grid (BASE). Both runs use the same LSM, GEOtop. The distributed runs have diverse spatial resolutions ranging from 25 m (10^6 cells) to 25 km (1 cell). The lumped sub-grid model is also run at differing resolution i.e. different levels of detail (number of samples) in the sampling of the sub-grid. The sub-grid scheme is evaluated with respect to the 25 m distributed simulation in all experiments, which we consider to be the baseline in this experiment, being the most finely discretised representation of the experiment domain. Experiments presented in results are: (Sect. 5.1) grid aggregated results compared directly with corresponding statistics of BASE. For this, the mean and standard deviation as well as the 25th and 75th percentiles are calculated; (Sect. 5.2) statistical description of sub-grid behaviour is evaluated by comparing the CDF of SUB simulation to CDF of BASE simulations using KS-test; and, (Sect. 5.3) fuzzy spatialised results are compared to BASE using the r-value and NRMSE to assess the predictive power of the scheme at the grid cell level. Finally, other aspects of TopoSUB performance are presented (Sects. 5.4–5.5). To keep computation times for the 25 m resolution BASE simulation reasonable, water movement in the soil was not considered.

4.3 Model settings

For testing, input PREDs of elevation, aspect, slope and sky view factor are used in the clustering algorithm, all computed from the input DEM at 25 m resolution. In all experiments we simulate output response variables of air temperature (Tair, °C), ground surface temperature at 10 cm depth (GST, °C), snow water equivalent (SWE, mm) and incoming shortwave radiation (SWin, W m^{-2}). These were chosen as suitable variables with which to test the performance of the scheme in terms of representing both surface processes and energy fluxes. Air temperature represents a simple check of the scheme

GMDD
5, 1041–1076, 2012

due to straightforward relationship with elevation in this study (using a standard lapse rate). SWin represents how well the scheme is able to represent topography with respect to parameters relevant to radiation modelling. Snow water equivalent and GST both represent important physical processes in mountain areas. TVs are analysed as mean annual values in all cases. Model parameters are set per default values in Table 2.

5 Results

5.1 Grid aggregated output

Grid aggregated results are analysed as deviations from the baseline distributed grid for increasing sample numbers (lumped scheme: samples, gridded scheme: pixels) that represent computational cost (Fig. 6). The simulated TVs approximate the high-resolution distributed results well requiring 10^3–10^4 times less computations. The convergence of results for the tested TVs at 100–200 samples in the sub-grid scheme suggests that we are able to reach a stable performance. A convergence of results with resolution is not observed in distributed simulations (except for the simple variable of air temperature) underscoring the importance of attention to scaling issues. Figure 6 also shows the improvement in aggregated information between a grid average computation (1 sample), as is common in climate models, and a 200 sample SUB simulation, that is capable of approximating the BASE simulation (which is explicitly modelling sub-grid processes).

5.2 Statistical description of sub-grid behaviour

CDFs are calculated using aggregated sample weights and provide a rapid description of sub-grid behaviour by presenting the distribution of simulated TVs without the need for spatialisation. This technique enables rapid assessments such as percentage of permafrost in the simulation domain or the total quantity of SWE. A good fit is seen

GMDD
5, 1041–1076, 2012

in all cases (Fig. 7). These statistics can be readily presented against topographic attributes to give more detailed understanding of the sub-grid e.g. permafrost extent by elevation band or exposures.

5.3 Spatialised output

Spatialised results are obtained by distributing LSM results based on the fuzzy membership of each pixel, resulting in a 25 m resolution mapping of TVs. Based on convergence of the NRMSE and correlation coefficient of the lumped scheme it can be seen that the majority of performance is gained until 64 samples, after 258 samples a stable performance is achieved (Fig. 8). This represents a reduction of computational effort of three to four orders of magnitude compared with distributed simulation BASE (depending on required quality level), a similar result to Sect. 5.1. Figure 9 gives density scatter plots of all TVs for 258 samples. A good correlation is reported in all cases. The lumped scheme is able to reproduce the distributed simulation with an NRMSE of 12–28 % depending on the TV. Figure 10 provides a visual comparison of the simulation results for GST.

5.4 Model stability

Figure 11 shows results of 40 simulations of TopoSUB at each of 4 resolutions: 25, 50, 100, 200 samples. The purpose being to investigate stability of the tool and any resolution dependency of its stability. Results of deviation of each simulation from mean values of mean and quantiles 25/75 of all 40 simulations indicate reasonable stability even at low resolutions as indicated by a low absolute deviation. A significant increase in stability is seen between 25–100 samples in all variables tested.

5.5 TopoSUB configurations

Figure 12 gives NRMSE for 128 sample SUB simulations for configurations: “fuzzy informed”, “fuzzy simple”, “crisp informed”, “crisp simple”. All TV results improve with

GMDD
5, 1041–1076, 2012

fuzzy membership. All TV results except SWin improve with informed scaling. This is because all TVs are heavily influenced by PRED "elevation", except for SWin. During calculation of β_{mean} SWin specific PREDs (e.g. sine of aspect) are down-weighted in favour of elevation – causing reduced performance for SWin results. This can be corrected for using the parameter W_{TV} (Eq. 5) depending upon user application. These results demonstrate that informed scaling and fuzzy membership are both useful tools to be applied in the pre-processor stage. Time-cost associated with these methods is incurred only once during pre-processing and benefit (improved performance) is received multiple times with LSM output.

6 Conclusions

This paper has described the TopoSUB scheme and has evaluated its performance by comparison with a distributed model, as well as given first indication of parameter choice. We have obtained promising results for the use of TopoSUB as robust and efficient tool for large area numerical simulations of processes in complex mountain topography. The informed sampling procedure is thought to be an appropriate method by which to sample a multidimensional sub-grid space without a priori knowledge of the relative importance of dimensions, with respect to simulated TVs. In the presented case, a sample number of approximately 64 samples proved sufficient to describe the wide range of elevations, slope expositions, slope angle and horizon elevations present in the study area, while 258 samples approximated well the results of a distributed 2-D simulation of 25 m resolution. While spatialisation proves a valuable tool for site specific studies, CDF's of sample results provide a rapid assessment of sub-grid behaviour.

There are obvious limitations which must be acknowledged due to the tool being based on a 1-D configuration, namely lateral mass transfers (such as snow redistribution, although this is commonly neglected even in distributed models) and flow modelling (surface/subsurface); processes which are not easily represented in a 1-D model.

GMDD
5, 1041–1076, 2012

The demonstrated ability of TopoSUB to efficiently approximate a distributed model as a reduced series of 1-D samples has several important applications:

1. Numerical, physically based models are important tools in furthering process understanding, but are inherently prone to a wealth of uncertainties, including initial and boundary conditions, parameterizations, model physics and parameter choice (Beven, 2001). Intensive sensitivity studies are often needed to define parameter values, and uncertainty usually needs to be quantified. TopoSUB would allow both of these tasks to be undertaken with little compromise on quality of result while focusing computing resources on long temporal scale or multiple repeat simulations.

2. TopoSUB is able to approximate the results of of a distributed grid that have been aggregated at coarse grid level. This suggests interesting prospects for the improvement of the representation of mountain fine-scale processes in coarse grid models.

3. TopoSUB can be used together with grid-computing infrastructure that is becoming increasingly common in many disciplines in order to parallelise a numerical model as an array of 1-D tasks.

We envisage TopoSUB to be a useful tool in a wide range of numerical modelling applications in complex terrain, due to flexible choice of inputs, numerical models and output options.

Supplementary material related to this article is available online at: http://www.geosci-model-dev-discuss.net/5/1041/2012/gmdd-5-1041-2012-supplement.zip.

GMDD
5, 1041–1076, 2012

Acknowledgements. This project was funded by the Swiss National Science Foundation grant "CRYOSUB: Mountain Cryosphere Sub-grid parameterization and Computation", project number 200021_121868 and kindly supported by the project X-sense funded through nanotera.ch, project number 200021_211868. This work was also supported by the AAA/SWITCH funded Swiss Multi Science Computing Grid project (www.smscg.ch) with computational infrastructure and support. Customized libraries (ggeotop and GC3Pie) and user support were kindly provided by GC3: Grid Computing Competence Center (www.gc3.uzh.ch). Special thanks go to Stefano Endrizzi for his invaluable support in the implementation of GEOtop in this project.

References

Avissar, R.: A statistical-dynamical approach to parameterize subgrid-scale land-surface heterogeneity in climate models, Surveys in Geophysics, 12, 155–178, 1991. 1044

Avissar, R. and Pielke, R. A.: A parameterization of heterogenous land surfaces or atmospheric numerical models and its impact on regional meteorology), Mon. Weather Rev., 117, 2113–2136, 1989. 1044

Barnett, T. P., Adam, J. C., and Lettenmaier, D. P.: Potential impacts of a warming climate on water availability in snow-dominated regions, Nature, 438, 303–309, 2005. 1042

Barry, R. G.: The status of research on glaciers and global glacier recession: a review, Progress in Physical Geography, 30, 285, 2006. 1043

Bartelt, P. and Lehning, M.: A physical SNOWPACK model for the Swiss avalanche warning:: Part I: numerical model, Cold Reg. Sci. Technol., 35, 123–145, 2002. 1043

Beven, K.: How far can we go in distributed hydrological modelling?, Hydrol. Earth Syst. Sci., 5, 1–12, doi:10.5194/hess-5-1-2001, 2001. 1058

Boeckli, L., Brenning, A., Gruber, S., and Noetzli, J.: A statistical approach to modelling permafrost distribution in the European Alps or similar mountain ranges, The Cryosphere, 6, 125–140, doi:10.5194/tc-6-125-2012, 2012. 1044

Burrough, P. A., van Gaans, P. F. M., and MacMillan, R. A.: High-resolution landform classification using fuzzy k-means, Fuzzy Sets and Systems, 113, 37–52, 2000. 1048

Burrough, P. A., Wilson, J. P., Gaans, P. F. M. V., and Hansen, A. J.: Fuzzy k-means classification of topo-climatic data as an aid to forest mapping in the Greater Yellowstone Area , USA, Landscape Ecology, 523–546, 2001. 1047

GMDD

5, 1041–1076, 2012

Dall'Amico, M., Endrizzi, S., Gruber, S., and Rigon, R.: A robust and energy-conserving model of freezing variably-saturated soil, The Cryosphere, 5, 469–484, doi:10.5194/tc-5-469-2011, 2011. 1053

Dimri, A. P.: Impact of subgrid scale scheme on topography and landuse for better regional scale simulation of meteorological variables over the western Himalayas, Clim. Dynam., 32, 565–574, 2009. 1044

Endrizzi, S. and Marsh, P.: Observations and modeling of turbulent fluxes during melt at the shrub-tundra transition zone 1: point scale variations, Hydrol. Res., 41, 471–491, doi:10.2166/nh.2010.149, 2010. 1053

Famiglietti, J. S. and Wood, E. F.: Multiscale modeling of spatially variable water and energy balance processes, Water Resour. Res., 30, 3061–3078, 1994. 1044

Genizi, A.: Decomposition of r2 in multiple regression with correlated regressors, Statistica Sinica, 3, 407–420, 1993. 1051

Giorgi, F. and Avissar, R.: Representation of heterogeneity effects in earth system modeling - Experience from land surface modeling, Rev. Geophys., 35, 413–438, 1997. 1043, 1044

Gruber, S.: Derivation and analysis of a high-resolution estimate of global permafrost zonation, The Cryosphere, 6, 221–233, doi:10.5194/tc-6-221-2012, 2012. 1042

Gruber, S., Hoelzle, M., and Haeberli, W.: Rock-wall temperatures in the Alps: modelling their topographic distribution and regional differences, Permafrost and Periglacial Processes, 15, 299–307, 2004. 1043

Gubler, S., Fiddes, J., Keller, M., and Gruber, S.: Scale-dependent measurement and analysis of ground surface temperature variability in alpine terrain, The Cryosphere, 5, 431–443, doi:10.5194/tc-5-431-2011, 2011. 1043

Harris, C., Vonder Muhll, D., Isaksen, K., Haeberli, W., Sollid, J. L., King, L., Holmlund, P., Dramis, F., Guglielmin, M., and Palacios, D.: Warming permafrost in European mountains, Global Planet. Change, 39, 215–225, 2003. 1043

Hartigan, J. A. and Wong, M. A.: A k-means clustering algorithm, Journal of the Royal Statistical Society, Series C, Applied statistics, 28, 100–108, 1979. 1046

Hebeler, F. and Purves, R. S.: The influence of resolution and topographic uncertainty on melt modelling using hypsometric sub-grid parameterization, Hydrol. Process., 22, 3965–3979, 2008. 1044

Immerzeel, W. W., Van Beek, L. P. H., and Bierkens, M. F. P.: Climate change will affect the Asian water towers, Science, 328, 1382, doi:10.1126/science.1183188, 2010. 1042

GMDD

5, 1041–1076, 2012

Isaksen, K., Holmlund, P., Sollid, J. L., and Harris, C.: Three deep Alpine-permafrost boreholes in Svalbard and Scandinavia, Permafrost and Periglacial Processes, 12, 13–25, 2001. 1043
Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., and Wu, A. Y.: An efficient-means clustering algorithm: Analysis and implementation, IEEE Transactions on Pattern Analysis and Machine Intelligence, 24, 881–892, 2002. 1046
Klok, E. J. and Oerlemans, J.: Model study of the spatial distribution of the energy and mass balance of Morteratschgletscher, Switzerland, J. Glaciol., 48, 505–518, 2002. 1043
Koster, R. and Suarez, M.: Modeling the land surface boundary in climate models as a composite of independent vegetation stands, J. Geophys. Res., 97, 2697–2715, 1992. 1044
Kotlarski, S.: A Subgrid Glacier Parameterisation for Use in Regional Climate Modelling, Ph.D. thesis, Max Planck Institute for Meteorology, 2007. 1044
Laternser, M. and Schneebeli, M.: Long-term snow climate trends of the Swiss Alps (1931–99), Int. J. Climatol., 23, 733–750, 2003. 1043
Leung, L. R. and Ghan, S. J.: A subgrid parameterization of orographic precipitation, Theor. Appl. Climatol., 52, 95–118, 1995. 1044
Liang, X. Z., Xu, M., Choi, H. I. L., Kunkel, K. E., Rontu, L., Geleyn, J. F., Müller, M. D., Joseph, E., and Wang, J. X. L.: Development of the regional climate-weather research and forecasting model (CWRF): Treatment of subgrid topography effects, in: Proceedings of the 7th Annual WRF User's Workshop, Boulder, CO, 2006. 1044
Mote, P. W., Hamlet, A. F., Clark, M. P., and Lettenmaier, D. P.: Declining mountain snowpack in western North America, B. Am. Meteorol. Soc., 86, 39–44, 2005. 1043
Okeke, F. and Karnieli, A.: Linear mixture model approach for selecting fuzzy exponent value in fuzzy c-means algorithm, Ecological Informatics, 1, 117–124, 2006. 1048
Paul, F. and Kotlarski, S.: Forcing a distributed glacier mass balance model with the regional climate model REMO. Part II: downscaling strategy and results for two swiss glaciers, J. Climate, 23, 1607–1620, doi:10.1175/2009JCLI3345.1, 2010. 1043
Paul, F., Kaab, A., and Haeberli, W.: Recent glacier changes in the Alps observed by satellite: consequences for future monitoring strategies, Global Planet. Change, 56, 111–122, 2007. 1043
Rigon, R., Bertoldi, G., and Over, T. M.: GEOtop: A distributed hydrological model with coupled water and energy budgets, J. Hydrometeorol., 7, 371–388, 2006. 1053

GMDD
5, 1041–1076, 2012

Riseborough, D., Shiklomanov, N., Etzelmüller, B., Gruber, S., and Marchenko, S.: Recent Advances in Permafrost Modelling, Permafrost and Periglacial Processes, 156, 137–156, doi:10.1002/ppp.615, 2008. 1043

Seth, A., Giorgi, F., and Dickinson, R. E.: Simulating fluxes from heterogeneous land surfaces: explicit subgrid method employing the biosphere-atmosphere transfer scheme (BATS), J. Geophys. Res., 99, 18651–18667, doi:10.1029/94JD01330, 1994. 1044

Sokal, R. R. and Sneath, P. H. A.: Principles of taxonomy, Science, 156, 1356, doi:10.1126/science.156.3780.1356, 1967. 1047

Walland, D. J. and Simmonds, I.: Sub-grid scale topography and the simulation of Northern Hemisphere snow cover, Int. J. Climatol., 16, 961–982, 1996. 1044

Wood, E. F., Sivapalan, M., Beven, K., and Band, L.: Effects of spatial variability and scale with implications to hydrologic modeling, J. Hydrol., 102, 29–47, doi:10.1016/0022-1694(88)90090-X, 1988. 1044

Zadeh, L. A.: Fuzzy sets*, Information and control, 8, 338–353, 1965. 1047

Zemp, M., Haeberli, W., Hoelzle, M., and Paul, F.: Alpine glaciers to disappear within decades, Geophys. Res. Lett., 33, L13504, doi:10.1029/2006GL026319, 2006. 1043

GMDD

5, 1041–1076, 2012

Table 1. Speeds t (min) of K-means in terms of parameters, cluster number k, pixel number p, and nstart n. The K-means algorithm run time is less sensitive to cluster number (test 1,2) than to pixel number (test 1,3). Speed up is obtained by sampling (test 4 + 5 = 2.16 mins) and represents approximately a ten fold speed up based on 10:1 sampling ratio.

Test	p	k	n	t
1	1 m	64	10	22.23
2	1 m	32	10	18.99
3	250 k	64	10	4.71
4	100 k	64	10	1.13
5	1 m	64	1	1.03

GMDD
5, 1041–1076, 2012

Table 2. Parameters used in TopoSUB.

name	description	default value
K	number of samples	128
PREDs	input predictors	(ele, slp, asp, svf)
TV	target variables	(GST, SWE, SWin, Tair)
M	fuzzy membership exponent	1.4
iter.max	maximum iterations of K-means	20
nstart1	number of random starts for sample K-means	10
nstart2	number of random starts for all data K-means	1
μ_{max}	number of membership dimensions	20
W_{TV}	weighting of TV in informed scale	(1,1,1,1)

GMDD
5, 1041–1076, 2012

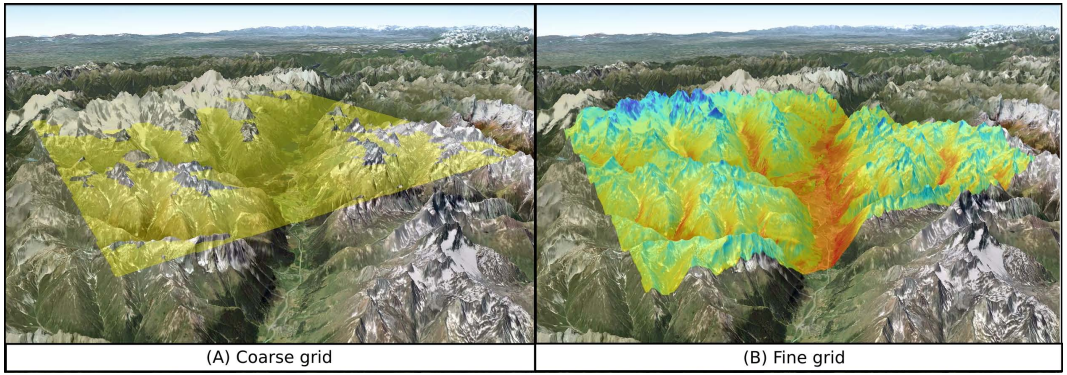

Fig. 1. The scale problem TopoSUB addresses – **(A)** the complex geometry of mountain topography is aggregated to a mean value in coarse grids which does not necessarily account for sub-grid heterogeneity **(B)** application of numerical models requires a fine grid in order to account for the effect of this heterogeneity, which is computationally expensive. TopoSUB allows for application of numerical models over large areas through a lumped approach that samples the most important aspects of this heterogeneity.

GMDD
5, 1041–1076, 2012

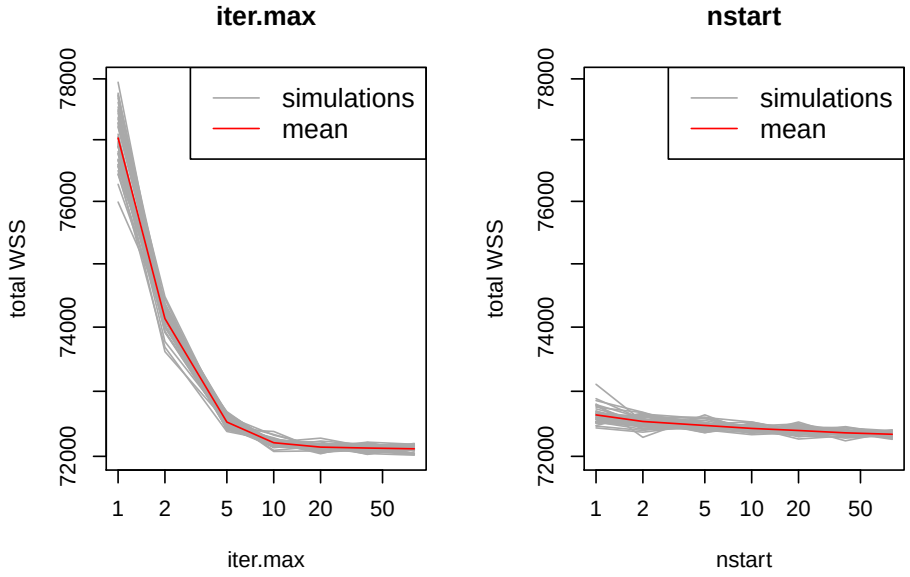

Fig. 2. The effect of maximum allowed iterations (iter.max) and number of random starts (nstart) of Kmeans on total within sum of squares (WSS) of resulting clusters (k = 128).

GMDD
5, 1041–1076, 2012

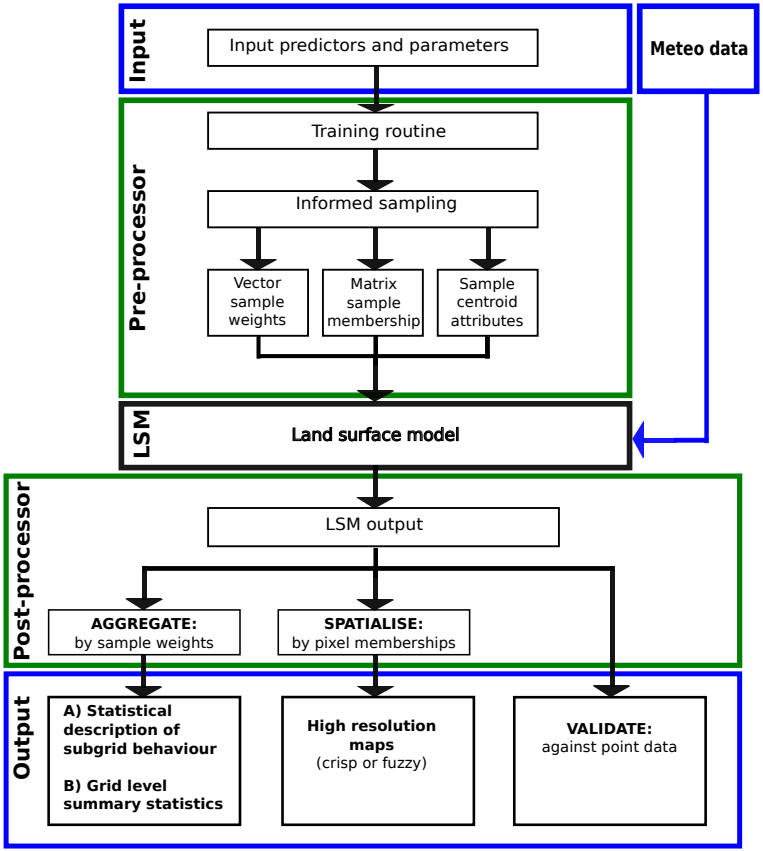

Fig. 3. Structure of the TopoSUB scheme with its two main modules: (1) pre-processor configures the sub-grid (runs once), (2) post-processor (runs multiple times together with the LSM).

GMDD
5, 1041–1076, 2012

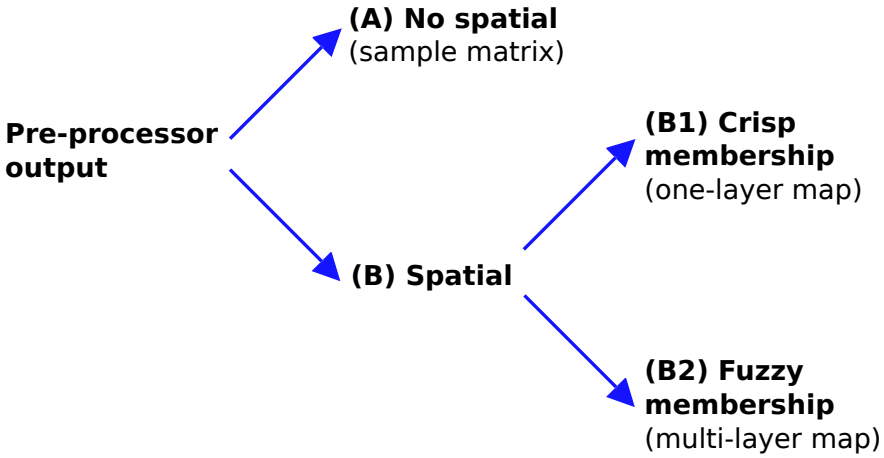

Fig. 4. Scheme of pre-processor output options, (A) no spatial, sample matrix of sample weights and environmental characteristics; (B1) spatial with crisp membership, single layer maps; and, (B2) spatial with fuzzy membership, multilayer maps with n number of chosen fuzzy membership dimensions.

GMDD
5, 1041–1076, 2012

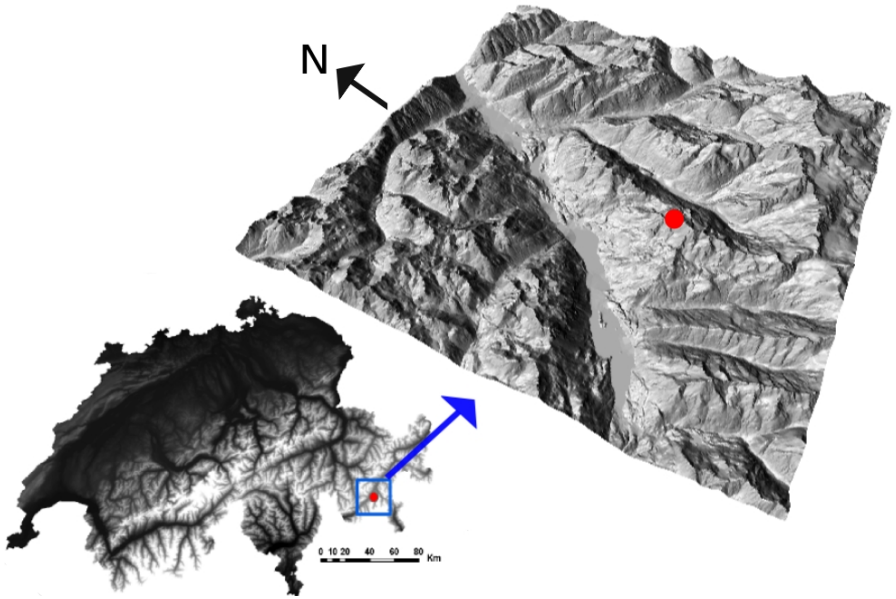

Fig. 5. Study region (25.6 × 25.6 km) in the upper Engadin, Switzerland. Location of driving climate station in red.

GMDD
5, 1041–1076, 2012

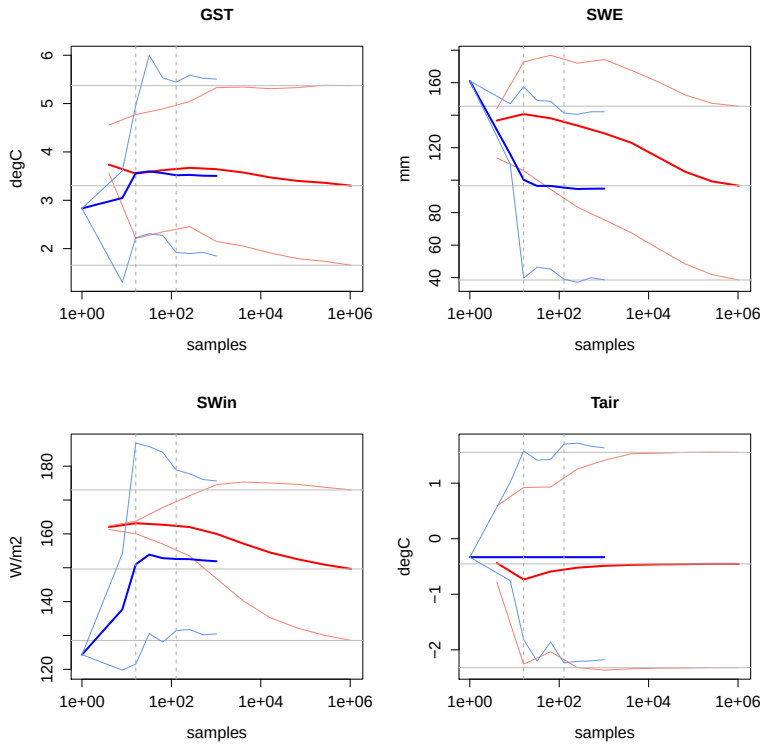

Fig. 6. Aggregated statistics: mean (in bold), 25th and 75th percentile (non-bold) of the sub-grid scheme at resolutions 1–1024 samples (blue) and distributed simulations (red) spanning resolutions of 4–10^6 pixels. Vertical lines indicate 16 and 128 samples. A stable performance is reached after the 128 sample level in all cases. TopoSUB is able to approximate aggregated 2-D simulation at 10^4 less computations. Note logarithmic × scale.

GMDD
5, 1041–1076, 2012

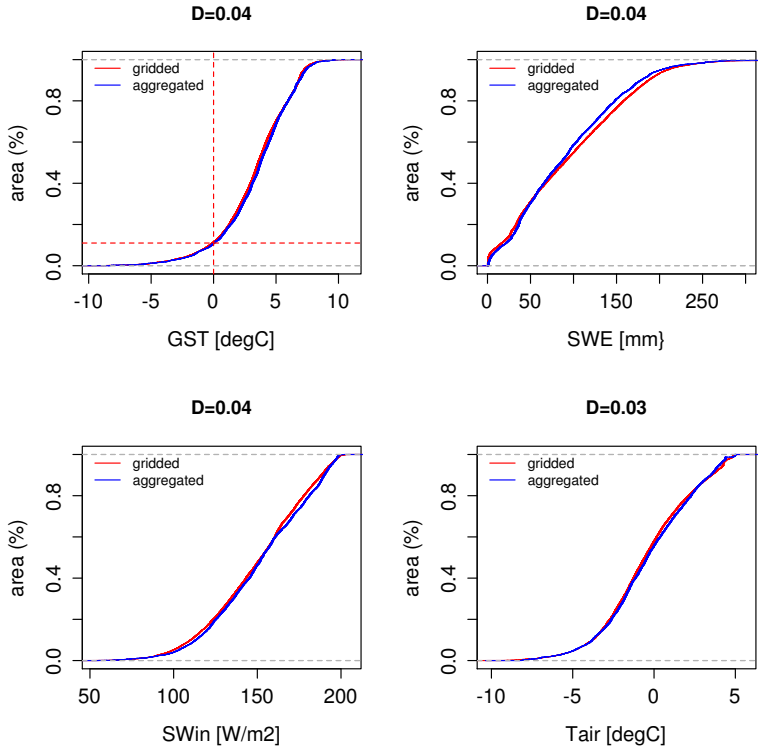

Fig. 7. CDFs of mean annual simulation results derived from the sub-grid scheme (blue) based on 258 samples and a distributed simulation (red) based on 10^6 pixels. A good fit is reported by a KS-test (D). Aggregated summary statistics can be computed directly from the CDF, for example percentage area with MAGST <0 °C (e.g. permafrost) (red dotted line).

GMDD
5, 1041–1076, 2012

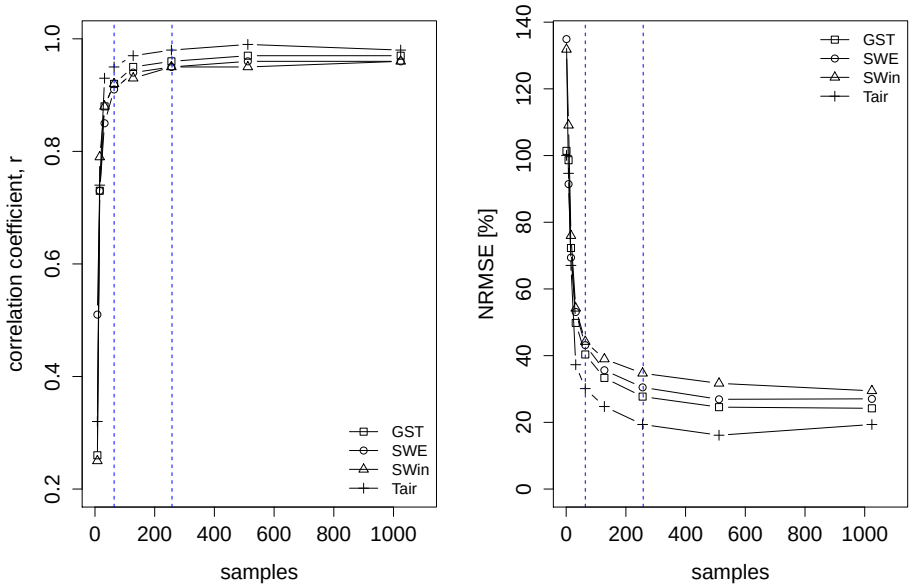

Fig. 8. The predictive power of the lumped scheme is assessed using the **(a)** r-value **(b)** NRMSE at resolutions 2–400 samples with respect to baseline distributed grid simulation. The majority of performance is gained until 64 samples (first dotted line). After 258 samples a reasonably stable performance is achieved (second dotted line).

GMDD
5, 1041–1076, 2012

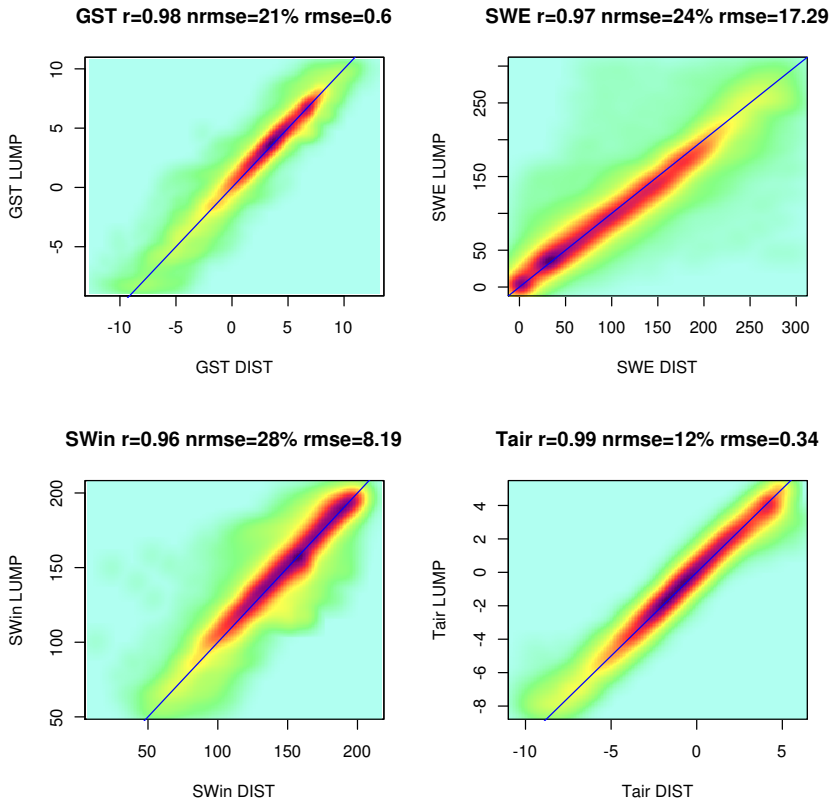

Fig. 9. Density scatter plot of 1-D/2-D after informed scaling and fuzzy spatialisation at 258 samples. All TVs are reproduced with low error as reported by r-value and RMSE (computed over 10^6 pixels).

GMDD
5, 1041–1076, 2012

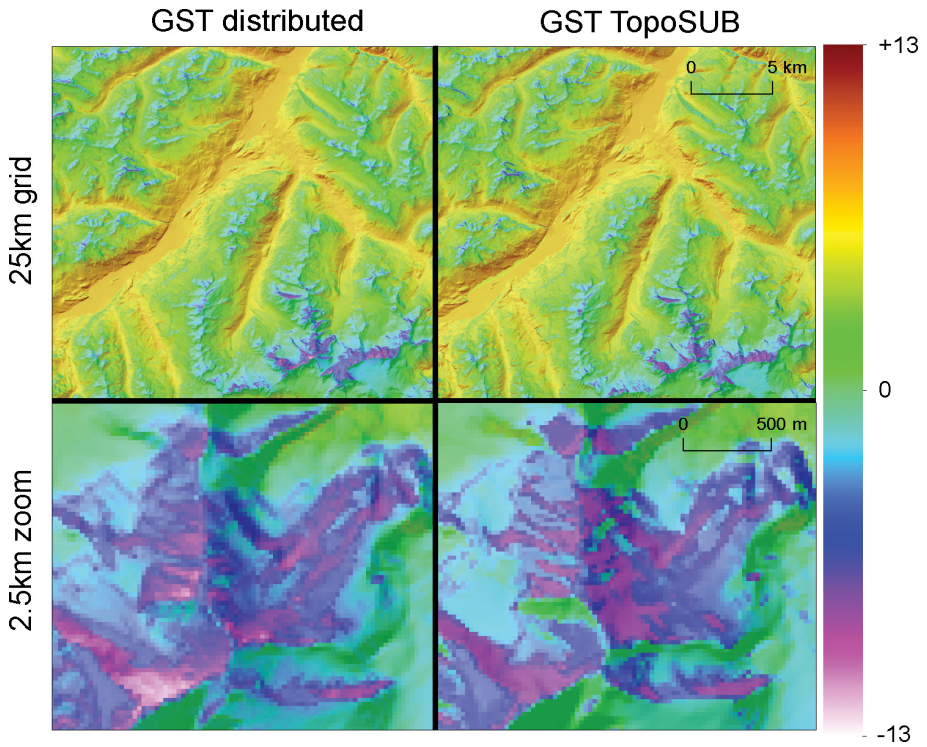

Fig. 10. GST computed at resolution of 25 m by distributed (2-D) and TopoSUB (1-D) over 25 × 25 km (top) and detailed view of 2.5 × 2.5 km (bottom). TopoSUB reproduces the spatial patterns of distributed simulation well, even at pixel resolution.

GMDD
5, 1041–1076, 2012

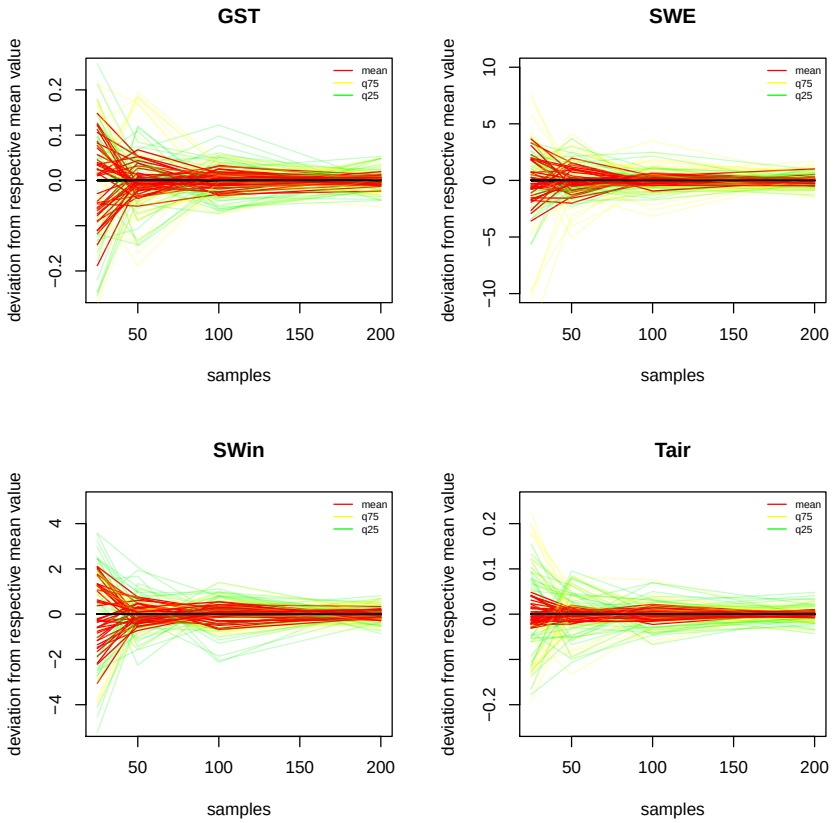

Fig. 11. Deviation from mean values of mean and quantiles 25/75 for 40 runs of TopoSUB. Results at resolutions of 25–200 samples indicate reasonable stability even at low resolutions. A significant increase in stability is seen between 25–100 samples in all variables tested.

GMDD
5, 1041–1076, 2012

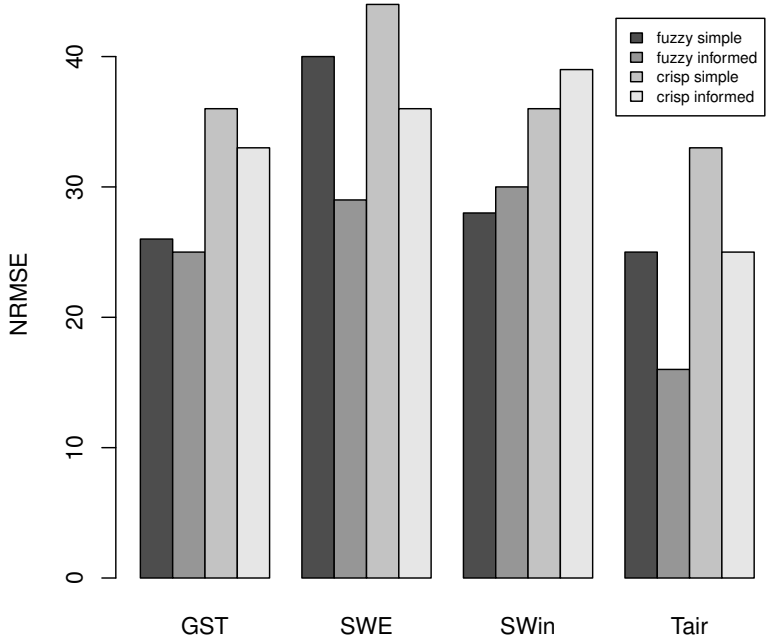

Fig. 12. NRMSE for 128 sample TopoSUB simulations for configurations: “fuzzy informed”, “fuzzy simple”, “crisp informed”, “crisp simple”. All TV results improve with fuzzy membership. All TV results except SWin improve with informed scaling.

GMDD
5, 1041–1076, 2012