Formetta, Giuseppe; Prosdocimi, Ilaria; Stewart, Elizabeth; Bell, Victoria. 2018. **Estimating the index flood with continuous hydrological models: an application in Great Britain**.



**The definitive peer-reviewed and edited version of this article is published in *Hydrology Research* 49 (1) 123-133 (2018). 10.2166/nh.2017.251 and is available at www.iwapublishing.com.**

# Estimating the index flood with continuous hydrological models: an application in Great Britain.

Giuseppe Formetta[1], Ilaria Prosdocimi[2], Elizabeth Stewart[1], and Victoria Bell[1]

[1]Centre for Ecology & Hydrology, Maclean Building, Crowmarsh Gifford, Wallingford OX10 8BB

[2]Department of Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY

Estimates of peak river discharge are essential for designing and managing hydraulic infrastructure such as dams, bridges, and flood alleviation schemes. Typically these are derived for an assigned annual exceedance probability (e.g. the 1 in 100 year flood), and the provision of accurate estimates is a critical issue in engineering hydrology, affecting both financial cost and human lives. In the UK, practitioners typically apply the Flood Estimation Handbook (FEH) statistical method which estimates the design flood as the product of a relatively frequent flow estimate (the index flood, IF) and a dimensionless regional growth factor used for estimating peak flows at higher return periods. For gauged catchments the IF is usually estimated from observations as the median annual maximum flow which has a two year return period. For ungauged catchments it is computed through a multiple linear regression model based on a set of morpho-climatic indices of the basin.

While the FEH IF methods provide peak flow estimates that are robust and defensible, they do not readily take into account catchment or rainfall heterogeneity (important for large catchments) or the effect of environmental change on river flows. Successful application to regions outside the UK currently requires a network of good quality, long-

term flow gauges to underpin the design flood method, not always present in less industrialised regions of the world.

With the aim of addressing these limitations, we present and assess a methodology to estimate the IF at national scale using continuous simulation from an area-wide physically-based hydrological model (Grid-to Grid or “G2G”). The new methodology is tested across Great Britain and compares well with estimates of the IF at 550 gauging stations ($R^2$=0.91) and similar performance simulating the annual maxima trend over time. The promising results for Great Britain support the aspiration that continuous simulation from large-scale hydrological models, supported by the increasing availability of global weather, climate and hydrological products, could be used to develop robust methods to help engineers estimate design floods in regions with limited gauge data or affected by environmental change.

## 1. Introduction

An accurate estimate of the design flood, i.e. the peak flow for an assigned probability of exceedance (NERC, 1975), is a critical requirement for reducing the social and economic impact of floods. Floods constitute 40% of worldwide natural disasters (EM-DAT, 2015) and often cause fatalities and damage to houses, businesses and infrastructure. Commonly, design flows are estimated with statistical models fitted to annual maxima (AMAX) measured at a gauged site (flood frequency analysis). Unfortunately hydrological records are often unavailable at the site of interest or, when available, they are too short to allow reliable statistical analyses. To overcome this limitation a standard approach is to adopt a “regionalization” procedure which introduces data from other sites into the flood frequency analysis, chosen on the basis that they exhibit similar hydrological behaviour. The regions from which these sites can be selected are typically defined using one of several different regionalization methods such as cluster analysis, the region of influence approach, the method of residuals and canonical correlation analysis. Several authors review these regionalization methods (Blöschl et al., 2013, Srinivas et al., 2008, Hrachowitz et al., 2013). The index flood method (Dalrymple, 1960) is one of the most popular regionalization procedures among engineers and practitioners (NERC, 1975; Hosking and Wallis, 2005; Institute of Hydrology, 1999). The method is based on the assumption that, for all the sites inside a “hydrologically homogeneous” region, the AMAX frequency distributions are identical apart from a local scaling factor (index flood, IF [$m^3\ s^{-1}$]). This assumption allows the computation of any p-th quantile at any location i-th as:

$$Q_i^p = IF_i \cdot q^p \tag{1}$$

where $IF_i$ is the index flood at location i and $q^p$ [-] is the regional growth curve, a dimensionless quantile function assumed to be identical for all the sites in the region. Various approaches have been developed to provide reliable estimates of IF, and Bocchiola et al., (2003) provides a summary of some of the most widely used. Broadly, if the site of interest is gauged, the IF can be estimated by direct methods, i.e. from the AMAX time series, using the sample mean (Dalrymple, 1960, Hosking and Wallis, 1993, NERC, 1975), the sample median (Robson and Reed, 1999), or using peak over threshold analysis (Chow et al., 1988, Robson and Reed, 1999). If the site of interest is ungauged, a variety of “indirect” methods have been proposed to estimate IF. The most commonly used are empirical methods (Hirsch et al., 1992; Meigh et al., 1997; Kjeldsen and Jones, 2009) that relate the IF evaluated by AMAX measurements to a set of morpho-climatic catchment descriptors such as area, slope, average annual rainfall, land use, etc. These methods include coefficients that are usually estimated by least squares (e.g. Stedinger and Tasker, 1985), maximum-likelihood (e.g. Kjeldsen et al., 2008), and Bayesian methods (e.g. Haddad et al., 2012).The uncertainty in the IF estimate attributable to the data used in the regression model calibration was quantified by Jaafar, W., Zurina and Han, (2012). Other indirect approaches for estimating IF and flow quantiles are based on on the use of artificial neural networks (e.g. Hall et al., 2002; Shu and Burn, 2004; Dawson et al., 2006) or on the connection between stochastic rainfall models and lumped flow routing models (Cordova and Rodriguez Iturbe, 1983; Brath et al., 1992; Calver et al., 2005; Kjeldsen et al., 2005; Rigon et al., 2011). Limitations of the latter modelling approach are: i) the simplified assumptions for the hydrological model component; ii) the requirement of catchment initial moisture conditions; iii) the assumption of high simplified and uniform rainfall storms in catchment.

Indirect estimation of IF based on continuous physically-based hydrological model simulations has also been explored in recent years. The advantages of such an approach include: i) taking into account catchment heterogeneity, ii) accommodation of temporal and spatial rainfall variability, and iii) ability to provide a consistent IF estimate for multiple points on the river network. Demonstrations of the use of continuous, physically-based model simulations for flood frequency analysis are provided for various catchments by Cameron et al., (2000), Calver et al., (2005), Moretti and Montanari, (2008), and Viviroli et al., (2009)., but to our knowledge, only Ravazzani et al., 2015 used continuous hydrological model simulation for estimating the IF. They applied the model FEST-WB (Montaldo et al. 2007, Rabuffetti et al. 2008) to reconstruct river flows for an alpine basin in the north part of Italy and to predict the IF.

For a gauged location, an estimate of the IF recommended by the FEH, is the median of the observed AMAX. This corresponds to the 2 year return period flow which is considered a good estimate of the bankfull river discharge. If less than 14 years of AMAX are available, the FEH suggests use of peak over threshold data. For ungauged sites, the Environment Agency Flood Estimation Guidelines 2012 recommends use of the regression model of Kjeldsen et al. (2008) to estimate the IF. Practitioners are also advised that data transfer from donor catchments to the site of interest can improve the accuracy of IF estimates (Kjeldsen and Jones, 2007). The “donors” are gauged catchments hydrologically similar to the site of interest (i.e. located upstream or downstream on the same river, or possessing similar size and land use).

Here we present a general methodology to estimate the IF at national scale using continuous hydrological simulation (Section 2). This approach aims to: i) integrate the indirect methods for IF estimation and address their limitations for larger and spatially heterogeneous catchments, and ii) provide effective tools for IF estimation in ungauged or poorly gauged catchments. The methodology is tested in Great Britain (Section 3) and assessed (Section 4) by comparison with estimates of the IF at 550 gauging stations.

## 2. Methodology

The area-wide physically-based hydrological model Grid-to-Grid (G2G, Bell et al., 2007a,b; 2009) has been used to estimate the IF at national scale. The G2G typically operates at a 1km$^2$ resolution across Britain and has been configured to represent spatial variability in catchment response. The model uses landscape information provided by gridded spatial datasets of elevation, soil and geology in preference to the identification of model parameters through catchment calibration, and for the application discussed here, a single model configuration and set of parameters is applied across Britain (i.e. with no catchment calibration). G2G model configuration and inputs are discussed in subsection 2.1. Model output consisting of river flow time series at each 1km$^2$ river grid-cell are used to construct maps of AMAX across Britain and to estimate the IF following the FEH methodology (Institute of Hydrology, 1999). Annual maxima in the UK are taken as the highest flow value recorded in a water year, which runs from October to September.

G2G modelled IFs were compared to measured IFs for 550 gauged sites using observations obtained from the National River Flow Archive (NRFA). Modelled and measured IFs were compared using a linear regression, together with an analysis of the sensitivity of model performance to morpho-climatic catchment descriptors. The agreement between the G2G-derived and the measured IF was evaluated by: i) quantifying the coefficient of determination, and ii) assessing the uncertainty in IF estimate using the factorial standard error (Kjeldsen, 2014). Maps of model residuals (differences between modelled and measured IF) provide additional information on regions and types of catchment where the model performs best (and worst). Finally the temporal trends of modelled and measured AMAX were compared to assess the model capability in detecting observed long term trends.

## 2.1 Grid-to-Grid model set-up and input data

The Grid-to-Grid Model (Bell et al., 2007a) is a grid-based hydrological model that simulates surface and sub-surface runoff, lateral movement of soil-moisture, and flow-routing along rivers. Over Britain it is typically applied at a 1km$^2$ grid resolution and a 15-minute time-step, and is configured using spatial datasets of topography, soil, and land cover. Applications include flood forecasting (e.g. Cole and Moore, 2009) and assessment of climate change impacts on floods and snowmelt (i.e. Bell et al., 2007b; Bell et al., 2009; Bell et al., 2016). The most recent version of the model as presented in Bell et al., (2016) was tested over the Great Britain for the period 1960-2011. Driving data consist of daily precipitation observations on a 1 $km^2$ grid, (CEH GEAR: Keller et al., 2015), monthly PE estimates on a 40 $km^2$ grid (MORECS: Hough and Jones, 1997), and daily minimum and maximum temperature observations on a 5$km^2$ grid for 1960–2014 (Perry et al., 2009) which were applied through the day using a sine curve and downscaled to 1 $km^2$ using a lapse rate and elevation data (Morris and Flavin, 1990). Model output consisting of 15 minutes river flows were used to provide AMAX values for 1$km^2$ river grid-cells across Britain.

## 3. Study Area and Data Availability

The study region includes 550 catchments from England, Scotland, and Wales. They are part of the United Kingdom peak flow dataset (version v4.1) obtained from "The National River Flow Archive" (NRFA, 2008; Dixon et al., 2013) and available at http://nrfa.ceh.ac.uk/. For the purposes of this analysis we used the instantaneous peak flow AMAX values and a set of catchment descriptors consisting of: the catchment area (AREA [$km^2$]); the average annual rainfall (SAAR, [mm]) for the period 1961-1990; the base flow index based on the Hydrology Of Soil Types classification presented in Boorman et al., 1995 (BFIHOST [-]), which reflects the geology of the site and has typical values that ranges from below 0.2 (highly impermeable) to above 0.8 (highly permeable); the mean distance between each pixel of the basin and the outlet (mean drainage path length, DPLBAR, [km]), and the extent of urban and suburban land cover during the year 2000 (URBEXT2000, [-]). Table 1 summarises these catchment properties in terms of the mean, minimum, maximum, and standard deviation value over the chosen set of 550 catchments. Of the 810 catchments for which peak flow data are available in Great Britain, 260 have been excluded for various reasons, including catchment size, and how well the gauged flows are thought to represent actual flows. Specifically, 225 catchments where DPLBAR<10 km and Area <50 km$^2$ have been excluded from the comparison of simulated and observed peak flows as modelled flows for these relatively small catchments were most likely to be adversely affected (underestimated) by the use of daily mean rainfall. These catchments have a faster hydrological response and probably the use of hourly rainfall data would be more appropriate to mimic the instantaneous peak flows. A modest number of catchments (35) were excluded due to strong anthropogenic influences including: i) the presence of an artificial channel that modifies the natural flow-paths; ii) unreliable rating curves due to the lack of high flow measures; and iii) strong influence of reservoirs or groundwater abstraction on the flow regime. Figure 1 presents a map of the study area, the location of the gauges selected for the analysis (black points), and the excluded gauges (white points).

# 4 Results and Discussion

## 4.1 Model Verification and Index Flood Map Estimation

A linear regression model was fitted to the measured and modelled log-transformed IF values for 550 catchments. The G2G model was executed for the whole simulation period (1960-2014) and the modelled IF in a given gauged station was computed using the modelled AMAX values corresponding to the period for which the measurements were available. Figure 2-a shows a scatterplot of 550 G2G and observation-derived IFs in logarithm scale, together with the derived linear regression model plot, and Table 2 shows the summary statistics of the linear regression model. The high values of the t-ratio, computed as the coefficient estimated value divided by its estimated standard deviation, give an indication that the estimated coefficients are statistically different from 0. The coefficient of determination $R^2$=0.91 summarizes the goodness of fit. Following Kjeldsen (2014), given the large number of catchments for which the model was evaluated (550) it is reasonable to assume that the prediction variance can be approximated by the variance of the regression model residuals, s=0.15. Under this assumption it is possible to evaluate the factorial standard error of the model FSE=1.47. The latter defines the 68% and 95% confidence intervals for the regression model as [$q \cdot FSE^{-1}$; $q \cdot FSE$] and [$q \cdot FSE^{-2}$; $q \cdot FSE^{2}$] respectively (Kjeldsen, 2014), where q indicates given discharge value. In our case q corresponds to the median of the AMAX. The FSE presented in this study is comparable with the FSE values of the regression models currently used in FEH which are based on the AMAX measurements of 600 gauging stations. The original FEH index flood regression model reported an FSE value of 1.56 (Robson and Reed, 1999) and the revised model lowered it to 1.431 (by assuming that the correlation between model errors is a function of the geographical distances between gauging stations (Kjeldsen et al., 2008)).

Figure 2-b presents a map of the residuals between modelled and measured IF using a logarithmic scale. The residuals are close to zero across most of Britain, with a modest underestimation in central and south west England, and a similarly modest overestimation in the South East. A significant factor contributing to the underestimation is the contribution of short-duration intense rainfall events to peak river flows in central and southern Britain, which will be poorly represented by daily gridded rainfall observations, while the overestimation in southern and eastern Britain can, for many groundwater-dominated catchments, be attributed to the effects of artificial abstractions which are not currently included in the G2G model formulation.

Figure 3-a presents a map of the modelled index flood ($m^3s^{-1}$) on a logarithmic scale, for the period 1960 to 2014. The IF is typically higher in the north and west of Britain, and in major rivers. The use of continuous G2G model simulation provides a consistent spatial and temporal dataset to explore whether there has been a significant change in the IF over the last 50 years. Figure 3-b presents a map of the change in the derived index floods between two periods: 1960 to 1986 and 1987 to 2014. The changes range from an increase in the IF of up to 45 $m^3s^{-1}$ (predominantly in the north and west) to a decrease of -40 $m^3s^{-1}$ in parts of Southeast Britain. This regional split is broadly in line with the increased trends detected in measured mean daily flows since the early 1960s in Scotland and, to a lesser extent, Wales and western England (Hannford and Marsh, 2008). However, the authors noted that the analysis of trends in some areas was limited by the available length of record.

The use of continous model simulation provides a method of estimating the IF with a 91% agreement with observation-derived estimates for 550 catchments across Britain. In order to investigate whether this agreement is influenced by catchment properties a series of analyses relating model fit with properties such as area, drainage path length, urban extent and baseflow index were undertaken. For each catchment property, the catchment values were divided into deciles (i.e. the nine values that divide the sorted data into ten equally sized subsamples) and measured and modelled IF for each catchment property subgroup were compared. Figure 4 presents 10 scatterplots and the coefficient of determination ($R^2$) of linear models fitted to the results for the catchment property: AREA. The title of each scatterplot specifies the AREA range [$km^2$] of each classes, for example the first plot is for catchments which range in area from 53 to 80 $km^2$, the second from 80 to 110 $km^2$, etc. Similar results are presented for percentage of urban extent (URBEXT2000) in Figure 5, and for baseflow index (BFIHOST), and drainage path length (DPLBAR) in Figure A1 and A2 in Appendix 1. The model fit is robust in the sense that is not strongly affected by the catchment properties. The decile range in $R^2$ is 0.82-0.90 for AREA, 0.78-0.92 for URBEXT2000, 0.81-0.93 for BFIHOST, and 0.84-0.91 for DPLBAR. These figures indicate relatively high levels of agreement between modelled and measured IF estimates, suggesting that the quality of the G2G estimated IF is relatively unaffected by different catchment properties and can provide estimates of consistent quality across various types of catchment (e.g. small, steep, or urbanized catchments).

## 4.2 Annual Maxima Trend Analysis

In the previous section we assessed whether AMAX output from a G2G continuous simulation could be used to estimate the measured IF by comparing the median of observed and simulated AMAXs over several decades. Typically, however, climate, anthropogenic or natural changes at the catchment scale can lead to long-term trends in observed annual maxima. For this reason it is important to ensure that if AMAX from continuous hydrological simulation are used in place of observed AMAX, they can also reproduce observed trends in river flows. This trend analysis has now been undertaken on 285 catchments, selected from the original 550, for which at least 40 year of measured flow data are available and the Mann Kendall test (MK, Kendall, 1975) with permutations provides a measure of the significance of potential trends in time. This method is presented in detail by Kundzewicz and Robson (2000) and has been used in several applications (i.e. Hannaford and Marsh, 2008, Hannaford and Marsh, 2006). The procedure is as follows: i) randomly re-order the AMAX time series to provide a large number of samples with no replacement; ii) perform Mann-Kendall trend test to each sample; iii) rank the trend test results; iv) compute the trend test for the original time series. If the derived trend for the original series falls outside the [0.05, 0.95] percentile range of the ranked values, it is deemed to be significant at the 95% confidence level, indicating a change in the magnitude of the AMAX over the 40-year period. The statistical tests have been performed on both measured and modelled AMAX providing test values (including the direction of the trend) and significance assessments for a trend in both the measured and modelled series. Results have been compared for the 69 catchments where the trend for the measured AMAX presented a significant test at the 95% significance level and are shown in Figure 6. No results are available for Scotland because the two criteria of at least 40 year of measured flow data are available and trend with a 95% significance level were not matched.

Figure 6 shows that: i) for 59 catchments positive trends were detected in both modelled and measured AMAX and ii) for 10 catchments the trend in the modelled series is not in agreement with the direction of the trend in the measured AMAXs series. These catchments are predominantly located in the south east part of England and for all of them the NRFA archive suggests that the runoff is affected by at least one of these reasons: a) reservoir in the catchment, b) presence of industrial or agricultural abstraction, and c) presence of water supply and groundwater abstractions. This anthropogenic influence which is not modelled in the current version of G2G may potentially explain the differences between measured and modelled AMAX trend in time for those basins.

## Conclusions

In this paper we demonstrate how use of continuous flow simulation by a national-scale distributed hydrological model (such as G2G) can be used to estimate key parameters such as the index flood (IF) required for flood estimation methods. The comparison between index floods estimated from current (FEH) and continuous simulation methods for 550 catchments throughout Great Britain indicates a good correlation between the two methods ($R^2$=0.91, factorial standard error FSE=1.47). We have also demonstrated that AMAX from continuous hydrological simulation can reproduce observed trends the measured annual maxima (agreement in 90% of the analysed catchments), indicating the potential utility of the methodology for conditions of non-stationarity.

This initial assessment of continuous simulation from a national-scale hydrological model (G2G) for estimating the IF is encouraging and demonstrates the new method can potentially overcome current methodological limitations such as the assumption of spatially homogeneous rainfall over the catchment and climate non-stationarity. Other benefits of the proposed new method include estimation of index floods in catchments subject to anthropogenic change, which at present can only be estimated using observed flows in naturalised catchments and require a correction to take into account the extent of urbanisation. Here, the accuracy of IF estimates from G2G continuous simulation is shown to be relatively unaffected by catchment properties such as area and urban extent, indicating that the methodology is robust for a variety of catchment types, so long as the continuous hydrological simulation is able to take into account the many factors (natural and anthropogenic) affecting river flows.

Countries such as Britain, for which an extensive network of flow and raingauges can support existing observation-based FEH methods, provide ideal test conditions for assessing the ability of alternative model-based flood estimation methods, such as continuous simulation from large-scale hydrological models, to underpin methods for flood estimation in data-sparse regions. It is to be hoped that the increasing availability and accuracy of global weather, climate and hydrological products can be used to develop a robust methodology to help engineers estimate design floods in regions with limited gauge data or affected by environmental change, potentially saving many lives.

## References


Bell, V. A., Kay, A. L., Jones, R. G., & Moore, R. J. (2007a). Development of a high resolution grid-based river flow model for use with regional climate model output. Hydrology and Earth System Sciences, 11(1), 532-549.

Bell, V. A., Kay, A. L., Jones, R. G., & Moore, R. J. (2007). Use of a grid-based hydrological model and regional climate model outputs to assess changing flood risk. *International Journal of Climatology*, *27*(12), 1657-1671.

Bell, V. A., Kay, A. L., Jones, R. G., Moore, R. J., & Reynard, N. S. (2009). Use of soil data in a grid-based hydrological model to estimate spatial variation in changing flood risk across the UK. Journal of Hydrology, 377(3), 335-350.

Bell, V.A.; Kay, A.L.; Davies, H.N.; Jones, R.G.. 2016 An assessment of the possible impacts of climate change on snow and peak river flows across Britain. Climatic Change, 136 (3). 539-553. 10.1007/s10584-016-1637-x

Blöschl, G. (Ed.). (2013). Runoff prediction in ungauged basins: synthesis across processes, places and scales. Cambridge University Press.

Bocchiola, D., De Michele, C., & Rosso, R. (2003). Review of recent advances in index flood estimation. Hydrology and Earth System Sciences Discussions,7(3), 283-296.

Boorman, D. B., Hollis, J. M., & Lilly, A. (1995). Hydrology of soil types: a hydrologically-based classification of the soils of United Kingdom. Institute of Hydrology.

Brath, A., Bacchi, B., and Rossi, R., (1992) The geo-morpho-climatic derivation of the probability distribution of the flood discharge. *Idrotecnica*, 4, 183-200.

Calver, A., Crooks, S., Jones, D.A., Kay, A., Kjeldsen, T. and Reynard, N., 2005. National river catchment flood frequency method using continuous simulation. Research Report to Defra, Centre for Ecology & Hydrology, Wallingford, UK.

Cameron, D., Beven, K., Tawn, J., & Naden, P. (2000). Flood frequency estimation by continuous simulation (with likelihood based uncertainty estimation). Hydrology and Earth System Sciences Discussions, 4(1), 23-34.

Calver, A., Crooks, S. C., Jones, D. A., Kay, A. L., Kjeldsen, T. C., & Reynard, N. S. (2005). National river catchment flood frequency method using continuous simulation.

Chow, V. T., Maidment, D. R., & Mays, L. W. (1988). Applied hydrology. McGraw Hill, New York, USA.

Cole, S. J., & Moore, R. J. (2009). Distributed hydrological modelling using weather radar in gauged and ungauged basins. Advances in Water Resources, 32(7), 1107-1120.

Córdova, J., & Rodríguez-Iturbe, I. (1983). Geomorphoclimatic estimation of extreme flow probabilities. Journal of Hydrology, 65(1), 159-173.

Dalrymple, T. (1960) Flood frequency analysis. Water Supply Paper 1543-A, US Geol. Survey, Reston, Virginia, USA

Dawson, C. W., Abrahart, R. J., Shamseldin, A. Y., & Wilby, R. L. (2006). Flood estimation at ungauged sites using artificial neural networks. Journal of Hydrology, 319(1), 391-409.

Dixon, H., Hannaford, J., & Fry, M. J. (2013). The effective management of national hydrometric data: experiences from the United Kingdom. Hydrological Sciences Journal, 58(7), 1383-1399.

EM-DAT (2015). EM-DAT: The OFDA/CRED international disaster database — www.emdat.be. Université Catholique de Louvain (Retrieved b http://www.emdat.be/ Access 10/04/16)

FEH (1999) Flood Estimation Handbook. Institute of Hydrology, Wallingford, UK.

Shu, C., & Burn, D. H. (2004). Artificial neural network ensembles and their application in pooled flood frequency analysis. Water Resources Research, 40(9).

Haddad, K., A. Rahman, and J. R. Stedinger (2012), Regional flood frequency analysis using Bayesian generalized least squares: A comparison between quantile and parameter regression techniques, Hydrol. Processes, 26, 1008–1021

Hall, M. J., Minns, A. W., & Ashrafuzzaman, A. K. M. (2002). The application of data mining techniques for the regionalisation of hydrological variables. Hydrology and Earth System Sciences Discussions, 6(4), 685-694.

Hannaford, J., & Marsh, T. (2006). An assessment of trends in UK runoff and low flows using a network of undisturbed catchments. International Journal of Climatology, 26(9), 1237-1253.

Hannaford, J., & Marsh, T. J. (2008). High-flow and flood trends in a network of undisturbed catchments in the UK. International Journal of Climatology, 28(10), 1325-1338.

Hirsch, R. M., Helsel, D. R., Cohn, T. A., Gilroy, E. J., & Maidment, D. R. (1992). Statistical analysis of hydrologic data. Handbook of hydrology., 17-1.

Hosking, J. R. M., & Wallis, J. R. (2005). Regional frequency analysis: an approach based on L-moments. Cambridge University Press.

Hosking, J. R. M., & Wallis, J. R. (1993). Some statistics useful in regional frequency analysis. Water Resources Research, 29(2), 271-281.

Hough, M. N., & Jones, R. J. A. (1997). The United Kingdom Meteorological Office rainfall and evaporation calculation system: MORECS version 2.0-an overview. Hydrology and Earth System Sciences Discussions, 1(2), 227-239.

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W & Fenicia, F. (2013). A decade of Predictions in Ungauged Basins (PUB)—a review. Hydrological sciences journal, 58(6), 1198-1255.

Institute of Hydrology. Flood Estimation Handbook (FEH), 5 Volumes. Wallingford, UK: Institute of Hydrology, 1999.

Jaafar, W., Zurina, W., & Han, D. (2012). Uncertainty in index flood modelling due to calibration data sizes. Hydrological Processes, 26(2), 189-201.

Kjeldsen, T. R., Stewart, E. J., Packman, J. C., Folwell, S. S., & Bayliss, A. C. (2005). Revitalisation of the FSR/FEH rainfall-runoff method. Final Report to DEFRA/EA, CEH, Wallingford.

Kjeldsen, T. R., & Jones, D. (2007). Estimation of an index flood using data transfer in the UK. Hydrological sciences journal, 52(1), 86-98.

Kjeldsen, T. R., Jones, D. A., & Bayliss, A. C. (2008). Improving the FEH statistical procedures for flood frequency estimation. Environment Agency.

Kjeldsen, T. R., & Jones, D. A. (2009). An exploratory analysis of error components in hydrological regression modeling. Water resources research, 45(2).

Kjeldsen, T. R. (2015). How reliable are design flood estimates in the UK?. Journal of Flood Risk Management, 8(3), 237-246.

Morris DG, Flavin RW (1990) A digital terrain model for hydrology. In: Proceedings of the 4th International Symposium on Spatial Data Handling, Zurich, Switzerland, 23–27 July 1990 1:250–262

Meigh, J. R., F. A. K. Farquharson, and J. V. Sutcliffe (1997), A worldwide comparison of regional flood estimation methods and climate, Hydrol. Sci. J., 42(2), 225–244.

Montaldo, N., Ravazzani, G., & Mancini, M. (2007). On the prediction of the Toce alpine basin floods with distributed hydrologic models. Hydrological processes, 21(5), 608-621.

Moretti, G. and Montanari, A., 2008. Inferring the flood frequency distribution for an ungauged basin using a spatially distributed rainfall–runoff model. Hydrology and Earth System Sciences, 12, 1141–1152. doi:10.5194/hess-12-1141-2008

National River Flow Archive (NRFA): Centre for Ecology and Hydrology, Wallingford, UK, available online: http://nrfa.ceh.ac.uk/, last acceess: 20 April 2016.

NERC (Natural Environment Research Council) (1975) Flood Studies Report. NERC, London, UK.

Perry, M., Hollis, D., & Elms, M. (2009). The generation of daily gridded datasets of temperature and rainfall for the UK. Met Office National Climate Information Centre, FitzRoy Road, Exeter, Devon EX1 3PB, UK.

Ravazzani, G., Bocchiola, D., Groppelli, B., Soncini, A., Rulli, M. C., Colombo, F., & Rosso, R. (2015). Continuous streamflow simulation for index flood estimation in an Alpine basin of northern Italy. Hydrological Sciences Journal, 60(6), 1013-1025.

Rabuffetti, D., Ravazzani, G., Corbari, C., & Mancini, M. (2008). Verification of operational Quantitative Discharge Forecast (QDF) for a regional warning system, the AMPHORE case studies in the upper Po River. Natural Hazards and Earth System Science, 8(1), 161-173.

Rigon, R., D'Odorico, P., and Bertoldi, G.: The geomorphic structure of the runoff peak, Hydrol. Earth Syst. Sci., 15, 1853-1863, doi:10.5194/hess-15-1853-2011, 2011.

Robson, A. J., & Reed, D. W. (1999). Flood estimation handbook. Institute of Hydrology, Wallingford.

Srinivas, V. V., Tripathi, S., Rao, A. R., & Govindaraju, R. S. (2008). Regional flood frequency analysis by combining self-organizing feature map and fuzzy clustering. Journal of Hydrology, 348(1), 148-166.

Stedinger, J. R., and G. D. Tasker (1985), Regional hydrological analysis: 1. Ordinary, weighted and generalized least squares compared, Water Resour. Res., 21, 1421–1432.

Viviroli, D., Zappa, M., Schwanbeck, J., Gurtz, J., & Weingartner, R. (2009). Continuous simulation for flood estimation in ungauged mesoscale catchments of Switzerland–Part I: Modelling framework and calibration results. Journal of Hydrology, 377(1), 191-207.

# Figures

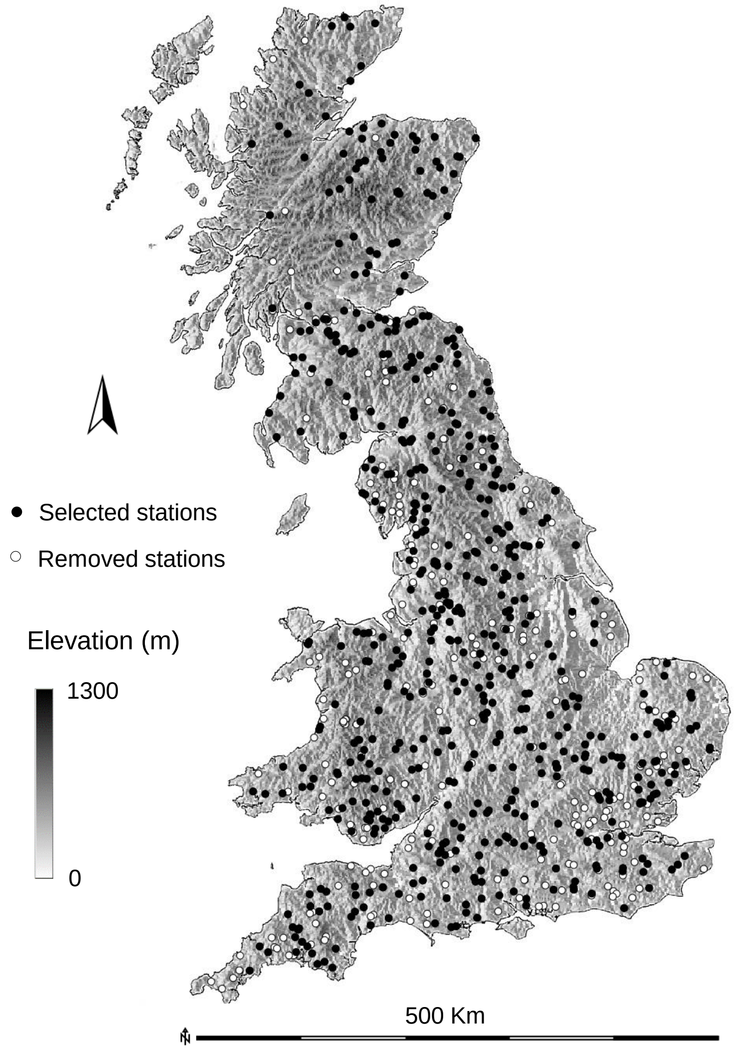


Figure 1: Location of the 550 catchments used in this study

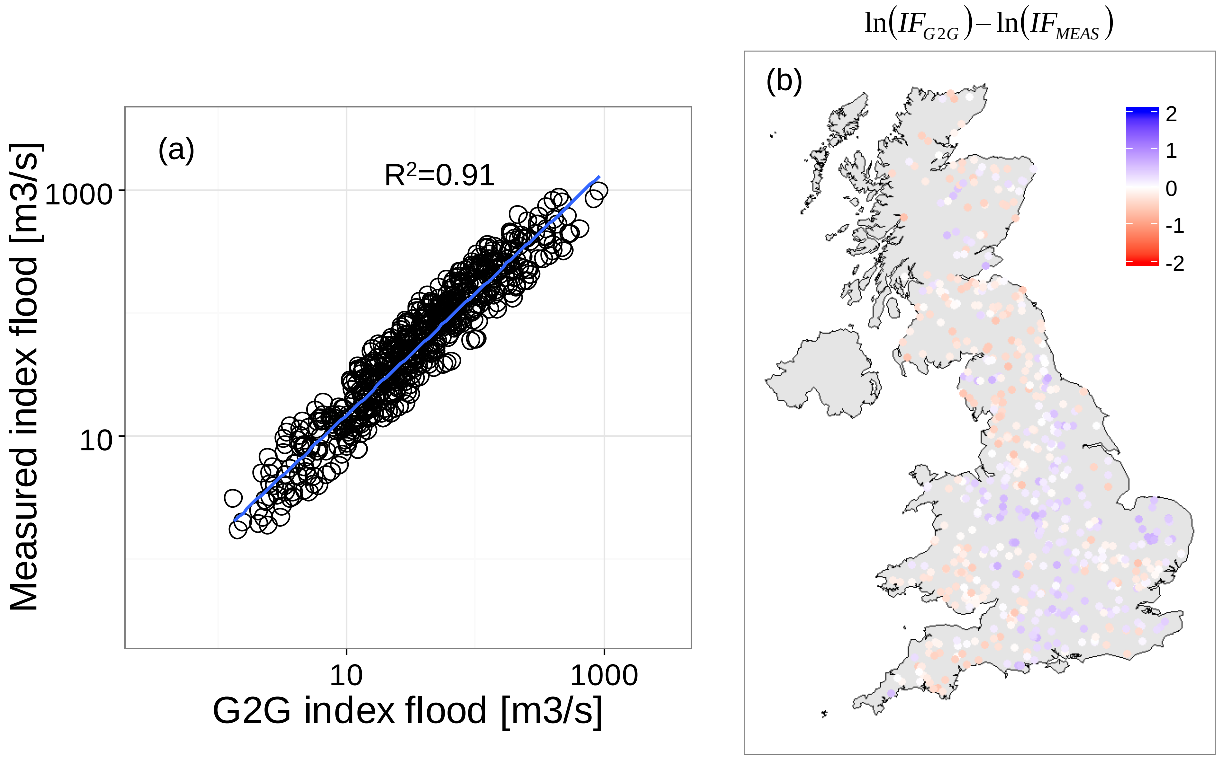


Figure 2: Linear regression model (a) and residual error in logarithm scale (b) for measured and modelled index floods for the 550 analysed catchments.

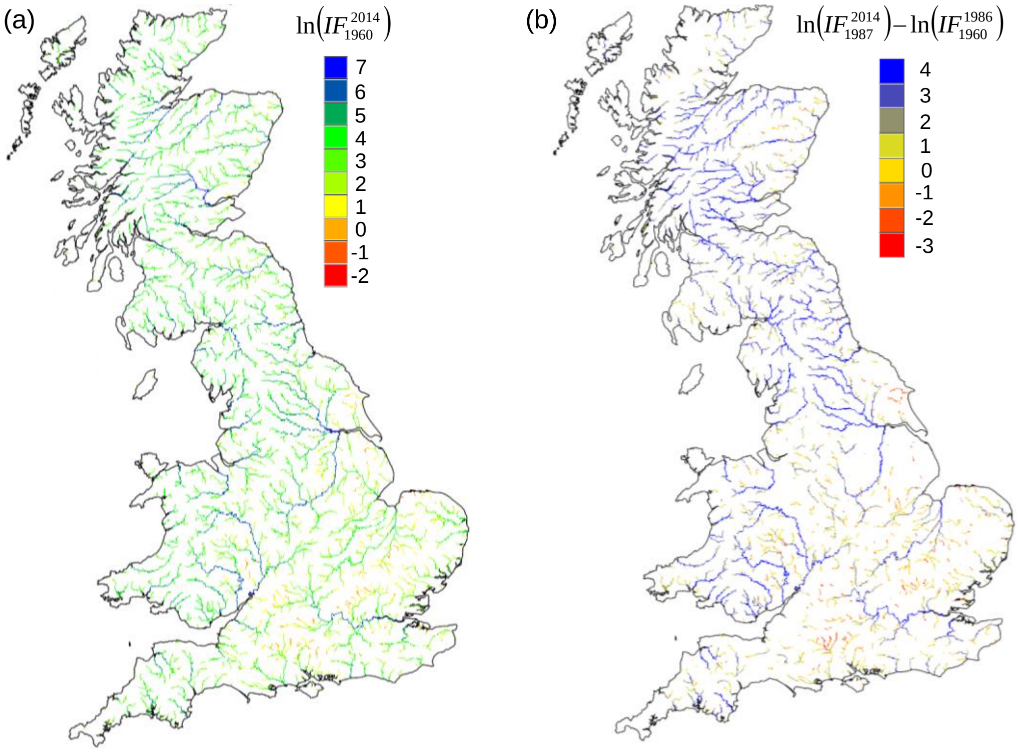


Figure 3: Maps of Britain showing, on a logarithmic scale: (a) Modelled index flood ($m^3s^{-1}$) for the period 1961-2011 (b) Change in the derived index flood ($m^3s^{-1}$) between 1961-1985 and 1986-2011.

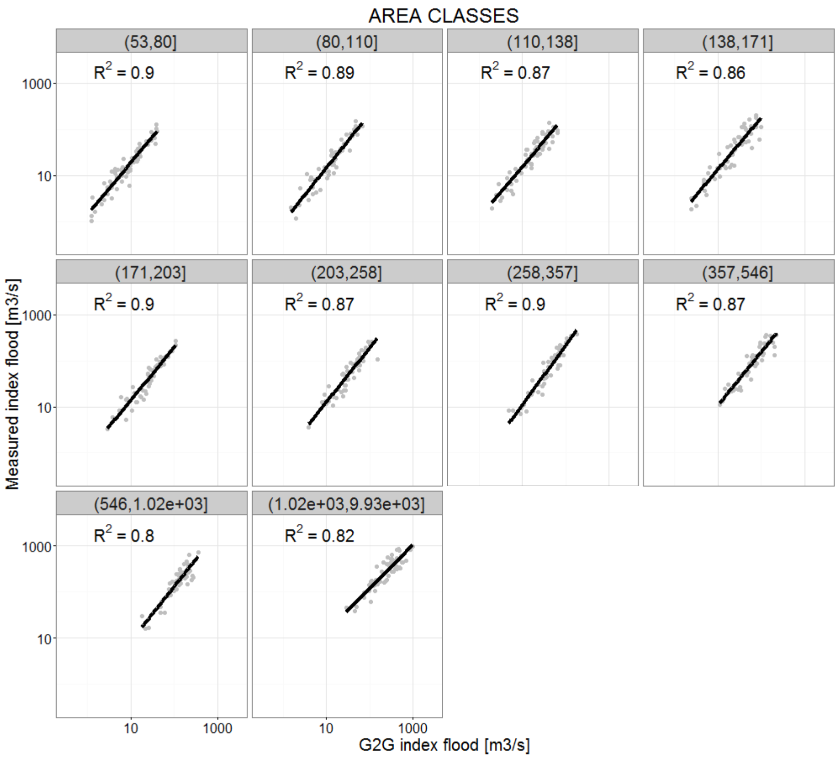


Figure 4: Scatterplots and coefficients of determination for modelled and measured index flood grouped by AREA classes.

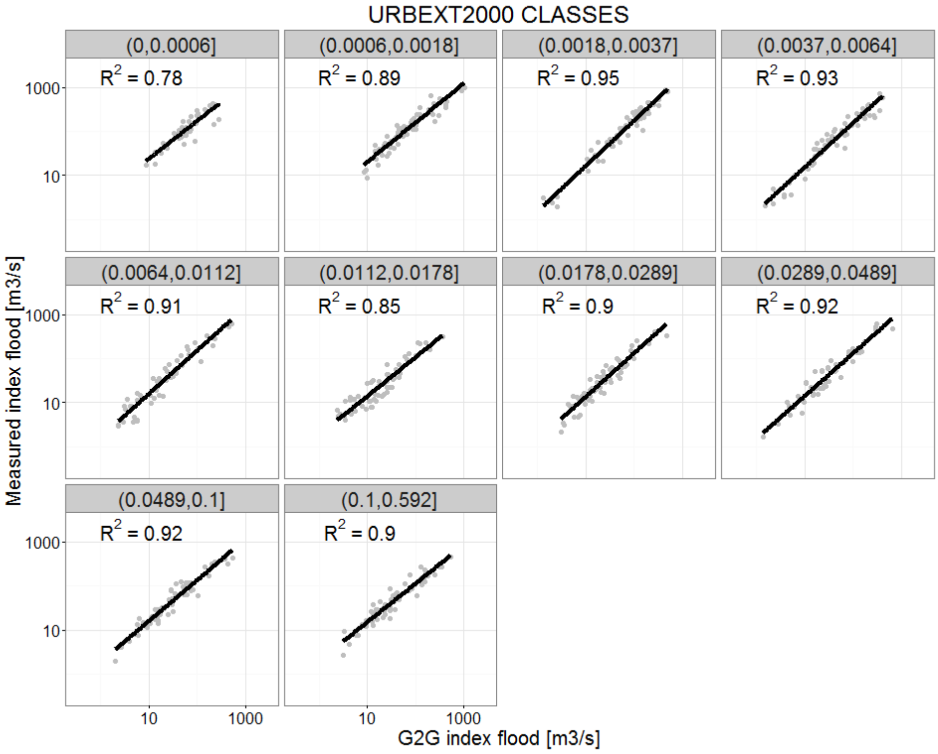


Figure 5: Scatterplots and determination coefficients for modelled and measured index flood grouped by URBEXT2000 classes.

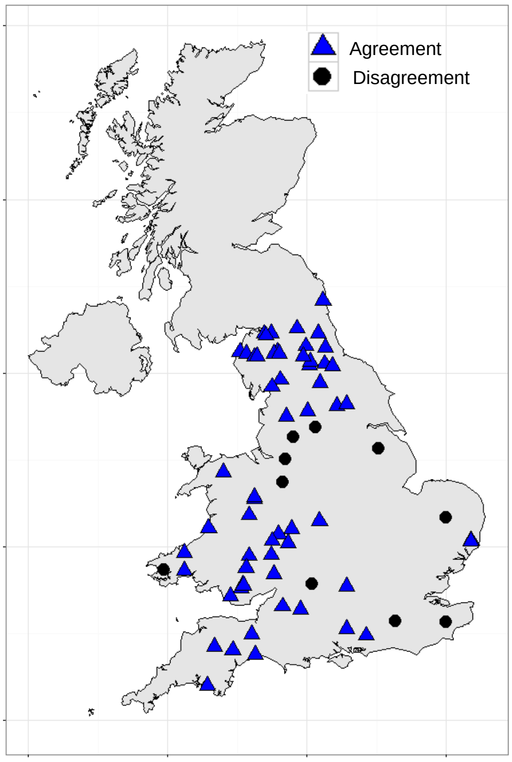


Figure 6: Comparison between the measured and modelled AMAX trend with time with a 95% significance. The catchments where both model and data agree are represented by blue triangles (positive); the points where they disagree are represented by black points.

# Tables

Table 1: Summary statistics (minimum, median, maximum, and standard deviation value) for the selected set of catchments indicators: AREA, SAAR, BFIHOST, DPLBAR, and URBEXT2000

| | AREA [km$^2$] | SAAR [mm] | BFIHOST [-] | DPLBAR [km] | URBEXT2000 [-] |
|---|---|---|---|---|---|
| Minimum | 55 | 558 | 0.24 | 10 | 0 |
| Median | 203 | 962 | 0.47 | 19 | 0.009 |
| Maximum | 9931 | 2913 | 0.96 | 140 | 0.592 |
| Stand. Dev. | 935 | 401 | 0.14 | 18 | 0.085 |

Table 2: Summary of the linear regression model linking the measured and modelled index floods

| Intercept | T-Stat intercept | Scaling exponent | T-Stat Scaling exponent | Residual Stand. Dev | $R^2$ |
|---|---|---|---|---|---|
| 0.41 | 8.995 | 0.99 | 76.681 | 0.386 | 0.910 |

# Appendix 1

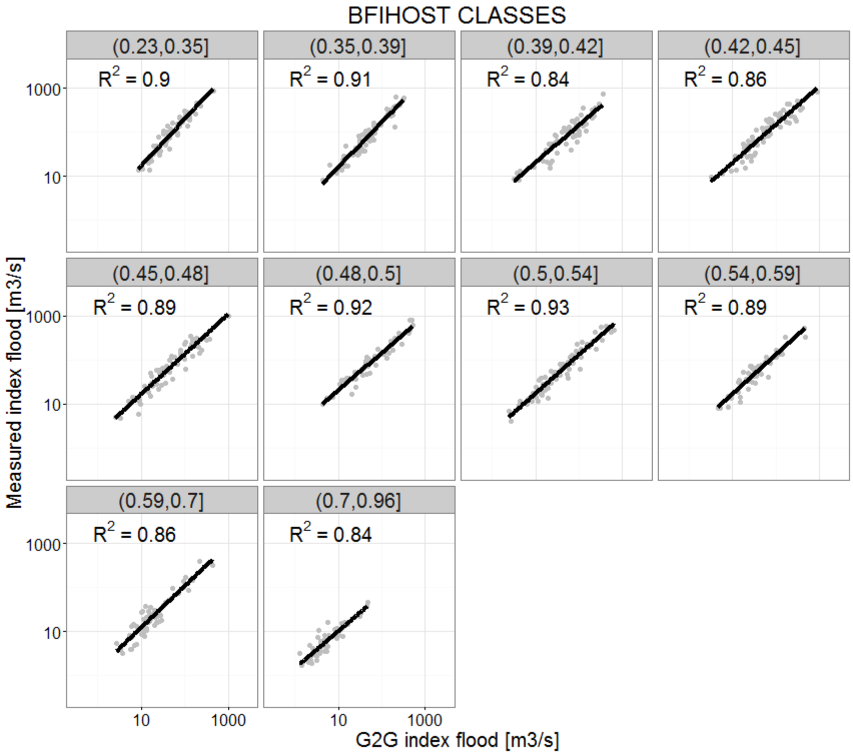


Figure A1: Scatterplots and determination coefficients for modelled and measured index flood grouped by BFIHOST classes.

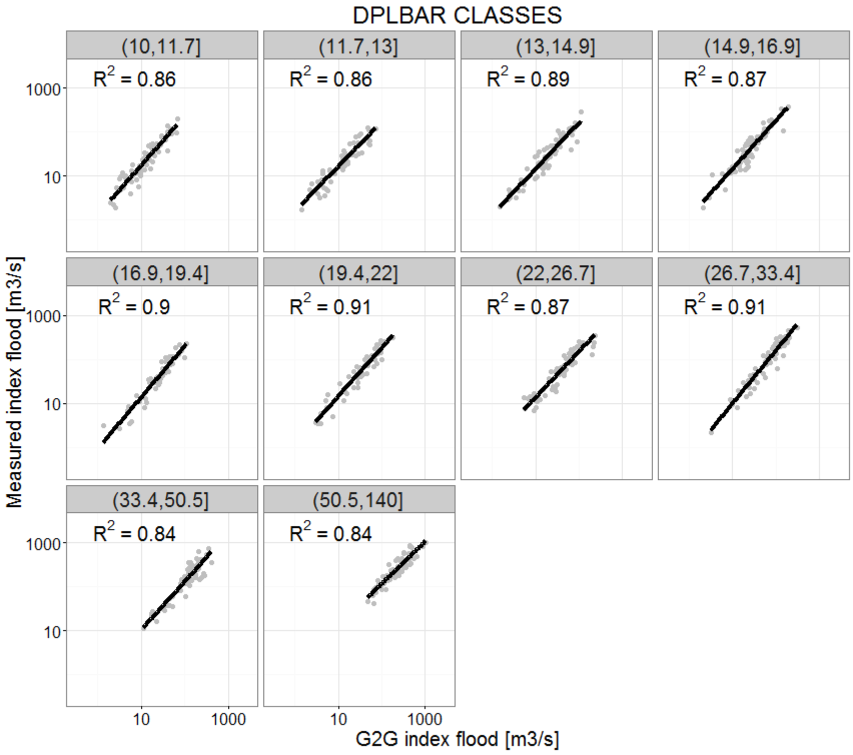


Figure A2: Scatterplots and determination coefficients for modelled and measured index flood grouped by DPLBAR classes.