Selecting suitable climate models for examining future changes in soil erosion and muddy flooding

*Neil Brannigan [1] | Donal Mullan [1] | Karel Vandaele [2] | Conor Graham [1] | Jennifer McKinley [1] | John Meneely [1]

[1] Geography, School of Natural and Built Environment, Queen's University Belfast, Belfast BT7 1NN, Northern Ireland, UK.

[2] Watering van Sint–Truiden, 3800 Sint– Truiden, Limburg, Belgium.

*Corresponding author: Neil Brannigan - nbrannigan06@qub.ac.uk

Short Informative

Soil erosion; muddy flooding; climate modelling; soil erosion modelling; climate change; mitigation.

Short Running Title

Selecting climate models for examining future changes in soil erosion.

Acknowledgements

The authors wish to thank the Department for the Economy (DfE), who provided the lead author with a PhD scholarship to complete this research.

Abstract

Climate models consistently project large increases in the frequency and magnitude of extreme precipitation events in the 21st century, revealing the potential for widespread impacts on various aspects of society. While the impacts on flooding receive particular attention, there is also considerable damage and associated cost for other precipitation – driven phenomena, including soil erosion and muddy flooding. Multiple studies have shown that climate change will worsen the impacts of soil erosion and muddy flooding in various regions. These studies typically drive erosion models with output from a single climate model or a few models with little justification. A blind approach to climate model selection increases the risk of simulating a narrower range of possible scenarios, limiting vital information for mitigation planning and adaptation. This study provides a comprehensive methodology to efficiently select suitable climate models for simulating soil erosion and muddy flooding. For a study region in Belgium using the WEPP soil erosion model, we compare the performance of our novel methodology against other model selection methods for a future period (2081 – 2100). The main findings reveal that our methodology is successful in generating the widest range of future scenarios from a small number of models, compared with other selection methods. This represents a novel targeted approach to climate model selection with respect to soil erosion by water but could be modified for other precipitation – driven impact sectors. This will ensure a broad range of climate impacts are simulated so the best- and worst-case scenarios can be adequately prepared for.

Key words: soil erosion; muddy flooding; climate modelling; soil erosion modelling; climate change.

1. Introduction

Climate models provide projections of future climate required for various climate change impact studies. These studies inform policymakers on necessary adaptation measures to mitigate climate change impacts. Most global and regional climate models (GCMs and RCMs) consistently project large increases in the frequency and magnitude of extreme events, while average daily rainfall intensities are also projected to rise throughout the 21st century (IPCC, 2013; Zhang, 2013). This is because temperatures are expected to increase by between 1.8°C and 4°C by the end of the 21st century (IPCC, 2013), leading to an intensified global hydrological cycle (Zhang, 2012). Extreme rainfall is highly correlated to changes in temperature, largely because of the Clausius – Clapeyron (CC) relation where the saturated vapour pressure of the atmosphere is described to increase at an approximate rate of 7% for every 1°C warming or 7% K (Mullan *et al.*, 2019). Furthermore, this rate is even higher for rainfall intensity (e.g. Sun *et al.*, 2007), with the most extreme precipitation events promoting an increase to 14% (Lenderink & Van Meijgaard, 2008).

These climatic changes have caused concern that processes driven by large-scale precipitation events, such as global soil erosion, will be exacerbated in future (e.g. Risbey and Entekhabi, 1996; Nearing *et al.*, 2005; Scholz *et al.*, 2008; Kundzewicz *et al.*, 2009; Zhang *et al.*, 2009). Soil erosion is already identified as one of the major environmental threats to arable land globally (Heitz *et al*., 2009; Maeda *et al*., 2010; Nearing *et al*., 2005; Panagos *et al*., 2015). Global soil erosion rates have previously been estimated to be around 10.2 ha^{-1} yr^{-1} (Yang *et al*., 2003), with erosion by water accounting for the most significant loss of soil (Panagos *et al*., 2015; Verstraeten *et al*., 2003; Yang *et al*., 2003), contributing to approximately 55% of global soil erosion totals (Bridges & Oldeman, 1999). Sediment loads and water discharge were previously found to change by 2% and 1.3%, respectively, for every 1% change in precipitation (Lu *et al*., 2013).

Climate change is therefore expected to continue to pose a serious threat to processes driven by large-scale precipitation events, with an increase in associated future financial costs. Consequently, there is a vital need to produce climate scenarios to assess how soil erosion and MF will be impacted by a changing climate. Previous modelling of future soil erosion reveals projected increases in soil loss in many parts of the world by 2100. For instance, Li and Fang (2016) gathered 205 available results from other soil erosion modelling studies and 136 of these revealed erosion rates to increase across the world under future climate scenarios. Of these results, 49 show that soil erosion rates will increase by more than 50%. As a symptom of soil erosion by water within a certain geographical range, muddy flooding (MF) has also been projected to increase in magnitude (Mullan *et al.*, 2016). MF is a term that describes runoff flowing from poorly vegetated arable land carrying large amounts of soil as suspended sediment or bedload (Boardman & Vandaele, 2016) that induces damage to public infrastructure and freshwater systems further downstream. Mullan *et al.* (2019) projected an earlier and longer MF season for a hillslope in eastern Belgium, along with an increase in the number of MF events each year. Total damages to private householders were previously estimated to range between €55 million to €165 million each year in Flanders, Belgium (Verstraeten and Poesen, 1999), with similar costs to public infrastructure. Total costs induced by water erosion globally are astronomically larger. Pimentel (2006) estimated total off-site water erosion costs of \$2.3 billion yr^{-1} for USA alone.

Despite this importance, there is currently limited methodological emphasis on how climate models are selected for modelling soil erosion. It is imperative that a thorough selection process is followed to select a manageable number of representative climate models for the study application. The Coupled Model Intercomparison Project Phase 5 (CMIP5) archive (Taylor *et al.*, 2012) contains outputs from 61 different general circulation models, such that all projections cannot be included for thoroughly studying the impacts of climate change. Constraints in computational and human resources mean that model choice must be limited to a practicable number, while an increase in the number of available models corresponds to an increase in the uncertainty remaining over future climate simulations. The uncertainty provided by the spread in climate model projections is a considerable concern in climate change impact studies, commonly larger than the uncertainty associated with model parameterisation and natural variability (Finger *et al.*, 2012; Lutz *et al.*, 2016; Minville *et al.*, 2008).

The absence of strategic climate model selection for soil erosion and MF applications has prompted a thorough climate model selection process to be developed and followed in this research. This methodology is partly inspired by Lutz *et al.* (2016) to combine two commonly applied selection concepts, but modified to become more targeted for specific application to soil erosion and MF. This is the primary aim of this research – to provide a comprehensive methodology to efficiently select suitable climate models for simulating processes driven by future large-scale precipitation events, with specific application to MF and soil erosion. It is intended that a wide range of sensible projections can be generated from a small number of models, such that the best- and worst-case future scenarios can be adequately prepared for without analysing dozens of model outputs.

2. Materials and Methods

2.1. Study Area

The Belgian loess belt (Figure 1) is an 8867 km^2 plateau that gently slopes north with a mean altitude of 115m. Belgium has a temperate maritime climate with mild winters and cool summers, influenced by the North Sea and Atlantic Ocean. As determined from the E-OBS high-resolution (0.25°) gridded dataset over Europe (Haylock *et al.*, 2008) between 1986 and 2019, mean annual temperatures range from 3.5°C in January to 18°C in July and August for the grid containing the study area (http://climexp.knmi.nl). Rainfall has an even distribution throughout the year, with average annual rainfall amounts ranging from 520 mm to 960 mm in the study area (~ 55 mm to 75 mm per month). Belgium also possesses the highest density of cultivated land in the country (Beckers *et al.*, 2018). Summer crops – such as maize, potatoes, and sugar beet – have increased in recent decades and now dominate the arable land in place of winter cereals. Cover crops such as mustard and phacelia are often encouraged to shield the soil during late spring and early summer while summer crops reach maturity (Bielders *et al.*, 2003; Mullan *et al.*, 2016).

This research focuses on a dry valley locally known as the 'Heulen Gracht' (50.76° N, 5.12° E) located within the 200 km^2 Melsterbeek catchment in the Limburg province of Belgium. The Heulen Gracht is a prominent landscape for academic and community research on MF problems and solutions (e.g., Boardman & Vandaele, 2020; Evrard *et al.*, 2007b, 2008), covering a 3 km^2 (300 hectares (ha)) area. This study focuses on the downstream half of the Heulen Gracht (approximately 1.3 km^2) to allow for precise drone measurements, such that an altitude of between 107 and 140 metres (m) was determined. Soil sampling in this study revealed a topsoil consisting of an average 17.5% sand, 75.1% silt, and 7.4% clay. Drone imagery also revealed that cropland covers 61% of the catchment surface, while grassland and orchards cover 27% and roads (1%) account for the remainder. The study area drains into Velm Village, which has a local reputation as a 'devastated village' after repeated flooding in recent decades. Two separate hillslopes have been selected for analysis in this study and the characteristics of each hillslope are described in Table 1. These hillslopes have been selected because they vary in steepness, length, crop management and mitigation measures.

2.2. Baseline Soil Erosion Modelling

The Water Erosion Prediction Project (WEPP) model (Flanagan and Nearing, 1995; v.2008.907) is a spatially-distributed continuous simulation model, providing long-term simulations of soil erosion and deposition along with other key soil, hydrology, and plant components at hillslope, field and small catchment (< 260 ha) scales (Laflen *et al.*, 1991; Ascough *et al.*, 1995; Li *et al.*, 2017). WEPP can predict soil erosion, sediment transport, and deposition across the landscape by applying a steady-state continuity equation to predict rill and inter-rill erosion processes. WEPP is widely applied within climate change – soil erosion research, with success demonstrated in a range of studies (e.g. Nearing, 1998; Stolpe, 2005), not least to investigate MF for a hillslope in the Melsterbeek catchment in Mullan *et al.* (2016, 2019).

2.2.1. Soil

A 30 cm bulk soil sample was taken every 10 m at each hillslope using a soil auger, reaching a maximum depth of 150 cm (Appendix 1). Lab analysis of the collected soil samples revealed a topsoil consisting of an average 17.5% sand, 75.1% silt, and 7.4% clay, which is consistent with previous topsoil sampling for the Heulen Gracht and other analogous catchments in the Belgian loess belt recorded in Evrard (2008). An average organic matter (OM) content of 4% is also consistent with results in Mullan *et al.* (2019) for an analogous catchment < 15 km away. Soil characteristics for the hillslopes are provided in Table 2. Critical shear, hydraulic conductivity, and rill and interrill erodibility values were estimated by WEPP. Estimated albedo was set at 0.1, CEC (meq/1) at 15, and initial soil saturation at 75% for each hillslope.

2.2.2. Slope

Slope profiles were established for both hillslopes following a high-resolution drone survey with an average ground sampling distance (GSD) of 2.3 cm. However, WEPP is incapable of processing slope data of this resolution for greater than 2 m length. WEPP allows up to 100 data points for both cumulative distance (ft) and slope (%), respectively, yet input value totals ≥ 50 tends to generate a distorted slope profile that produces misleading results. This limitation had not been previously reported in literature, perhaps owing to lower resolution data being used. To manage this, the maximum number of input values prior to distortion had been applied to simulate WEPP for each hillslope. Given that both hillslopes differ in length, the number of input values used to simulate

WEPP differs between each hillslope, such that the sampling distance is 2.3 m and 6.6 m for Hillslopes 1 and 2, respectively.

2.2.3. Land Management

Land use data between 2008 and 2018 had been collected from Geopunt Vlaanderen (https://www.geopunt.be/kaart), which is an opensource database provided by the Flemish Government. Crop rotation dates had been sourced by the local soil erosion expert, Dr Karel Vandaele. Plant growth parameters were calculated by WEPP for each crop without additional modifications (Flanagan and Nearing, 1995). The high-resolution imagery (GSD 2.3 cm) captured by the drone survey allowed for accurate measurements of the dimensions of certain features, such as a grass buffer strip or grassed waterway at the bottom of a given hillslope. Land management details required to simulate WEPP for both hillslopes are displayed in Appendix 2.

2.2.4. Climate

Baseline climate data were simulated using the stochastic weather generator CLIGEN (Nicks *et al.*, 1995), which draws on the statistical properties of observed climate measurements to generate long-term daily climate data. All required CLIGEN input parameters are presented in Table 3. Previous studies (e.g. Nearing, 1990) have demonstrated that the precipitation variables provide the greatest influence on soil loss and runoff projections using WEPP. Mean precipitation per wet day is calculated using monthly means, skewness, and standard deviation values. Series of wet and dry days are determined from the transitional probabilities of a wet day following a wet day (Pw/w) and a wet day following a dry day (Pw/d). Rainfall intensity is calculated from determinations of monthly half hour precipitation (MX.5P) and time to peak storm intensity (Time PK). Time PK is a dimensionless variable that represents an empirical probability distribution of the time to peak storm intensity as a fraction of storm duration, such that this is the only variable that is not calculated for each given month (Mullan *et al.*, 2019; Yu, 2003).

High resolution (0.25°) observed (E-OBS) daily temperature and precipitation data from 1950 - 2019 (Haylock et al., 2008) were downloaded from the Royal Netherlands Meteorological Institute (KNMI) Climate Explorer site (http://climexp.knmi.nl) for the grid containing the study area. The Niel-bij-Sint-Truiden climate station (pinned in Figure 1, less than 3 km from the Heulen Gracht) provided sub-hourly precipitation data between 2004 and 2014 to determine Time Pk and MX.5P. The remaining parameters – solar radiation, wind speed and direction, and relative humidity – were sourced from nearby Maastricht, Netherlands. Equation 1 was used to convert relative humidity to dew point temperature (Alduchov & Eskridge, 1996). CLIGEN was simulated for 330 years to represent 30 cycles of each 11-year crop rotation in WEPP, as recommended by Mullan *et al.* (2019).

$$TD = \frac{243.04\left(\ln\left(\frac{RH}{100}\right) + \frac{(17.625 * T)}{(243.04 + T)}\right)}{\left(17.625 - \ln\left(\frac{RH}{100}\right) - \frac{(17.625 * T)}{(243.04 + T)}\right)} \tag{1}$$

where TD = dew point temperature; $\ln$ = natural logarithm; RH = relative humidity; and T = mean temperature.

2.3. Climate Model Selection

Due to computational and human resource limits, the range of viable soil erosion models is much narrower than the range of potential climate models. It is routine in previous climate change – soil erosion studies to select a small subset of climate models for impact analysis, yet the reasons for selecting specific models are often arbitrary or justified based on some simple statistical information relating to the models. We compare three different approaches in this study – each of which will now be outlined. The method that yields the widest range in projections for key soil erosion and MF diagnostics (while using an equal number of scenarios) will be determined the most desirable method for model selection. Although a wide envelope of uncertainty makes adaptation and planning decisions difficult, it is important to capture the widest possible spread to account for a wide array of potential climate futures – without the need to apply dozens of climate models to soil erosion impact studies.

2.3.1. Past – Performance and Envelope (PPE) Method

While climate models are commonly chosen based upon their past-performance (e.g., Pierce *et al.*, 2009; Biemans *et al.*, 2013) – i.e., their ability to closely simulate present and near-past climate – it is plausible that potential climate scenarios may be omitted. Alternatively, the ‘envelope approach’ ensures that a broad range of projections for a given climatological variable is represented from a selected ensemble of models. However, by neglecting the skill provided by the model in simulating present and near-past climate, this approach assumes that all models are equally plausible. It is only mean annual changes that define model selection using the envelope approach (Lutz e*t al.*, 2016). With these limitations in mind, the revised methodology applied in this research (herein referred to as the PPE method) is inspired by the concept provided in Lutz e*t al.* (2016) to combine the past-performance and envelope approaches for selecting a manageable number of the most suitable climate models. PPE adapts and departs from the envelope approach in Lutz *et al.* (2016) to be specifically applicable to precipitation driven phenomena, while certain key precipitation characteristics necessary to run CLIGEN in WEPP are compared to assess model past – performance.

Precipitation data (mm/day) from each model were downloaded for both a moderate radiative forcing scenario – representative concentration pathway (RCP4.5) – and a high radiative forcing scenario – RCP (RCP8.5) for the future period (2081-2100), with E-OBS 1950 – now 0.25° Europe observed data (1986-2005) used as a historical baseline. RCP4.5 provided 102 model runs, while RCP8.5 provided 77 model runs. For both RCPs, the average ΔP between the future period and the observed period was calculated for all models. All available initial condition ensemble members were included for all models since each initial condition ensemble member leads to a different future.

To avoid selecting outliers, the 10th and 90th percentile values for ΔP for both RCPs were determined. These percentile values represented ‘wet’ (90th percentile) and ‘dry’ (10th percentile) sides. All models, irrespective of time step scale, were added to the initial selection of models to calculate the percentile values, thereby ensuring that all projected possible scenarios were fully represented. However, since all models must provide data at a daily time step for empirical – statistical downscaling at a later stage, the number of ‘available’ models for selection was subsequently significantly reduced. The three daily time step models with the lowest distance from each side were selected by subtracting the precipitation value (% mm/day) from each percentile value. The selected wet and dry models for RCP4.5 and RCP 8.5 are provided in Table 4. The reduced number of models available at a daily time step rendered overlap in selected models between ‘cold-dry’ and ‘warm-dry’ models, and ‘cold-wet’ and ‘warm-wet’ models alike. In other words, the selection process was unable to explicitly distinguish between ‘cold’ and ‘warm’ models, instead providing more simply ‘wet’ and ‘dry’ sides for model selection.

Certain precipitation characteristics necessary to run CLIGEN in WEPP were compared between observed and historical modelled data for the selected models (Table 4). Any negative values were converted to positive. The purpose of this step is to further narrow down model choice to models that most closely simulate observed metrics of precipitation that are important for MF. These metrics are the mean; SDev; skew; P(w/w); P(w/d); and number of wet days (NWD) as introduced in Table 3. Table 5 ranks the model performance, with first rank corresponding to least difference. The three models with the least difference in values for each RCP were selected, with the final selected models provided in Table 6. HADGEM2-AO (Table 4) was excluded from this step since we lacked the necessary computer memory to extract data for this model.

2.3.2. Equilibrium Climate Sensitivity (ECS) Method

Selecting climate models based on the range of highest and lowest ECS values provided by the IPCC (Kattsov et al., 2013) is popular in soil erosion research (e.g. Mullan *et al.*, 2016, 2019). ECS considers changes in water vapour, clouds, lapse rate, and surface albedo to calculate the warming for doubling of atmospheric CO_2 compared to preindustrial climate once a new climatic equilibrium is achieved. Accordingly, ECS has been used to describe the severity of future climatic changes (Knutti *et al.*, 2017).

In keeping with the criteria applied for PPE, ECS values below the 10th percentile and above the 90th percentile of all ECS values for the CMIP5 models were excluded. This provided a pool of 23 different models. The three models nearest to the 10th percentile and the three models nearest to the 90th percentile were selected. Models nearest to the 10th percentile were simulated under RCP4.5, while the models nearest to the 90th percentile were simulated under RCP8.5. The selected ECS models are displayed in Table 6.

2.3.3. Random Selection (RS) Method

Previous climate change impact studies have also included CMIP5 models with little to no justification of selection (e.g. Fazeli Farsani *et al.*, 2019; Sardari *et al.*, 2019; Sha *et al.*, 2019). Consequently, three iterations of random model selection (herein referred to as RS) have been undertaken to determine whether different combinations of models selected at random can provide a wider range in soil erosion and MF diagnostics compared to the carefully tuned methodological approaches applied for PPE and ECS. Of course, a wider range in projections provided by RS would be simply by chance and it does not consider past performance, unlike PPE. The RS models are separated into three groups in Table 6 – Random Group 1 (RG1), Random Group 2 (RG2) and

Random Group 3 (RG3) - randomly assigned as RCP4.5 or RCP8.5. There is noticeably some overlap in models selected from the PPE and ECS selection within the RS in Table 6.

2.4. Spatial Downscaling

Climate information for each model from the *Earth System Grid Federation* (ESGF) (https://esgf-node.llnl.gov/search/esgf-llnl/) is provided at GCM/ ESM grid box scale. These models aim to represent the full Earth system and use RCP scenarios to produce projections of future climate (Hawkins *et al.*, 2013). Spatial downscaling is required to reduce the grid box scale to match the observed climate dimensions and this has been applied to all models (Table 6). The original grid box scale for each model is provided in Appendix 3.

Observed precipitation (1986-2005) was plotted against the ranked quantiles of the reference period (1986-2005) for the selected models on a monthly basis using QQ-plots (Mullan *et al.*, 2019). Polynomial functions were applied to the precipitation data for each model, and appropriate ordering (mostly third order) was applied to each model to avoid clearly anomalous precipitation data points. Alternatively, observed TMAX and TMIN were calibrated using the change factor (CF) approach, as outlined in Hawkins *et al.* (2013). The CF method (Equation 2) changes the simulated modelled output of mean and daily variance by using the observed daily variability (Arnell *et al.*, 2003; Gosling *et al.*, 2009). This method was previously found to be more robust than those using model variability, such as the bias correction method (Hawkins et al., 2013). While the CF approach is widely accepted for calibrating temperature data, the positive definite nature of precipitation makes calibration more complex (Hawkins *et al.*, 2013).

$$T_{CF}(t)=\overline{T_{RAW}}+\frac{\sigma_{T,RAW}}{\sigma_{T,REF}}\left(O_{REF}(t)-\overline{T_{REF}}\right), \qquad (2)$$

where $T_{CF}(t)$ is the change factor for temperature; $\overline{T_{RAW}}$ is the raw modelled temperature for a future period and $\overline{T_{REF}}$ is the observed, where the bar above the symbol represents the time mean; $\sigma_{T,RAW}$ indicates the standard deviation of the daily raw model output for the future period and $\sigma_{T,REF}$ indicates the standard deviation of the daily model output for the reference period ; O_{REF} indicates the daily observations.

2.5. Temporal Downscaling

Temporal downscaling is also required to generate daily scenarios from monthly scenarios, which is necessary to perturb CLIGEN within WEPP. Temporal downscaling was applied to all models in Table 6. Raw historical (1986-2005) and future (2006-2100) precipitation data from the selected models were downscaled to produce daily scenarios. Raw TMAX and TMIN data for the historical and future periods were also downscaled to produce daily scenarios, as needed for WEPP simulations.

Transitional probabilities (Pw/w and Pw/d) were determined by categorising historical precipitation into wet months, dry months, and all months. Wet months were defined when monthly precipitation totals equalled or exceeded the 90th percentile of the mean monthly precipitation totals for each respective month during the reference period (1986-2005). Dry months were defined when monthly precipitation totals did not meet this percentile value. Linear relationships were established between historical monthly precipitation totals and the transitional probabilities for wet months, dry months, and all months. These transfer functions were forced with future monthly precipitation totals to calculate future transitional probabilities. Mean P was calculated following the method in Zhang *et al.* (2004). Equation 3 was applied to calculate the unconditional probability of precipitation occurrence (π):

$$\Pi = \frac{P_w / d}{1 + \frac{P_w}{d} - P_w / d} \qquad (3)$$

the new $Mean\,P$ is then calculated using Equation 4:

$$Mean\,P = \frac{R_m}{N_d \pi} \qquad (4)$$

where $Mean\,P$ is described previously, R_m is the projected mean precipitation totals for a given month, and $N_d \pi$ is the expected number of wet days in the month.

Table 7 is adapted from Mullan *et al.* (2019) to detail how the CLIGEN parameters were adjusted to represent future climate changes. Aside from Mean P, P(W/W), P(W/D), AV TMAX, and AV TMIN, all remaining CLIGEN parameter monthly values were calculated by developing linear relationships using the historical data (1986-2005). A summary of all steps described for each model selection method is provided in Figure 2.

3. Results

3.1. Mean Annual Changes

Figures 3 and 4 demonstrate that PPE provides the widest range in mean annual precipitation response projected from six separate CMIP5 models, with ‘dry’ and ‘wet’ models determining the soil erosion and MF diagnostic response.

PPE consistently projects a wider spread in future scenarios compared to ECS at both hillslopes. At Hillslope 1, the range (highest minus lowest model value) in sediment yield projections for PPE is 1.9 t/ha higher (an increase by 271%) than ECS projections, while the range in soil loss is 1 kg/m^2 higher for PPE compared to ECS (an increase by 243%). Similar observations in sediment yield and soil loss are illustrated for Hillslope 2 – the ranges in sediment yield and soil loss projected by PPE are 239% and 216%, respectively, higher than ECS. Differences in runoff projections are marginally closer, as PPE projects an increase compared to ECS by 39% and 173% at Hillslopes 1 and 2, respectively.

RG1 and RG2 projections are marginally closer to PPE. At Hillslope 1, the range in PPE sediment yield is higher than RG1 and RG2 by 26% and 37%, respectively. Observations at Hillslope 2 are similar – 25% and 23% higher than PPE for RG1 and RG2, respectively. Differences in soil loss projections between both methods closely reflect differences in sediment yield observations. While PPE also demonstrates the widest range in runoff at Hillslope 2, RG2 provides the widest range at Hillslope 1. PPE runoff projections reveal a narrower model spread by 20% compared to RG2 at Hillslope 1.

3.2. Return Periods

Return period analysis reveals important information concerning the frequency and magnitude of events. Return period intervals of 2, 5, 10, 20, 25, 50 and 100 years are displayed for each model selection method in Table 8.

As for mean annual projections (Figures 3 and 4), Table 8 reveals that PPE frequently generates the widest model response for sediment yield at both hillslopes. While this is typically observed for all return period intervals, differences in projections between each selection method are most clearly represented at a 1 in 100-year event. For a 1 in 100-year event at Hillslope 2, the range in sediment yield for PPE is higher than ECS, RG1, and RG2 by 20.2 t/ha, 18.9 t/ha, and 11.5 t/ha, respectively. RG2 sediment yield return periods are closest to PPE, with only marginal decreases from PPE at

Hillslope 2. However, unlike for Hillslope 2, PPE is unable to provide the widest spread in sediment yield projections for all return periods at Hillslope 1. This is observed for RG1 where the range in sediment yield is higher than PPE by 3.7 t/ha, despite RG1 projecting among the lowest sediment yield at Hillslope 2 (28.8 t/ha).

Table 8 also reveals that RG2 generates the widest response in daily precipitation for all return period intervals at both hillslopes, closely followed by PPE. The highest difference between these methods is observed for a 1 in 50-year event, where daily precipitation for RG2 exceeds PPE by 11.1 mm. ECS consistently projects the lowest daily precipitation for all return periods at both hillslopes.

4. Discussion

4.1. PPE Success

As shown in Figures 3 and 4, selecting climate models based on most increased wetness and least increased wetness distinctly provides a broader range in projected soil erosion diagnostics compared to selecting models based on highest and lowest ECS values. PPE also largely demonstrates a wider spread in projections compared to almost all random scenarios, with only minor exceptions. To this end, PPE is successful in generating the widest range of sensible future scenarios, which has not been achieved elsewhere for modelling soil erosion by water.

4.2. Precipitation Variability Drives Soil Erosion Response

Mullan *et al.* (2019) previously found that model projections of rainfall intensity (Mx.5 P) correlated very strongly with projected sediment yield for a hillslope in Flanders, such that this variable alone could confidently explain future sediment yield projected by each model. In this research, differences in the range of sediment yield and soil loss for each model selection method is best explained by both SDEV P and Skew P together, where Figure 5 closely resembles sediment yield and soil loss results in Figures 3 and 4. The remaining precipitation variables (e.g., Mx.5 P, Mean P, P(W/D), P(W/W), Time Pk) make relatively minor contributions in separating the range in projections provided by model selection methods. This is consistent with findings in Zhang (2012) where a larger SKEW P combined with a larger SDEV P typically provide more events with greater magnitudes of daily precipitation in WEPP, while the opposite was true with a smaller SKEW P and SDEV P. In capturing models that project the most and least increased wetness, PPE returns the widest range in

SDEV P and Skew P, which in turn provides the widest range in sediment yield and soil loss projections. While the impact of SDEV P is less clear for runoff, there also appears to be some correlation in runoff distributions when comparing Figures 3 and 4 to Figure 5. These results suggest that the variability in the probability of a wet day occurring within a given month largely determines the response for key soil erosion diagnostics in WEPP.

Though RG2 narrowly provided a wider runoff distribution than PPE at Hillslope 1, precipitation variability also adequately explains the wide range generated by RG2 and PPE compared to other methods. While the model difference in RG2 mean annual precipitation is considerably higher (80mm) than PPE (Table 9), all remaining model selection groups also provide higher mean annual precipitation ranges such that precipitation amount alone cannot explain these results. Instead, mean annual precipitation and Skew P together explain this runoff response. Model distributions of Skew P for PPE and RG2 are considerably higher than all remaining methods (Table 9). For instance, while ECS displays a higher range in mean annual precipitation and a similar SDEV P, a much lower Skew P value (Table 9) dictates that model variance in runoff projections is 2.88 mm and 5.52 mm lower than PPE and RG2, respectively, at Hillslope 1 (Figure 3). Differences in SKEW P also determine the increased range in daily precipitation for RG2 and PPE compared to other selection methods for all return periods (Table 9).

4.3. Impact of Hillslope Characteristics

Differences in hillslope characteristics may support explanations for variance in runoff projections between PPE and RG2 at both hillslopes. As introduced in Table 1, Hillslope 1 has an average slope gradient of 8.1° and a 21 m wide grass buffer strip, while Hillslope 2 has an average slope gradient of 5° and a 3.1 m wide grassed waterway. It is possible that the steeper slope gradient at Hillslope 1 accentuates the impact of precipitation amount to determine runoff response. It is also intuitive to suggest that the expansive area of grassland is more effective in reducing runoff volumes under lower precipitation amounts. The role of grassland in increasing surface roughness has been previously discussed for similar studies (Evrard *et al.*, 2008), with runoff reductions of around 90% reported elsewhere (Schmitt *et al.*, 1999). Consequently, a higher range in mean annual precipitation may ensure that RG2 projects a wider model spread in runoff response compared to PPE at Hillslope 1, despite a narrower range in Skew P. Conversely, the reduced slope gradient at Hillslope 2 may decrease the influence of mean annual precipitation in determining runoff response, such that the impact of a higher Skew P range for PPE may become more dominant. The 3.1 m wide grassed waterway may be insufficient to play any discernible role in reducing runoff volumes.

Similar analysis may be attributed to the higher sediment yield projections for a 1 in 100-year event at Hillslope 1 for RG1 compared to PPE. RG1 generates the widest response in precipitation amount of all methods, while possessing moderately high SDEV P values (Table 9). These values may elicit a higher sediment yield response at the steeper sloped Hillslope 1, while the relative influence of a moderately low range in Skew P increases at the more gently sloped Hillslope 2 (Table 9). The influence of these precipitation variables for different hillslope gradients remains speculative and should be further studied to confidently ascertain anomalies observed for RG1 and RG2.

4.4. Addressing Limitations

Unlike PPE, ECS values provided by the IPCC (Kattsov et al., 2013) do not discriminate between different ensemble members and the range of initial conditions provided by each model. To this end, ECS is inherently unlikely to be capable of generating the same range in projections as PPE with a smaller pool of available models. However, this caveat does not limit results in this study. Instead, this supports PPE for selecting suitable climate models to model geomorphic processes determined by precipitation, such as soil erosion on cultivated fields.

As suggested in Section 4.3, the relative influence of mean annual precipitation amounts and the variability in monthly precipitation together with hillslope characteristics (e.g., slope gradient; land management) in providing runoff discrepancies between hillslopes in WEPP should be further investigated. Results in this study could not conclusively determine the cause for performance differences in model selection groups for separate hillslopes.

It should be cautioned that this study only compared ECS and random model selection against PPE, considering that the former two methods are commonly chosen for hydrological climate change impact studies. Having demonstrated success in this study, it may be worthwhile to compare PPE against other popular selection approaches (e.g. Houle *et al.*, 2012; Evans *et al.*, 2013). However, since PPE represents a blend of two sensible selection approaches and is carefully tailored for processes driven by large scale precipitation events, the wide range provided by the selected models for PPE can be considered robust.

As noted in several studies (Boardman & Vandaele, 2016; Butler, 2005; Mullan, 2013; Mullan *et al.*, 2012; Verstraeten *et al.*, 2003), a range of future land use changes should also be included in future research investigating climate change impacts on soil erosion and MF. This will enable adequate stress-testing of the resilience and adaptation of current mitigation measures (e.g. Mullan *et al.*, 2016, 2019), or perhaps identify a need for mitigation where currently absent.

5. Conclusions and Implications

Previous climate change – soil erosion impact studies typically applied a single model or a few models to drive erosion models with little justification for selection. This approach to climate model selection limits the provision of vital information for mitigation planning and adaptation by increasing the risk of simulating a narrow range of possible scenarios. The PPE method devised in this study is successful for efficiently selecting suitable climate models to simulate soil erosion and MF, which has not been achieved elsewhere for modelling soil erosion by water. The highest range in future (2081-2100) mean annual sediment yield and soil loss was projected by PPE, while the range in projected runoff was also among the highest of all methods at both hillslopes. Return period analysis largely reflects mean annual results.

No single precipitation variable could explain mean annual sediment yield, soil loss, and runoff results. Instead, the standard deviation (SDEV P) and skewness (Skew P) of precipitation together most closely replicate the distribution statistics of sediment yield and soil loss for all methods at both hillslopes. When coupled with mean annual precipitation amounts, mean annual runoff results could be reasonably explained by Skew P values for all methods at both hillslopes.

PPE model selection allows for adequate preparation for the worst- and best-case scenarios at the study area by generating the broadest range in projections for key soil erosion and MF diagnostics. A sensible range is generated by PPE, since PPE blends and precisely transforms both envelope-based and past-performance approaches for specific application to soil erosion and MF. Relevant impact sectors such as soil erosion and MF, and other hydrological phenomena should consider applying this method to examine the impact of future climatic changes.

Conflict of Interest Statement

The authors do not wish to declare any conflict of interest.

References

Alduchov, O. A., & Eskridge, R. E. (1996). Improved Magnus form approximation of saturation vapor pressure. *Journal of Applied Meteorology*, *(35)4*, 601–609. https://doi.org/https://doi.org/10.2172/548871

Arnell, N. W., Hudson, D. A., & Jones, R. G. (2003). Climate change scenarios from a regional climate model: Estimating change in runoff in southern Africa. *Journal of Geophysical Research D: Atmospheres*, *108*(16). https://doi.org/10.1029/2002jd002782

Ascough, J. ., Baffaut, C., Nearing, M. ., & Flanagan, D. . (1995). Watershed Model Channel Hydrology and Erosion Processes. In *USDA Water Erosion Prediction Project: Hillslope Profile and Watershed Model Documentation.*

Beckers, V., Beckers, J., Vanmaercke, M., Van Hecke, E., Van Rompaey, A., & Dendoncker, N. (2018). Modelling farm growth and its impact on agricultural land use: A country scale application of an agent-based model. *Land*, *7*(3). https://doi.org/10.3390/land7030109

Bielders, C. L., Ramelot, C., & Persoons, E. (2003). Farmer perception of runoff and erosion and extent of flooding in the silt-loam belt of the Belgian Walloon Region. *Environmental Science and Policy*, *6*(1), 85–93. https://doi.org/10.1016/S1462-9011(02)00117-X

Boardman, J. (2006). Soil erosion science: Reflections on the limitations of current approaches. *Catena*, *68*(2–3), 73–86. https://doi.org/10.1016/j.catena.2006.03.007

Boardman, J., & Vandaele, K. (2016). Effect of the spatial organization of land use on muddy flooding from cultivated catchments and recommendations for the adoption of control measures. *Earth Surface Processes and Landforms*, *41*(3), 336–343. https://doi.org/10.1002/esp.3793

Boardman, J., & Vandaele, K. (2020). Managing muddy floods: Balancing engineered and alternative approaches. *Journal of Flood Risk Management*, *13*(1), 1–10. https://doi.org/10.1111/jfr3.12578

Bridges, E. M., & Oldeman, L. R. (1999). Global assessment of human-induced soil degradation. *Arid Soil Research and Rehabilitation*, *13*(4), 319–325. https://doi.org/10.1080/089030699263212

Butler, J. J. (2005). Muddy Flooding on the South Downs. *Online Papers Archived by the Institute of Geography*, *School of Geosciences*, *University of Edinburgh*. http://www.era.lib.ed.ac.uk/handle/1842/830

Evans, J. P., Ji, F., Abramowitz, G., & Ekström, M. (2013). Optimally choosing small ensemble members to produce robust climate simulations. *Environmental Research Letters*, *8*(4). https://doi.org/10.1088/1748-9326/8/4/044050

Evrard, O. (2008). *Muddy floods in the Belgian loess belt : Problems and solutions*.

Evrard, O., Bielders, C. L., Vandaele, K., & van Wesemael, B. (2007a). Spatial and temporal variation of muddy floods in central Belgium, off-site impacts and potential control measures. *Catena*, *70*(3), 443–454. https://doi.org/10.1016/j.catena.2006.11.011

Evrard, O., Persoons, E., Vandaele, K., & van Wesemael, B. (2007b). Effectiveness of erosion mitigation measures to prevent muddy floods: A case study in the Belgian loam belt. *Agriculture, Ecosystems and Environment*, *118*(1–4), 149–158. https://doi.org/10.1016/j.agee.2006.02.019

Evrard, O., Vandaele, K., van Wesemael, B., & Bielders, C. L. (2008). A grassed waterway and earthen dams to control muddy floods from a cultivated catchment of the Belgian loess belt. *Geomorphology*, *100*(3–4), 419–428. https://doi.org/10.1016/j.geomorph.2008.01.010

Fazeli Farsani, I., Farzaneh, M. R., Besalatpour, A. A., Salehi, M. H., & Faramarzi, M. (2019). Assessment of the impact of climate change on spatiotemporal variability of blue and green water resources under CMIP3 and CMIP5 models in a highly mountainous watershed. *Theoretical and Applied Climatology*, *136*(1–2), 169–184. https://doi.org/10.1007/s00704-018-2474-9

Finger, D., Heinrich, G., Gobiet, A., & Bauder, A. (2012). Projections of future water resources and their uncertainty in a glacierized catchment in the Swiss Alps and the subsequent effects on hydropower production during the 21st century. *Water Resources Research*, *48*, 1–20. https://doi.org/10.1029/2011WR010733

Flanagan, D. C., & Nearing, M. A. (1995). USDA-Water Erosion Prediction Project: Hillslope profile and watershed model documentation. *Nserl Rep*, *10*(July), 1–123.

Gosling, S. N., McGregor, G. R., & Lowe, J. A. (2009). Climate change and heat-related mortality in six cities Part 2: Climate model evaluation and projected impacts from changes in the mean and variability of temperature with climate change. *International Journal of Biometeorology*, *53*(1), 31–51. https://doi.org/10.1007/s00484-008-0189-9

Hawkins, E., Osborne, T. M., Ho, C. K., & Challinor, A. J. (2013). Calibration and bias correction of climate projections for crop modelling: An idealised case study over Europe. *Agricultural and Forest Meteorology*, *170*, 19–31. https://doi.org/10.1016/j.agrformet.2012.04.007

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., & New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006. *Journal of Geophysical Research Atmospheres*, *113*(20). https://doi.org/10.1029/2008JD010201

Heitz, C., Spaeter, S., Auzet, A. V., & Glatron, S. (2009). Local stakeholders' perception of muddy flood risk and implications for management approaches: A case study in Alsace (France). *Land Use Policy*, *26*(2), 443–451. https://doi.org/10.1016/j.landusepol.2008.05.008

Houle, D., Bouffard, A., Duchesne, L., Logan, T., & Harvey, R. (2012). Projections of future soil temperature and water content for three Southern Quebec forested sites. *Journal of Climate*, *25*(21), 7690–7701. https://doi.org/10.1175/JCLI-D-11-00440.1

IPCC. (2013). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. In T. F. Stocker, D. Qin, G. K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, & P. M. Midgley (Eds.), *Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA*. https://doi.org/https://doi.org/10.1017/CBO9781107415324

Kattsov, V., Federation, R., Reason, C., Africa, S., Uk, A. A., Uk, T. A., Baehr, J., Uk, A. B., Catto, J., Canada, J. S., & Uk, A. S. (2013). Evaluation of climate models. *Climate Change 2013 the Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, *9781107057*, 741–866. https://doi.org/10.1017/CBO9781107415324.020

Knutti, R., Rugenstein, M. A. A., & Hegerl, G. C. (2017). Beyond equilibrium climate sensitivity. *Nature Geoscience*, *10*(10). https://doi.org/https://doi.org/10.1038/nclimate1716

Kundzewicz, Z. W., Mata, L. J., Arnell, N. W., Döll, P., Jimenez, B., Oki, T., Şen, Z., & Shiklomanov, I.

(2009). *The implications of projected climate change for freshwater resources and their management resources and their management*. *6667*. https://doi.org/10.1623/hysj.53.1.3

Laflen, J., Lane, L., Water, G. F.-J. of soil and, & 1991, U. (1991). WEPP:A new generation of erosion prediction technology. *Jswconline.Org*, *46*(1), 34–38. http://www.jswconline.org/content/46/1/34.short

Lenderink, G., & Van Meijgaard, E. (2008). Increase in hourly precipitation extremes beyond expectations from temperature changes. *Nature Geoscience*, *1*(8), 511–514. https://doi.org/10.1038/ngeo262

Li, P., Mu, X., Holden, J., Wu, Y., Irvine, B., Wang, F., Gao, P., Zhao, G., & Sun, W. (2017). Comparison of soil erosion models used to study the Chinese Loess Plateau. *Earth-Science Reviews*, *170*, 17–30. https://doi.org/10.1016/j.earscirev.2017.05.005

Li, Z., & Fang, H. (2016). Impacts of climate change on water erosion: A review. *Earth-Science Reviews*, *163*, 94–117. https://doi.org/10.1016/j.earscirev.2016.10.004

Lu, X., Siemann, E., Shao, X., Wei, H., & Ding, J. (2013). Climate warming affects biological invasions by shifting interactions of plants and herbivores. *Global Change Biology*, *19*(8), 2339–2347. https://doi.org/10.1111/gcb.12244

Lutz, A. F., Maat, W., Biemans, H., & Shrestha, A. B. (2016). Selecting representative climate models for climate change impact studies : an advanced envelope-based selection approach. *International Journal of Climatology*, *4005*(January), 3988–4005. https://doi.org/10.1002/joc.4608

Maeda, E. E., Pellikka, P. K. E., Siljander, M., & Clark, B. J. F. (2010). Potential impacts of agricultural expansion and climate change on soil erosion in the Eastern Arc Mountains of Kenya. *Geomorphology*, *123*(3–4), 279–289. https://doi.org/10.1016/j.geomorph.2010.07.019

Minville, M., Brissette, F., & Leconte, R. (2008). *Uncertainty of the impact of climate change on the hydrology of a nordic watershed*. 70–83. https://doi.org/10.1016/j.jhydrol.2008.05.033

Mullan, D. (2013). Soil erosion under the impacts of future climate change: Assessing the statistical significance of future changes and the potential on-site and off-site problems. *Catena*, *109*, 234–246. https://doi.org/10.1016/j.catena.2013.03.007

Mullan, D., Favis-Mortlock, D., & Fealy, R. (2012). Addressing key limitations associated with modelling soil erosion under the impacts of future climate change. *Agricultural and Forest Meteorology*, *156*, 18–30. https://doi.org/10.1016/j.agrformet.2011.12.004

Mullan, D., Matthews, T., Vandaele, K., Barr, I. D., Swindles, G. T., Meneely, J., Boardman, J., & Murphy, C. (2019). Climate impacts on soil erosion and muddy flooding at 1.5 versus 2°C warming. *Land Degradation and Development*, *30*(1), 94–108. https://doi.org/10.1002/ldr.3214

Mullan, D., Vandaele, K., Boardman, J., Meneely, J., & Crossley, L. H. (2016). Modelling the effectiveness of grass buffer strips in managing muddy floods under a changing climate. *Geomorphology*, *270*, 102–120. https://doi.org/10.1016/j.geomorph.2016.07.012

Nearing, M. A. (1998). Why soil erosion models over-predict small soil losses and under-predict large soil losses. *Catena*, *32*(1), 15–22. https://doi.org/https://doi.org/10.1016/S0341-8162(97)00052-0

Nearing, M. A. (1990). Sensitivity analysis of the WEPP hillslope erosion model. *Transactions of the ASAE*, *33*(3), 839–849.

Nearing, M. A., Jetten, V., Baffaut, C., Cerdan, O., Couturier, A., Hernandez, M., Le Bissonnais, Y., Nichols, M. H., Nunes, J. P., Renschler, C. S., Souchère, V., & Van Oost, K. (2005). Modeling response of soil erosion and runoff to changes in precipitation and cover. *Catena*, *61*(2-3 SPEC. ISS.), 131–154. https://doi.org/10.1016/j.catena.2005.03.007

Nicks, A. D., Lane, L. J., & Gander, G. A. (1995). Chapter 02. Weather generator. *USDA–Water Erosion Prediction Project Hillslope Profile and Watershed Model Documentation*, *July*, 2.1–2.22.

Panagos, P., Borrelli, P., Poesen, J., Ballabio, C., Lugato, E., Meusburger, K., Montanarella, L., & Alewell, C. (2015). The new assessment of soil loss by water erosion in Europe. *Environmental Science and Policy*, *54*, 438–447. https://doi.org/10.1016/j.envsci.2015.08.012

Pimentel, D. (2006). Soil erosion: A food and environmental threat. *Environment, Development and Sustainability*, *8*(1), 119–137. https://doi.org/10.1007/s10668-005-1262-8

Risbey, J. S., & Entekhabi, D. (1996). Observed Sacramento Basin streamflow response to climate impact studies. *Journal of Hydrology*, *184*, 209–223. https://doi.org/https://doi.org/10.1016/0022-1694(95)02984-2

Sardari, M. R. A., Bazrafshan, O., Panagopoulos, T., & Sardooi, E. R. (2019). Modeling the impact of climate change and land use change scenarios on soil erosion at the minab dam watershed. *Sustainability (Switzerland)*, *11*(12). https://doi.org/10.3390/su10023353

Schmitt, T. J., Dosskey, M. G., & Hoagland, K. D. (1999). Filter Strip Performance and Processes for Different Vegetation, Widths, and Contaminants. *Journal of Environmental Quality*, *28*(5), 1479–1489. https://doi.org/10.2134/jeq1999.00472425002800050013x

Scholz, G., Quinton, J. N., & Strauss, P. (2008). Soil erosion from sugar beet in Central Europe in response to climate change induced seasonal precipitation variations. *Catena*, *72*(1), 91–105. https://doi.org/10.1016/j.catena.2007.04.005

Sha, J., Li, X., & Wang, Z. L. (2019). Estimation of future climate change in cold weather areas with the LARS-WG model under CMIP5 scenarios. *Theoretical and Applied Climatology*, *137*(3–4), 3027–3039. https://doi.org/10.1007/s00704-019-02781-4

Stolpe, N. B. (2005). A comparison of the RUSLE, EPIC and WEPP erosion models as calibrated to climate and soil of south-central Chile. *Acta Agriculturae Scandinavica Section B: Soil and Plant Science*, *55*(1), 2–8. https://doi.org/10.1080/09064710510008568

Sun, Y., Solomon, S., Dai, A., & Portmann, R. W. (2007). How often will it rain? *Journal of Climate*, *20*(19), 4801–4818. https://doi.org/10.1175/JCLI4263.1

Taylor, K. E., Stouffer, R., & Meehl, G. A. (2012). An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, *93 (4)*(November), 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Verstraeten, G., & Poesen, J. (1999). The nature of small-scale flooding, muddy floods and retention pond sedimentation in central Belgium. *Geomorphology*, *29*(3–4), 275–292. https://doi.org/10.1016/S0169-555X(99)00020-3

Verstraeten, G., Poesen, J., Govers, G., Gillijns, K., Van Rompaey, A., & Van Oost, K. (2003). Integrating science, policy and farmers to reduce soil loss and sediment delivery in Flanders, Belgium. *Environmental Science and Policy*, *6*(1), 95–103. https://doi.org/10.1016/S1462-9011(02)00116-8

Yang, D., Kanae, S., Oki, T., Koike, T., & Musiake, K. (2003). Global potential soil erosion with reference to land use and climate changes. *Hydrological Processes*, *17*(14), 2913–2928. https://

doi.org/10.1002/hyp.1441

Yu, B. (2003). An assessment of uncalibrated CLIGEN in Australia. *Agricultural and Forest Meteorology*, *119*(3–4), 131–148. https://doi.org/10.1016/S0168-1923(03)00141-2

Zhang, X. C. (2012). Verifying a temporal disaggregation method for generating daily precipitation of potentially non-stationary climate change for site-specific impact assessment. *International Journal of Climatology*, *33*(2), 326–342. https://doi.org/10.1002/joc.3425

Zhang, X. C. (2013). Adjusting skewness and maximum 0.5 hour intensity in CLIGEN to improve extreme event and sub-daily intensity generation for assessing climate change impacts. *Transactions of the ASABE*, *56*(5), 1703–1713. https://doi.org/10.13031/trans.56.10004

Zhang, X. C., Liu, W. Z., Li, Z., & Zheng, F. L. (2009). Simulating site-specific impacts of climate change on soil erosion and surface hydrology in southern Loess Plateau of China. *Catena*, *79*(3), 237–242. https://doi.org/10.1016/j.catena.2009.01.006

Zhang, X. C., Nearing, M. A., Garbrecht, J. D., & Steiner, J. L. (2004). Downscaling Monthly Forecasts to Simulate Impacts of Climate Change on Soil Erosion and Wheat Production. *Soil Science Society of America Journal*, *68*(4), 1376–1385. https://doi.org/10.2136/sssaj2004.1376

Tables

Table 1: Hillslope characteristics.

	Slope Length (m)	Mean Slope Gradient (°)	Crop Cover	Mitigation Measure
Hillslope 1	65.0	8.1	Potatoes; Maize; Sugar Beet (44 m)	Grass Buffer Strip (21 m)
Hillslope 2	132.7	5.0	Potatoes; Maize; Sugar Beet (129.6 m)	Grassed Waterway (3.1 m)

Table 2: Mean measured soil input parameters at each hillslope.

	Sand (%)	Silt (%)	Clay (%)	OM (%)
Hillslope 1	15.1	76.7	8.3	4.1
Hillslope 2	14.3	78	7.7	3.9

Table 3: Description of CLIGEN input parameters and associated nomenclature (Mullan *et al.*, 2019).

Parameter		Unit
Mean daily precipitation for each wet day for a given month	Mean P	in
Standard deviation of Mean P for a given month	SDev P	in
Skewness of Mean P for a given month	Skew P	in
Conditional probability of a wet day following a wet day for a given month	Pw/w	%
Conditional probability of a wet day following a dry day for a given month	Pw/d	%
Mean maximum temperature for a given month	AV TMAX	°F
Mean minimum temperature for a given month	AV TMIN	°F
Standard deviation of TMAX for a given month	SD TMAX	°F
Standard deviation of TMIN for a given month	SD TMIN	°F
Mean solar radiation for a given month	SOL.RAD	L/d[a]
Standard deviation of SOL.RAD for a given month	SD SOL	L/d[a]
Mean maximum half hourly precipitation for a given month	MX.5P	in
Mean dew point temperature for a given month	DEW PT	°F
Time to peak storm intensity	Time Pk	[b]
Mean percent of time that wind blows from 1 of 16 cardinal directions for a given month	% DIR [c]	%
Mean wind speed related to % DIR [c] for a given month	MEAN	m/s^{-1}
Standard deviation of MEAN for a given month	SDev MEAN	m/s^{-1}
Skewness of MEAN for a given month	Skew MEAN	m/s^{-1}
Mean percent of days that mean wind speed is less than 1 ms^{-1} for a given month	CALM	%

Table 4: Selected models for RCP4.5 and RCP8.5. Models are ordered by distance to percentile, ‘1’ representing least distance and ‘3’ most distance. All models are r1i1p1 unless otherwise stated.

	RCP4.5			RCP8.5		
	1	2	3	1	2	3
Wet	GISS-E2-R r6i1p3	GISS-E2-R r6i1p1	MRI-CGCM3	IPSL-CM5A-LR r4i1p1	ACCESS1-3	IPSL-CM5A-LR r2i1p1
Dry	HADGEM2-AO	CNRM-CM5	HADGEM2-ES r2i1p1	CanESM2 r3i1p1	HADGEM2-ES r2i1p1	HADGEM2-ES

Table 5: Variance between observed and historical modelled data for models selected in Section 2.3.1. April to September are selected for analysis since these months were previously considered as key months for MF (e.g. Mullan et al., 2016). A rank of 1 equals closest performance to observed.

	April to September							
	Mean	**SDev**	**Skew**	**P(w/w)**	**P(w/d)**	**NWD**	**Sum**	**Rank**
MRI-CGCM3	-0.88	-0.01	-0.40	0.17	0.21	3.40	5.08	1
IPSL-CM5A-LR r2i1p1	-1.51	-0.98	-0.41	0.18	0.12	2.99	6.20	2
HADGEM2-ES-r2i1p1	-2.19	-1.07	0.21	0.16	0.19	3.03	6.86	3
HADGEM2-ES-r1i1p1	-2.04	-1.04	-0.10	0.19	0.17	3.34	6.87	4
IPSL-CM5A-LR r4i1p1	-2.34	-1.75	-0.33	0.13	0.10	2.29	6.93	5
ACCESS1-3	-1.61	0.00	0.30	0.24	0.33	4.71	7.18	6
GISS-E2-R p3 r6i1p3	-1.07	0.31	-0.57	0.30	0.38	5.66	8.29	7
GISS-E2-R p1 r6i1p1	-0.96	0.57	-0.45	0.30	0.40	5.73	8.41	8
CNRM-CM5	-1.46	-0.70	1.51	0.24	0.30	4.58	8.78	9
CanESM2 r3i1p1	-2.40	-1.18	0.81	0.26	0.33	4.82	9.80	10

Table 6: Selected models from the PPE, ECS, and RG 1-3. Models are separated by RCP 4.5 and 8.5, selected randomly for RG 1-3. Otherwise, the order of the selected models within each RCP grouping displayed is random. Unless otherwise stated, all models are r1i1p1.

RCP4.5	RCP8.5
PPE	
MRI-CGCM3	IPSL-CM5A-LR r2i1p1
HADGEM2-ES-r2i1p1	HADGEM2-ES-r1i1p1
GISS-E2-R p3 r6i1p3	IPSL-CM5A-LR r4i1p1
ECS	
GFDL-ESM2G	GFDL-CM3
GFDL-ESM2M	ACCESS1-0
GISS-E2-H	CSIRO-Mk3-6-0
RG1	
GISS-E2-R p3 r6i1p3	IPSL-CM5A-LR r4i1p1
HADGEM2-ES-r2i1p1	HADGEM2-ES-r1i1p1
IPSL-CM5B-LR	CanESM2 r3i1p1
RG2	
MRI-CGCM3	ACCESS1-0
GFDL-ESM2M	GFDL-CM3
IPSL-CM5A-MR	CNRM-CM5
RG3	
GISS-E2-H	MIROC5
IPSL-CM5A-MR	IPSL-CM5A-LR r2i1p1
GFDL-ESM2G	CSIRO-Mk3-6-0

Table 7: Details of modifications required for key CLIGEN parameters to represent future climate changes.

CLIGEN Parameter	Derivation Method
Mean P	Equations 3 and 4
SDev P	Calculated from future Mean P
SKEW P	Calculated from future Q99
P(W/W)	See Section 2.5.
P(W/D)	See Section 2.5.
AV TMAX	Modified from future AV TMAX
AV TMIN	Modified from future AV TMIN
TMAX SD	Calculated from future AV TMAX
TMIN SD	Calculated from future AV TMIN
SOL.RAD	Linear regression – plotted against future AV TMAX
SD.SOL	Linear regression – plotted against future AV TMAX
MX.5P	Linear regression – plotted against future AV TMIN
DEW PT	Linear regression – plotted against future AV TMIN
Time PK	Linear regression – plotted against future SDev P

Table 8: Comparing the range (highest minus lowest model value) in sediment yield (SY) and daily precipitation (Pr.) projected by each model selection method at Hillslopes 1 and 2 for different return period intervals. The highest projected range for each return period interval has been coloured red.

		Hillslope 1		Hillslope 2	
	Return Period	SY Range (t/ha)	Pr. Range (mm)	SY Range (t/ha)	Pr. Range (mm)
PPE	2	1.7	17.6	9.1	18.7
	5	3.3	27.2	16.0	28.0
	10	5.1	34.8	23.0	35.1
	20	8.0	45.6	31.0	45.6
	25	9.3	53.8	33.8	53.8
	50	12.8	64.4	37.0	64.4
	100	14.2	85.1	47.7	85.1
ECS	2	0.4	6.0	2.8	6.6
	5	1.4	9.0	5.8	9.4
	10	1.9	13.0	7.9	13.0
	20	2.0	15.7	11.0	16.3
	25	2.7	20.0	11.3	20.3
	50	3.4	22.8	16.9	22.8
	100	10.2	30.5	27.5	30.5
RG 1	2	1.5	13.2	7.3	10.1
	5	2.6	20.0	10.9	16.0
	10	4.1	23.3	16.2	20.1
	20	6.5	29.0	23.9	29.6
	25	7.7	32.0	26.8	33.4
	50	10.3	32.3	25.4	31.6
	100	17.9	33.4	28.8	38.9
RG 2	2	1.2	19.9	8.1	20.9
	5	2.8	28.8	15.1	29.6
	10	3.9	37.0	19.2	37.2
	20	5.3	48.6	22.6	48.6
	25	5.7	57.6	23.1	57.6
	50	8.9	75.5	36.9	75.5
	100	12.0	89.8	47.5	89.8
RG 3	2	0.4	10.6	2.8	9.5
	5	1.4	14.2	6.7	13.3
	10	2.3	16.8	9.9	18.2
	20	2.8	20.0	13.9	23.5
	25	3.2	20.5	14.7	28.4
	50	4.8	24.8	22.6	34.7
	100	6.7	38.0	36.2	44.9

Table 9: Ranges (highest minus lowest model value) in mean annual precipitation, SDEV P and Skew P for all model selection groups. SDEV P and Skew P are dimensionless. While the PPE and RG2 models possess among the lowest SDEV P, the higher Skew P appears to separate runoff response for these latter models relative to the remaining models.

	Precipitation (mm)	SDEV P	Skew P
PPE	97.2	0.06	1.92

ECS	184.8	0.06	1.01
RG 1	204.5	0.07	1.58
RG 2	176.9	0.06	1.70
RG 3	159.0	0.08	1.56

Figures

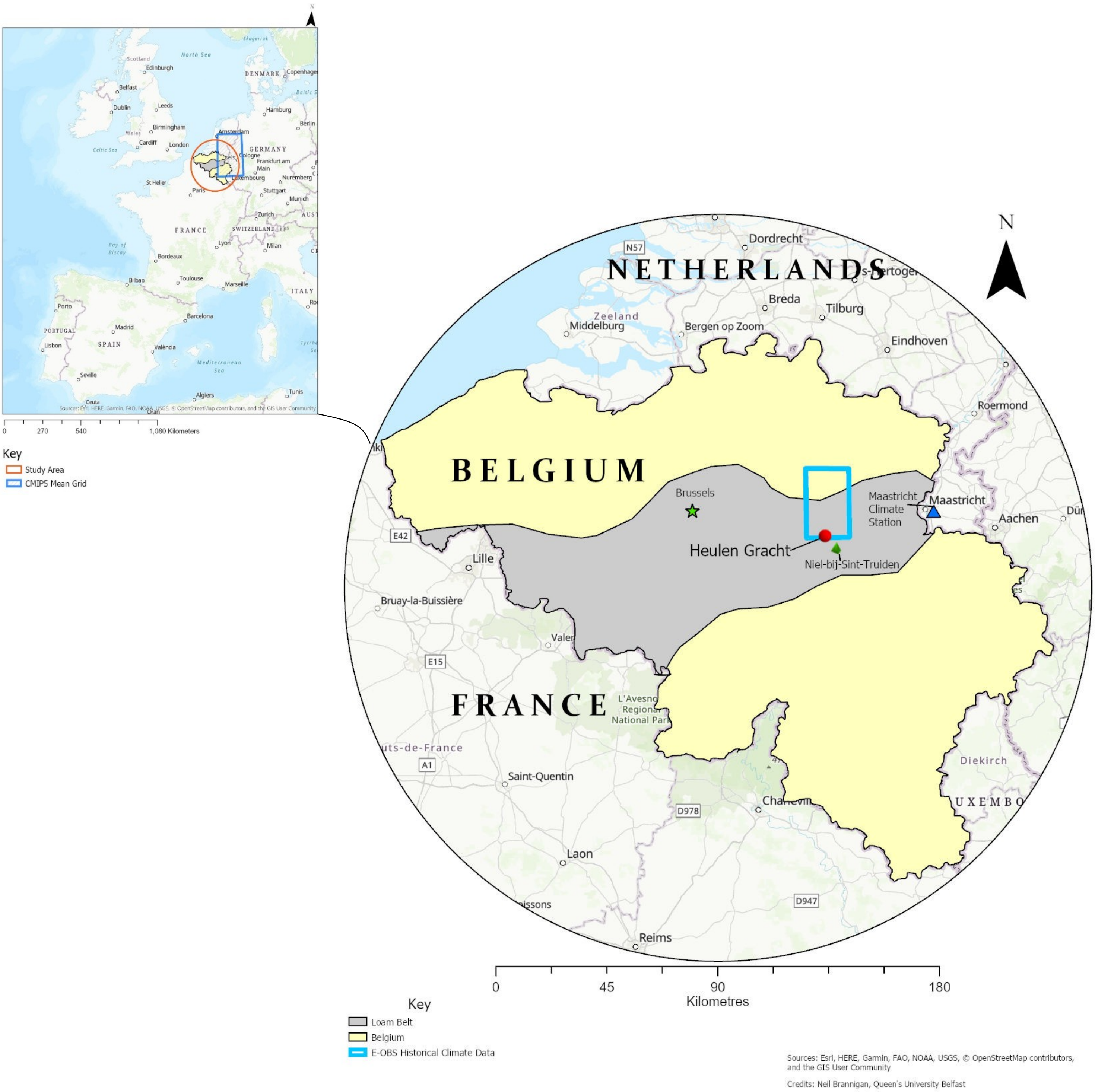

Figure 1: The location of the study area within the Belgian loess belt

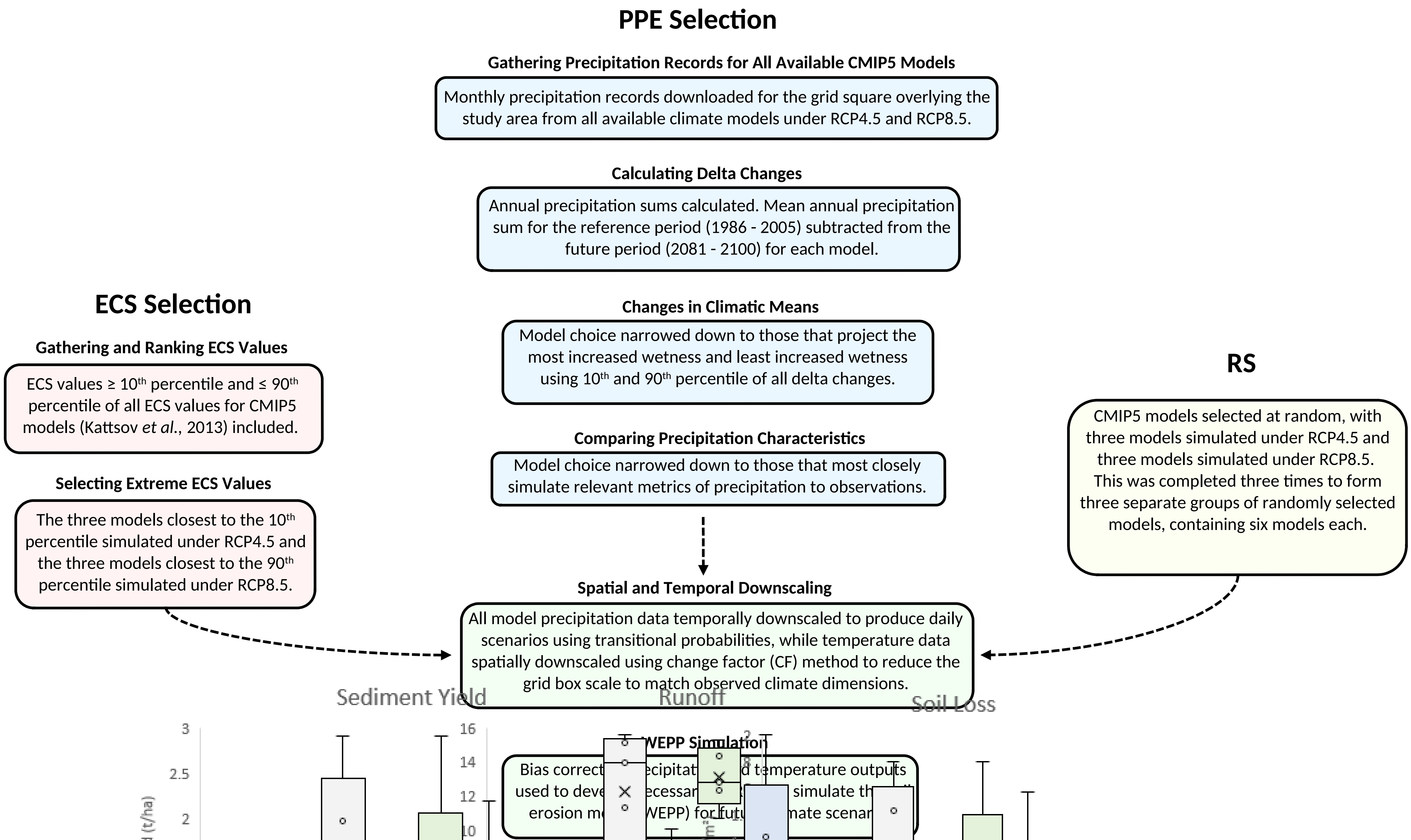

Figure 2: Summary of steps provided for all model selection methods to project climatic conditions for the future period (2081-2100) and simulate in a soil erosion model.

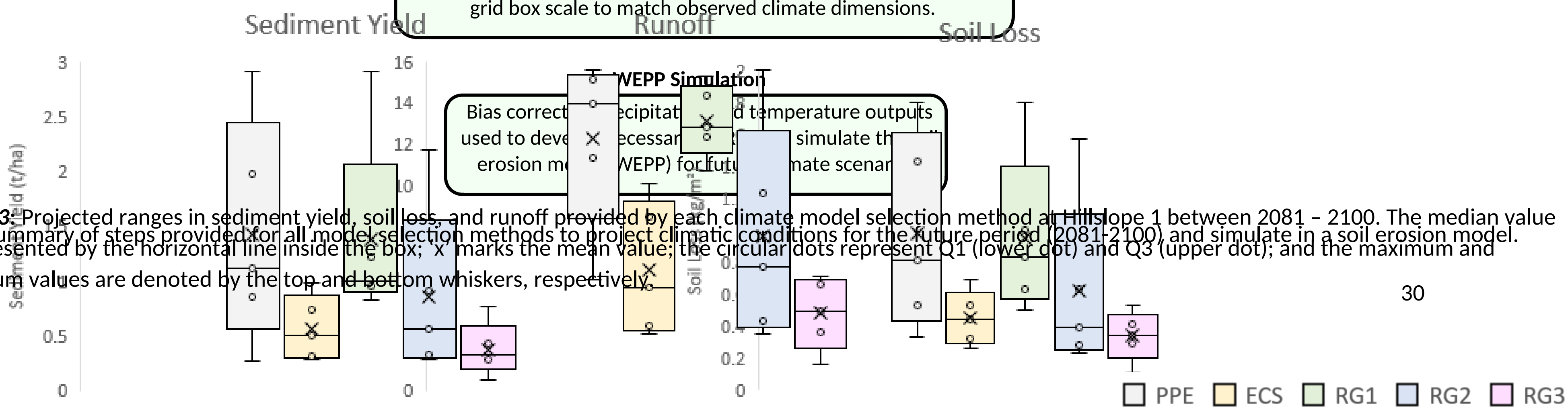

Figure 3: Projected ranges in sediment yield, soil loss, and runoff provided by each climate model selection method at Hillslope 1 between 2081 – 2100. The median value is represented by the horizontal line inside the box; 'x' marks the mean value; the circular dots represent Q1 (lower dot) and Q3 (upper dot); and the maximum and minimum values are denoted by the top and bottom whiskers, respectively.

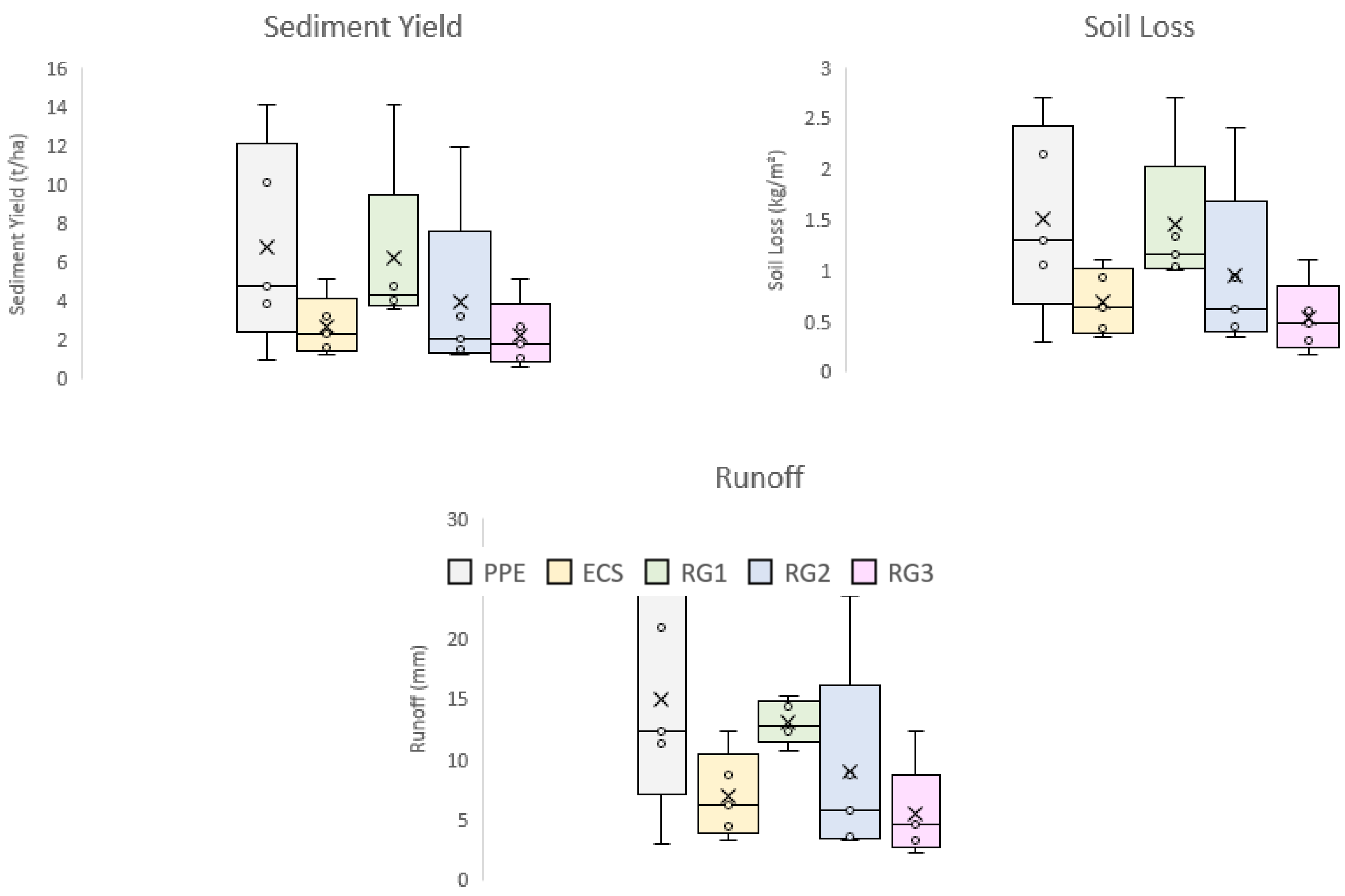

Figure 4: Same as for Figure 3, but for Hillslope 2.

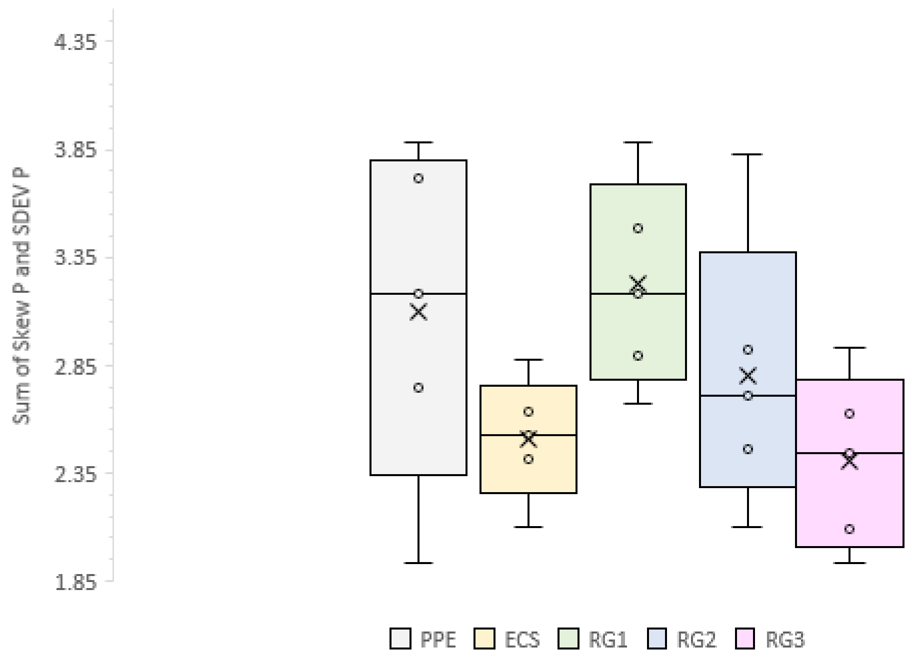

Figure 5: Skew P and SDEV P summed together for each climate model selection method. The distribution of these results closely compares to the sediment yield and soil loss model selection method distributions in Figures 3 and 4.