Supplement Information

Supplement Information S1. Model sensitivity to precipitation lapse rate

In HYOGA2, precipitation is assumed to be constant for all elevation bands in each grid cell. Nevertheless, in order to examine the effect of precipitation lapse rate, a simple sensitivity test assuming a fixed precipitation lapse rate (a 10% increase per 100 m, similar to the value used in Radić and Hock (2011)) was applied to glaciers where observed mass balance can be obtained from DM05. Supplement Figure S1(d) shows the result of this sensitivity test: correlation coefficients between the computed mass balance with the fixed precipitation lapse rate and the observed mass balance of the available stations in DM05. Supplement Figure S1(d) is similar to Supplement Figure S1(c) which is the result of "without precipitation lapse rate"; for example, these two figures are similar in terms of the mean and standard deviation of correlation coefficients. As such, the result indicated that the modeled mass balance is not sensitive to this parameter in the retrospective simulation.

The same sensitivity test was then extended to a future warmer climate derived from a single GCM simulation, MIROC5. Thirty four stations among the 114 stations in Supplement Figure S1, where the model was well calibrated and glacier remained in 2100, were selected for the test. On average, the difference in cumulative mass balance since 1948 between the control simulation and the simulation with the fixed precipitation lapse rate was 4% in current climate (~2005) and 10% at the end of the future climate (~2100). Nonetheless, as was described in the main text, we basically did not use precipitation lapse rates in the simulations to avoid the incorporation of an additional uncertain parameter. In reality, precipitation lapse rates on a global scale are unknown. We would argue that the incorporation of adequate precipitation lapse

rate is a subject of future research and is beyond the scope of this letter. Further studies are expected to examine the effect of precipitation lapse rate in a future warming climate and to set it adequately on a global scale. A future study is also expected to investigate the physical reason behind the effect of precipitation lapse rate on mass balance estimation.

Supplement Information S2. Model Calibration and Initialization

The calibration parameters of HYOGA2 are the degree-day factors (DDFs) for ice and snow. The globally distributed DDFs for HYOGA2 were obtained following a method proposed by Hirabayashi *et al*. (2010). The individual glaciers obtained from the RGI inventory were first aggregated into one large glacier in 0.5° grid cells. Then, the DDFs for each cell were calibrated until yielding the maximum agreement with the cell-specific long-term (1948–1980) average of total glacier mass balance observed at 295 measurement sites. The mass balance observations, glacier areas and locations were available from Dyurgerov and Meier (2005) (DM05). If there were no mass balance observations within five grid cells from the target cell, and the glacier grid cell was located within one of the five regions with available data, the average regional mass balances of Dyurgerov and Meier (1997) and Serreze *et al*. (2000) were used for calibration. These regional balances were obtained as the area-weighted means of the observed mass balances. In total, 2545 out of the 3120 aggregated glaciers at 0.5° grid cells were calibrated against these cell-specific or regional estimates of long-term glacier mass balance. For the rest of the 575 grid cells in South America, Australasia, Africa, the sub-Antarctic islands, and the small Arctic islands, the long-term average mass balance in 1948–1980 was assumed to be zero. For estimating the DDFs, the upper and lower limits were set at 1 to 4 (mm °C^{-1} day^{-1}) for

snow and 4 and 20 (mm $°C^{-1}$ day^{-1}) for ice. These ranges were subjectively decided based on published papers with the objective of obtaining reasonable mass balance changes compared to the observed values within this model structure. In some glaciers, therefore, these DDF ranges may not reflect a similar range of locally observed DDFs. Because the calibrated DDFs, as the only calibration parameter, might have compensated for all potential errors in the model simulations, model improvements such as consideration of debris effects could improve the calibration results. Inclusion of such additional glacier processes along with collection of locally derived DDFs may improve the calibrated DDF values and initial ranges.

In the original HYOGA, the initial glacier area was considered to be a calibration parameter. It was calibrated until the calculated glacier area was close to the observed glacier area at the time of observation. Because HYOGA2 estimated the mass change in each individual glacier, the calibration of the initial glacier area led to unstable and erroneous simulations in relatively small glaciers. Assuming that the change in glacier area from the beginning of the model simulation (1948) to the observation year was negligible for the future area change of all glaciers, the glacier area recorded in the RGI was taken as an initial area on January 1, 1948. For glaciers located at relatively low altitudes, where a warming climate has been recorded in the past few decades, the retreating trends of glaciers may already have been obvious in the 20th century. For these glaciers, it was assumed that the change in area and associated change in volume in the past were smaller than the potential change in the future climate. Hence, the area feedback to altitude was not calculated for the past. For initialization of the snow amount, the state of snow thickness in each elevation band obtained at the end of the 33-year (1948–1980) simulation was used as an initial condition for the snow amount in the simulation used for analysis.

Supplement Information S3. Glacier Inventory

The glacier inventory used in this study was the version 2.0 of the Randolph Glacier Inventory (RGI, http://www.glims.org/RGI/randolph.html), which was processed by Marzeion *et al*. (2012). The location (longitude and latitude), maximum and minimum altitude, and glacier area could be obtained from the inventory. Only glaciers with complete data on location and area (257,325, with a total area of 515,706 km^2) were selected from the RGI (Figure 1). The number of selected MGI larger than 2 km^2 where climate input was available was 21,304, with a total area of 463,410 km^2. These glaciers covered ~90% of the total glacier area in the inventory. In addition, 2,930 LT2s (MGI < 2 km^2 in area) with a total area of 898 km^2, were selected from the RGI to represent all LT2s at each 0.5° grid and were considered for computation.

For glaciers with the median altitude available in the inventory but without maximum or minimum altitudes, the maximum or minimum altitude was calculated by assuming that the median altitude was one-third higher than the minimum altitude, as suggested by Hirabayashi *et al*. (2010). For glaciers without a median altitude, but with either a maximum or minimum altitude, the other value was calculated by assuming that the difference between the maximum and minimum altitudes was 1500 m (determined subjectively from glaciers with available data). In addition, for glaciers without data for both maximum and minimum altitudes, the maximum sample value of the 30 arc second (approximately 900 m) elevation data of the Shuttle Radar Topography Mission (SRTM) within each 0.5° grid was assumed to be the maximum altitude of the glacier. The minimum altitude was taken to be 1500 m lower than the maximum altitude. Finally, the maximum and minimum altitudes for each glacier

were adjusted such that they did not exceed the range of maximum and minimum altitudes of SRTM data within the corresponding 0.5° grid.

For the glaciers in the inventories without data for maximum glacier length L_{max}, the L_{max} (m) was calculated from the maximum Z_{max} (m) and minimum Z_{min} (m) elevation as

$$L_{max} = K (Z_{max} - Z_{min}). \qquad \text{(S1)}$$

The constant ratio K (0.004188) was calculated from glaciers where all three data were available in the extended World Glacier Inventory (WGI-XF; Cogley, 2010). To avoid potential model instability, the minimum value of L_{max} was set to 25 m, i.e., the minimum length that could be resolved in the 50 m vertical model grid.

As the RGI does not provide the date of observation for individual glaciers, the initial glacier area in the HYOGA2 historical simulation was assumed to be the same as the RGI glacier area. Because most of the surveys or images used to define the outline of RGI were dated after 2000, it was assumed that the impact of the initial glacier area on future mass balance simulations was minimal.

The mass balance estimation of MGI in HYOGA2 was sensitive to the input climate forcing, especially to surface temperature. This study used data from a global database, which may include some grids with unrealistic surface temperatures, resulting in erroneous mass balance estimation. This is apparent in High Mountain Asia, where the gauge observations for the correction of global climate datasets is limited. To identify such grids and exclude them from the analysis, upper and lower limits of annual mass balance were set at 2 and -5 m in the retrospective simulations, respectively. This range was subjectively decided from the available observed mass changes. We realized that real mass changes in glaciers may be inside or outside this range. However, for the purpose of global modeling, we applied this threshold to

exclude potential erroneous calculated mass changes caused by the model. Glaciers showing erroneous mass balance calculations (before 1960) outside the defined limits were excluded in the analysis. In total, 2,104 glaciers from the RGI were excluded from the regional mass balance analysis. The total area of the excluded glaciers was 66,472 km^2. We assumed that the exclusion of these glaciers would not significantly alter the regionally averaged total mass estimates in the model. The ratio of the initial glacier areas calculated in the model (in total 450,505 km^2) and the areas of all the small glaciers and ice caps registered in the RGI (in total 515,706 km^2) at each 0.5° grid cell was used to scale up the results to estimate the global volume loss from glaciers.

Supplement Information S4 . Extension of Observation-based Climate Data

The retrospective simulations for the 64-year period from 1948 to 2011 inclusive of HYOGA2 were forced by an extended version of the observation-based global 0.5° gridded dataset of daily precipitation and near-surface temperature (H08; Hirabayashi *et al*., 2005, 2008a). The daily temperature of H08 was based on monthly temperature and monthly diurnal temperature ranges of the Climate Research Unit version TS 2.1 data for 1948–2002 (CRU; Mitchell and Jones, 2005). For extended periods from 2003 to 2011, the monthly temperature of Fan and van den Dool (2008) was scaled using the mean ratio of monthly temperatures from the CRU and Fan and van den Dool (2008), from 1986 to 2002. The monthly diurnal temperature range of the CRU TS3.1 (2008–2009) and monthly means of maximum and minimum temperatures of the Global Telecommunications System (GTS, 2010–2011) were used for extension of the monthly diurnal temperature range. The daily precipitation data of H08 was based on version 6 of the monthly precipitation product from the Global Precipitation

Climatology Centre (Rudolf and Rubel, 2005 for 1948–2010), and the daily precipitation product from the GTS (for obtaining the monthly mean for 2011 and wet-day number after 2003) and APHRODITE V1003R1 (Yatagai *et al.*, 2009). Hirabayashi *et al.* (2005) described the disaggregation of monthly climatic variables into a daily time series using a stochastic weather generator, and Hirabayashi *et al.* (2008b) explained the under-catch correction for precipitation data. Daily observed values of APHRODITE with correction of the gauge under-catch bias were used where and when they were available.

Supplement Information S5. Bias correction of the GCM Data

Because of unavoidable overall biases between GCM simulations and observed data (e.g., temperature, precipitation), GCM data should be corrected before being used as inputs in assessment models such as glacier models. It is necessary to bring the statistical properties of present-day simulations in line with observations and to use this information to correct future projections (Seneviratne *et al.*, 2012). The original GCM data were first disaggregated onto a 0.5° grid by linear interpolation from four neighboring grid cells. Then, the biases in the daily time series of future land-surface temperature and precipitation were corrected.

The bias correction method explained by Watanabe *et al.* (2012) was used to correct the GCM-simulated temperature and precipitation. The GCM simulation data (30 years from 1950 to 1980 at 0.5° spatial resolution) were compared against observation-based daily climate data (H08; Hirabayashi *et al.*, 2008a) to estimate the biases. The corrections were then conducted in two steps following the method of Haerter *et al.* (2011), in which the monthly and daily variances were corrected simultaneously. First, the monthly bias was corrected by comparison of the monthly

GCM and observation data. The daily variability was then corrected by calculating the difference between the daily time series and monthly means of the GCM data and observations.

During the bias correction, statistical parameters related to monthly temperature and precipitation, instead of monthly values, were adjusted. A brief introduction of the bias correction methods for temperature and precipitation are presented here.

The monthly bias correction method for temperature can be expressed as

$$\bar{x}_{mon,cor} = \bar{x}_{mon,obs} + \bar{x}_{mon,21c} - \bar{x}_{mon,20c} \qquad (S2)$$

$$\sigma_{mon,cor} = \sigma_{mon,21c}\, \sigma_{mon,obs}\, (\sigma_{mon,20c})^{-1} \qquad (S3)$$

$$x_{mon,cor} = f(P(x_{mon,21c}; \bar{x}_{mon,21c},\, \sigma_{mon,21c}); \bar{x}_{mon,cor},\, \sigma_{mon,cor}) \qquad (S4)$$

where $\bar{x}_{mon,cor}$ and $\sigma_{mon,cor}$ are the corrected mean and standard deviation of monthly means of GCM temperature, respectively, $\bar{x}_{mon,21c}$ and $\bar{x}_{mon,20c}$ are the means of future and current GCM temperature, respectively, $\bar{x}_{mon,obs}$ is the mean of the observation-based gridded temperature product from H08, $\sigma_{mon,21c}$ and $\sigma_{mon,20c}$ are the standard deviation for the future and current GCM temperature, and $\sigma_{mon,obs}$ is the standard deviation of H08 temperature. $P(x;a,b)$ represents a cumulative distribution function (CDF) of the data x, which are normally distributed with mean a and standard deviation b, and $f(x;a,b)$ represents an inverse function of the CDF $P(x;a,b)$. The daily time series was then obtained from the bias-corrected monthly time series as

$$x_{day,cor} = \sigma_{day,cor}\, (\sigma_{day,21c})^{-1}\, (x_{day,21c} - \bar{x}_{day,21c}) + x_{mon,cor} \qquad (S5)$$

where $\sigma_{day,21c}$ is the standard deviation of the daily time series of future GCM temperature, $\sigma_{day,cor}$ is the standard deviation of the daily time series of corrected

temperature obtained by applying the equidistant CDF matching method (EDCDF method; Li *et al*., 2010) to the standard deviations of daily time series of current/future GCM and observed temperatures. The equidistant CDF matching method was originally applied to monthly climate data.

The bias correction for monthly precipitation was performed by fitting the precipitation statistics to a two-parameter gamma distribution, expressed as

$$k_{mon,cor} = \bar{k}_{mon,21c} \, \bar{k}_{mon,obs} \, (\bar{k}_{mon,20c})^{-1} \qquad \text{(S6)}$$

$$\theta_{mon,cor} = \bar{\theta}_{mon,21c} \, \bar{\theta}_{mon,obs} \, (\bar{\theta}_{mon,20c})^{-1} \qquad \text{(S7)}$$

where k and θ are the scaling and shape parameters of the gamma distribution, respectively. The subscript notations are the same as those used for temperature (Equations (S2) through (S4)). Once the bias in monthly GCM precipitation was corrected, the EDCDF method, as in the case of temperature, was used to compute the bias-corrected daily time series of the precipitation.

Supplement Information S6. Model validation

The old and new versions of the glacier model HYOGA were validated using observation-based mass balance of glaciers. Figure S1 compares the correlation coefficient of the computed mass balance and available observations. The mean and standard deviation of the correlation coefficients among the available stations of DM05 were 0.47±0.24 for HYOGA, 0.48±0.25 for HYOGA2 with calibrated DDFs, and 0.42±0.3 for HYOGA2 with globally distributed DDFs. It is clear that HYOGA2 with globally calibrated DDFs had a similar performance to the old model. Hence, the overall time series of annual mass balance global means were similar between the old and new models (Figure 2).

Supplement References (for other references see main text)

Cogley JG. 2010. A more complete version of the World Glacier Inventory. *Annals of Glaciology* **50**: 32–38. doi: 10.3189/172756410790595859.

Dyurgerov MB, Meier MF. 1997. Mass balance of mountain and subpolar glaciers: a new global assessment for 1961–1990. *Arctic and Alpine Research* **29**:379–391.

Fan Y, van den Dool H. 2008. A global monthly land surface air temperature analysis for 1948-present. *Journal of Geophysical Research-Atmospheres* **113**: D01103. doi: 10.1029/2007JD008470.

Haerter J, Hagemann S, Moseley C, Piani C. 2011. Climate model bias correction and the role of timescales. *Hydrology and Earth System Sciences 1***5:** 1065–1079. doi:10.5194/hess-15-1065-2011.

Hirabayashi Y, Kanae S, Motoya K, Masuda K, Döll P. 2008b. A 59-year (1948–2006) global near-surface meteorological data set for land surface models. Part II: Global snowfall estimation. *Hydrological Research Letters* **2**: 65–69. doi: 10.3178/HRL.2.65.

Li H, Sheffield J, Wood E. 2010. Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. *Journal of Geophysical Research-Atmospheres* **115**: D10101. doi:10.1029/2009JD012882.

Mitchell T, Jones P. 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology* **25**: 693–712. doi:10.1002/joc.1181.

Rudolf B, Rubel F. 2005. Observed Global Climate, Landolt-Börnstein (Numerical Data and Functional Relationships), M. Hantel (eds). Springer: Berlin; 11.11-11.53.

Seneviratne SI, Nicholls N, Easterling D, Goodess CM, Kanae S, Kossin J, Luo Y, Marengo J, McInnes K, Rahimi M, Reichstein M, Sorteberg A, Vera C, Zhang X. 2012. Changes in climate extremes and their impacts on the natural physical environment. In: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC), Field CB, Barros V, Stocker TF, Qin D, Dokken DJ, Ebi KL, Mastrandrea MD, Mach KJ, Plattner G-K, Allen SK, Tignor M, Midgley PM. (eds). Cambridge University Press, Cambridge, UK, and New York, NY, USA; 109–230.

Serreze MC, Walsh JE, Chapin III FS, Osterkamp T, Dyurgerov M, Romanovsky V, Oechel WC, Morison J, Zhang T, Barry RG. 2000. Observational evidence of recent change in the northern high-latitude environment. *Climatic Change* **46**:159–207.

Yatagai A, Arakawa O, Kamiguchi K, Kawamoto H, Nodzu MI, Hamada A. 2009. A 44-year daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *SOLA* **5**: 137–140. doi:10.2151/sola.2009-035.

Supplement Table

Supplement Table SI. Summary of the GCMs selected in this study. The institution and model names were taken from http://cmip-pcmdi.llnl.gov/cmip5/availability.html. Size information was extracted from data headers.

Model	Institution	Size (no. of grids)	
		North-South	East-West
CCCma-CanESM2	Canadian Centre for Climate Modelling and Analysis, Canada	64	128
CNRM-CM5	Centre National de Recherches Meteorologiques/ Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique, France	128	256
GFDL-ESM2G	Geophysical Fluid Dynamics Laboratory, USA	90	144
INM-CM4	Institute for Numerical Mathematics, Russia	120	180
MIROC5	Atmosphere and Ocean Research Institute, National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology, Japan	128	256
MPI-ESM-LR	Max Planck Institute for Meteorology (MPI-M), Germany	96	192
MRI-CGCM3	Meteorological Research Institute, Japan	160	320
NCAR-CCSM4	National Center for Atmospheric Research, USA	192	288
NCC-NorESM1-M	Norwegian Climate Centre, Norway	96	144

Supplement Figures
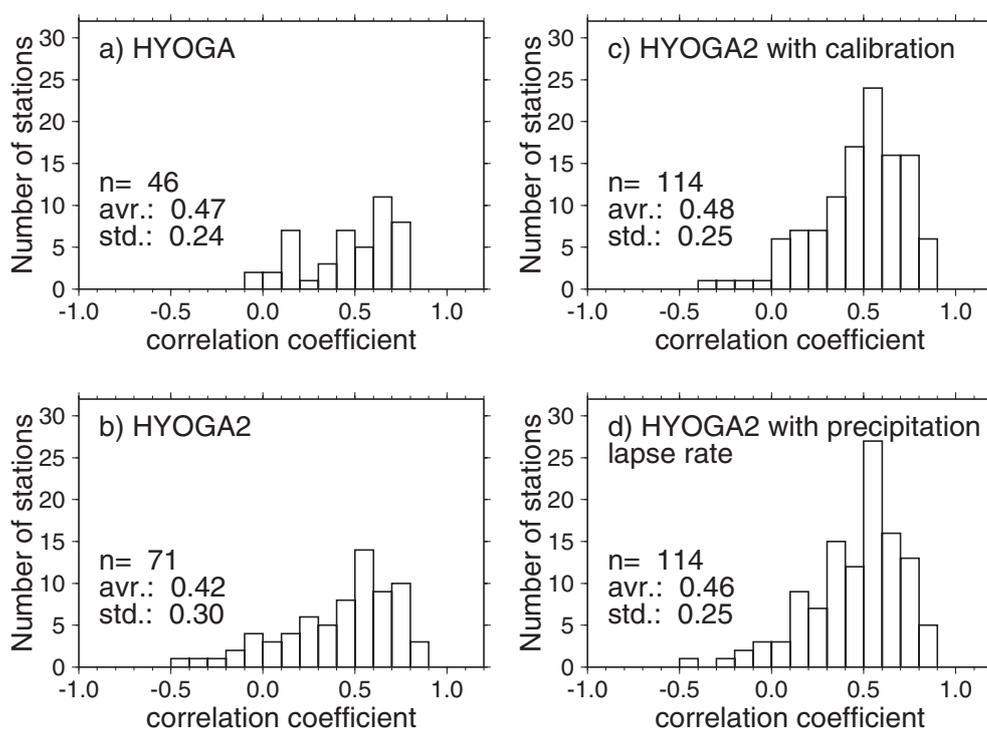
Supplement Figure S1. Frequency of correlation coefficients between observed and modeled annual mass balances of the old model HYOGA (a), HYOGA2 with globally distributed parameters (this study; b), HYOGA2 with calibration (c), and the sensitivity test of HYOGA2 with a 10% (100m)$^{-1}$ increase in precipitation with altitude (d). Number of glaciers (n), averages (avr.,) and standard deviations (std.) of correlation coefficients among all stations are given in the figures.