Comparison of Threshold Hydrologic Response across Northern Catchments

1

2   **Comparison of Threshold Hydrologic Response across Northern Catchments**

3

4

5   Genevieve Ali[1*], Doerthe Tetzlaff[2], Jeffrey J. McDonnell[2,3], Chris Soulsby[2], Sean Carey[4], Hjalmar

6   Laudon[5], Kevin McGuire[6], Jim Buttle[7], Jan Seibert[8,9], Jamie Shanley[10]

7

8   [1]Department of Geological Sciences, University of Manitoba, Canada

9   [2]Northern Rivers Institute, School of Geosciences, University of Aberdeen, Scotland, UK

10   [3] Global Institute for Water Security, University of Saskatchewan, Canada

11   [4]School of Geography and Earth Sciences, McMaster University, Canada

12   [5]Department of Forest Ecology and Management, SLU, Umeå, Sweden

13   [6]Virginia Water Resources Research Center, Virginia Tech, USA

14   [7]Department of Geography, Trent University, Canada

15   [8]Department of Geography, University of Zurich, Switzerland

16   [9]Department of Earth Sciences, Uppsala University, Sweden

17   [10]U.S. Geological Survey, Montpelier, USA

18

19   [*]Corresponding author: Genevieve Ali, Genevieve.Ali@umanitoba.ca

20

21

22

23   **Abstract**

24

25   Nine mid- to high-latitude headwater catchments —part of the North-Watch (Northern

26   Watershed Ecosystem Response to Climate Change) program — were used to analyze threshold

27   response to rainfall and snowmelt-driven events, and link the different responses to the

28   catchment characteristics of the nine sites. The North-Watch data include daily time-series of

29   various lengths of multiple variables such as air temperature, precipitation and discharge.

30   Rainfall and meltwater inputs were differentiated using a degree-day snowmelt approach.

31   Distinct hydrological events were identified, and precipitation-runoff response curves were

32   visually assessed. Results showed that eight of nine catchments showed runoff initiation

33   thresholds and effective precipitation input thresholds. For rainfall-triggered events, catchment

34   hydroclimatic and physical characteristics (e.g., mean annual air temperature, median flow path

35   distance to the stream, median sub-catchment area) were strong predictors of threshold

36   strength. For snowmelt-driven events, however, thresholds and their governing factors

37   controlling precipitation-runoff response were difficult to identify. The variability in catchments

38   responses to snowmelt was not fully explained by runoff initiation thresholds and input

39   magnitude thresholds. The quantification of input intensity thresholds (e.g., snow melting and

40   permafrost thawing rates) is likely required for an adequate characterization of nonlinear spring

41   runoff generation in such northern environments.

42   **Keywords:** thresholds, rainfall, snowmelt, quickflow, dynamic storage deficit, North-Watch

43   program

44 **1   Introduction**

45

46   Understanding of the response of streams to precipitation inputs in northern headwater

47 catchments is still limited (Tetzlaff *et al*., 2013a). In many temperate humid catchments where

48 most of the process work has been done, hydrological threshold behaviours have been

49 described where changes in runoff response are strongly dependent on antecedent soil

50 moisture conditions and/or disproportional to forcing inputs across the whole possible range of

51 inputs (e.g., Dickinson and Whiteley, 1970; Tani, 1997; Phillips, 2003; Tromp-Van Meerveld and

52 McDonnell, 2006a; Detty and McGuire, 2010). Many studies have shown that critical values of

53 precipitation amounts or (soil moisture) storage capacities need to be exceeded for hydrological

54 response initiation (e.g., Whipkey, 1965; Mosley, 1979; Tani, 1997; Uchida *et al*., 2005; Tromp-

55 Van Meerveld and McDonnell, 2006a); these precipitation input thresholds have been

56 considered by some to be emergent catchment properties (Weiler *et al.*, 2005; Lehmann *et al.*,

57 2007) and by others as catchment hydrological signatures (Spence, 2007). Thresholds may also

58 provide a useful tool for catchment comparison and model calibration and validation, as they

59 facilitate the grouping of similar hydrological responses (e.g., Sivakumar, 2005; Graham and

60 McDonnell, 2010).

61

62   While threshold detection and explanation appears to be a useful research avenue for

63 advancing catchment process understanding, work to date has focused mostly on small

64 catchments and hillslopes (Tani, 1997), has been highly qualitative in the quantification of

65 threshold strength (Tromp-Van Meerveld and McDonnell, 2006a) and has not yet explored

66   these dynamics in northern watersheds, as experimental work in high-latitude environments is

67   much harder to conduct. Indeed, with only a few exceptions (e.g., Detty and McGuire, 2010;

68   Graham and McDonnell, 2010; Penna *et al.*, 2011), threshold detection studies have been

69   performed largely at the hillslope scale given the availability of high-frequency (e.g., hourly and

70   sub-hourly) precipitation-runoff data. The majority of threshold detection studies have dealt

71   with rainfall events (and not snowmelt) in mostly humid temperate environments (e.g., Tani,

72   1997; McGlynn and McDonnell, 2003; Tromp-Van Meerveld and McDonnell, 2006a, b; Lehmann

73   *et al.*, 2007; Detty and McGuire, 2010). Event rainfall critical threshold values have been

74   identified for specific sites, e.g., 20 mm (Mosley, 1979; Tani, 1997), 23 mm (Penna *et al.*, 2011),

75   35 mm (Whipkey, 1965) or 55 mm (Tromp-Van Meerveld and McDonnell, 2006a): it is likely that

76   these differences in rainfall storage thresholds is controlled by catchment characteristics such as

77   mean soil depth, depth of overburden, or interception capacity of the overlying vegetation and

78   litter layer, although those aspects are rarely reported in detail in associated publications.

79

80   In contrast to temperate environments, little information exists for northern

81   catchments in terms of the linearity or non-linearity of their runoff responses to precipitation

82   inputs. Several research initiatives such as the Northern Research Basins (NRB) working group

83   have been established to gain a better understanding of runoff generation processes in cold

84   regions, particularly processes that are heavily influenced by snow, ice and frozen ground (Kane

85   and Yang, 2004). Given the limited amount of hydrometric equipment deployed in northern

86   catchments in comparison to temperate environments, hydrologists tended to transfer theories

87   developed in temperate regions to cold landscapes to explain the spatio-temporal variability of

88   runoff volume and magnitude, regardless of whether runoff generation is rainfall or snowmelt-

89   driven (Quinton and Marsh, 1999). The transferability of traditional runoff generation theories

90   to cold catchments is not straightforward given: (i) the major differences in the control factors

91   prevailing in low and high-latitude regions; and (ii) the tremendous heterogeneity of landscapes

92   and dominant processes even within high-latitude regions. The complexity of threshold

93   response in northern Canadian catchments has been documented; notably by Allan and Roulet

94   (1994), Goodyear (1997), Spence and Woo (2002, 2003, and 2006) and Buttle *et al*. (2004),

95   among others. The (ubiquitous) existence of runoff initiation thresholds and effective

96   precipitation thresholds in northern catchments, however, remains unclear as water storage

97   and release are not only governed by antecedent soil moisture but also snowpack and

98   permafrost properties. Site intercomparison work is needed to quantify how hillslope or

99   catchment characteristics might explain differences in threshold values, if they do indeed exist

100   in northern catchments, and how rainfall and snowmelt-driven hydrological dynamics might

101   compare in cold landscapes. This is especially important in light of projected climate changes

102   that predict spatially variable effects on northern streamflow regimes depending on the future

103   magnitude and onset of snowmelt runoff generation (Tetzlaff *et al*., 2013a).

104

105   Here we explore the linearity of runoff response to precipitation inputs for nine mid- to

106   high-latitude catchments from the North-Watch (Northern Watershed Ecosystem Response to

107   Climate Change; http://abdn.ac.uk/northwatch) program. North-Watch is a cross-regional inter-

108   catchment comparison initiative that aims to assess the physical, chemical and ecological

109   response of northern catchments to climate change. Extensive temperature, precipitation and

110   discharge data available at the daily timestep for each study catchment were processed using a

111   degree-day methodology to differentiate rainfall from snowmelt water inputs. Hydrograph

112   analysis was then used to identify distinct hydrological events and examine water input,

113   dynamic water storage, and runoff dynamics. Three specific questions guided the analyses: (i)

114   Do northern catchments exhibit threshold response to precipitation inputs? (ii) If so, is there a

115   (significant) difference in threshold behaviours between rainfall-triggered and snowmelt-driven

116   hydrological events? and (iii) Which catchment characteristics best explain differences in

117   threshold values among the sites? The overall goal was to understand how hydrological event

118   type (rainfall-triggered versus snowmelt-driven) and input or water storage dynamics interplay

119   to determine catchment runoff response patterns in mid- to high-latitude environments.

120

121

122   **2   Methods**

123

124   ***2.1   Study sites***

125

126   The nine study sites are part of the North-Watch program and were chosen as both long-

127   term hydroclimatic and detailed topographic data were available. The catchments are located

128   within Scotland, the United States, Canada and Sweden (Figure 1) and are among the most

129   intensively studied long-term headwater research sites across the circum-boreal region. They

130   span different hydroclimatic zones, including northern temperate, subarctic and boreal

131   environments; mean annual air temperatures range from -2.2^{o}C to 9.2^{o}C across the sites while

132   mean annual precipitation ranges from 478 mm to 2632 mm. Some of their other characteristics

133   have been discussed in detail by Carey *et al*. (2010, 2013), Tetzlaff *et al*. (2013b), and Laudon *et*

134   *al*. (2013b) and are summarized in Table 1.

135

136   Briefly, in Scotland, the Strontian site is situated in the maritime northwest, the Allt

137   a'Mharcaidh site is in the western subarctic Cairngorms and the Girnock site is in the Northeast.

138   The three catchments have drainage areas ranging from 8 to 30 km^2 and include steep montane

139   regions and flat, lower-lying areas. Mean annual temperatures range from 5.7^{o}C to 9.1^{o}C and

140   geology consists largely of igneous and metamorphic rocks (Robins, 1990). Typically, superficial

141   glacial drift is superimposed on the solid geology and determine the presence of fine textured

142   peats and peaty gleys in valley bottoms and on gentle slopes; freely draining soils such as

143   Podzols or alluvial soils are present on steeper slopes (Tetzlaff *et al.*, 2007). The Strontian

144   catchment is partly forested (mainly *Pinus sylvestris*), especially on lower slopes while the Allt

145   a'Mharcaidh and Girnock sites are characterized by heather (*Calluna* spp.) on steeper slopes at

146   higher altitudes and blanket bog (*Spagnum* spp.) in poorly drained areas (Bayfield and Nolan,

147   1998).

148

149   Two of the US sites are located in the Northeast (Hubbard Brook and Sleepers River)

150   while a third is in the Northwest (H.J. Andrews). In the White Mountains of New Hampshire,

151   Hubbard Brook Experimental Forest (WS3, 0.41 km^2) is covered by second-growth northern

152   hardwood species. Short, cool summers and long, cold winters are common in this humid

153   continental climate (Likens and Bormann, 1990; Bailey *et al.*, 2003) with a mean annual air

154   temperature of 6.4^{o}C and 1381 mm of precipitation, 25% to 35% of which falls as snow. Geology

155   largely consists of pelitic schist overlain by basal and ablation tills of varying thickness. Sleepers

156   River (W9, 0.41 km^2) in Vermont is also primarily forested with northern hardwoods of sugar

157   maple, ash, beech, and yellow birch (Shanley and Chalmers, 1999). The catchment has a mean

158   annual air temperature of 4.7^{o}C and receives 1256 mm of precipitation annually, 25% of which

159   typically falls as snow. Bedrock is mostly quartz-mica phyllite with calcareous granulite overlain

160   by dense silty till. In Oregon, the catchment under study is the 5.8 km² Mack Creek in the H.J.

161   Andrews Experimental Forest. Its geology is andesitic and basaltic lavaflows and it is mostly

162   covered by old-growth Douglas fir (*Pseudotsuga menziesii*) forest. Mack Creek is not only the

163   steepest catchment (with the highest relief of 860 m) among all North-Watch study sites but

164   also the warmest and the wettest. Winters are usually wet and mild and summers rather warm

165   and dry (Anderson, 1992) as the catchment has a mean annual air temperature of 9.2^{o}C and

166   mean annual precipitation of 2158 mm. Greater than 80% of precipitation occurs from

167   November to April, most of which falls as snow.

168

169   In Canada, focus was on the Wolf Creek catchment (Granger basin, 7.6 km^2) and one of

170   the Dorset catchments (Harp 5, 1.19 km^2). Wolf Creek is the second most northerly catchment

171   and is the coldest and driest (mean annual air temperature of -2.2^{o}C) of all North-Watch sites as

172   it is subjected to a sub-arctic continental climate on the fringe of the Coast Mountains of Yukon.

173   Permafrost underlies 70% of the catchment while the geology is primarily sedimentary,

174   comprised of limestone, sandstone, siltstone and conglomerate, overlain by a mantle of glacial

175   till ranging from 1-4 m in thickness (Carey and Quinton, 2005). Given the cold temperatures and

176 low annual precipitation (478 mm), vegetation generally consists of shrubs (*Salix*) and alpine

177 tundra at higher elevations (McCartney *et al.*, 2006). The Dorset site in Ontario is located in the

178 southern Boreal ecozone (Eimers *et al.*, 2008) in a humid continental climate with a mean

179 annual temperature of 4.9°C and precipitation of 980 mm. In contrast to Wolf Creek, soil frost is

180 rare as it primarily occurs in wetlands and only in winter, and the bedrock is a Precambrian

181 shield overlain by a thin layer of till. Vegetation is deciduous or mixed forest on well-drained

182 soils whilst poorly drained soils have mixed or coniferous forest.

183

184   Lastly, the Krycklan catchment (site 7, 0.50 km^2) in Sweden on the Fennoscandian shield

185 has a mean annual temperature of 2.4°C and is the second driest of all North-Watch sites with a

186 mean annual precipitation value of 651 mm, 40% of which falls as snow. It is underlain by

187 metasediments and podzol soils. The spatial distribution of vegetation species is highly

188 dependent on topography: dry upslope areas are primarily forested with mature Scots Pine

189 (*Pinus sylvestris*), wetlands are usually covered with *Sphagnum*, and other flat, low-lying areas

190 are covered with Norway Spruce (*Picea abies*) (Laudon *et al.*, 2013).

191

192 ***2.2   Hydrograph analysis and input-output response assessment***

193

194   Multi-year precipitation, temperature and discharge data were analysed to identify

195 distinct hydrological events and relate water inputs, dynamic storage deficits (i.e., overall

196 catchment shallow soil storage deficit – see details below), and runoff initiation prior to the

197 analysis of hydrologic thresholds. For each catchment, the longest continuous measurement

198 period available of the daily precipitation, air temperature and discharge timeseries was used

199 (Table 2). From the discharge timeseries, computer-based baseflow separations were

200 performed. Three different baseflow estimation methods were used: the fixed interval, the

201 sliding interval and the local minimum methods (Sloto and Crouse, 1996). As the differences

202 between the three methods were rather small, the fixed interval baseflow estimates were

203 retained for further analyses. Using the precipitation timeseries in conjunction with daily air

204 temperature data, water inputs were separated into two categories: rainfall and snowmelt.

205 Rainfall was assumed to be all precipitation falling when air temperature was above 0°C. Snow

206 accumulation was modelled by adding all precipitation when air temperature was below 0°C.

207 The snowpack was assumed to melt with a degree-day factor of 4 mm °C^{-1} day^{-1} when air

208 temperature was above 0°C (Juston *et al.*, 2009). The uniform threshold temperature and

209 degree-day factor across all sites were used for simplicity and because of the lack of consistent

210 energy balance data with which to estimate snowmelt.

211

212   To delineate hydrological events for all North-Watch sites, the following rules were

213 applied:

214   (i) A hydrological event is defined as the occurrence of a water input event followed by a

215 runoff event;

216   (ii) The beginning of a water input event is defined by a day with nonzero water input

217 (i.e., water input ≥ 1 mm) after a minimal 1-day dry period;

218   (iii) The beginning of a hydrological event corresponds to the beginning of a water input

219 event;

220   (iv) The beginning of the associated runoff event is defined by the first initial hydrograph

221   rise after the beginning of the water input event;

222   (v) The end of a water input event is defined by a day with precipitation input is less than

223   1 mm;

224   (vi) The end of a runoff event is defined by a day at the end of a recession period with no

225   water input or less than a 15% difference between the daily baseflow and the daily discharge

226   values;

227   (vii) The end of a hydrological event corresponds to the end of a runoff event.

228   Given the use of rainfall and snowmelt water inputs, two types of hydrological events could be

229   discriminated: rainfall-triggered events (i.e., rain > 0, snowmelt = 0), and snowmelt-driven

230   events (i.e., snowmelt > 0, with occasionally rain > 0 as well). Beyond the rainfall vs. snowmelt

231   event classification, no discrimination was made between rain-on-snow events and radiation-

232   driven melt events. Across all datasets, the identified water input events always led to a

233   discharge increase, albeit sometimes very small. Some hydrological events were associated with

234   a runoff coefficient (ratio of total runoff to total water input) > 100%: these events were

235   retained for further analyses only if they involved nonzero snowmelt water inputs to justify such

236   high runoff coefficient values.

237

238   Once all hydrological events were identified, the following state variables were

239   calculated:

240   - W_{input} is the sum of all water inputs (rainfall and snowmelt) for the duration of an event;

241 - $W_{storage}$ is the amount of water input required before runoff starts. Building upon the Soil

242   Conservation System (SCS) basic rainfall-runoff equation, here $W_{storage}$ is computed as

243   the initial abstraction: the sum of all water inputs (rainfall and snowmelt) which occur

244   between the beginning of the water input event and the initial rise in the storm

245   hydrograph (e.g., Steenhuis *et al.*, 1995; Lyon *et al.*, 2004). The initial abstraction can

246   therefore be seen as dynamic storage and be used as a proxy measure for the overall

247   catchment shallow soil storage deficit prior to each event;

248 - Quickflow is the difference between the discharge and the baseflow timeseries, and Q_{flow}

249   is the sum of all quickflow produced between the beginning and the end of a runoff

250   event.

251 To estimate $W_{storage}$ (i.e., the initial abstraction) from the event hyetographs and hydrographs,

252 only hydrological events with a minimum 1-day delay between the start of the water input

253 event and the initiation of runoff response were considered; doing so made it possible to avoid

254 dealing with high frequency (hourly), short-term input-output dynamics which are not well

255 captured by daily data. For each catchment, Q_{flow} was plotted against both W_{input} and then

256 against $W_{input} - W_{storage}$, in both cases separately for rainfall-triggered and snowmelt-driven

257 events (i.e., four plots in total). The variable $W_{input} - W_{storage}$ was used as it represents the

258 effective precipitation after the overall catchment storage deficit has been overcome. Given the

259 size of the catchments considered, the routing of effective precipitation to the catchment outlet

260 was assumed to occur rather quickly (within hours) and therefore considered instantaneous for

261 the selected data resolution (i.e., daily time scale).

262

263   Several recent studies have shown examples of nonlinear hydrological response with

264   relationships that are reminiscent of a hockey stick shape (e.g., Tani, 1997; Weiler *et al.*, 2005;

265   Detty and McGuire, 2010; Graham and McDonnell, 2010). Given the presence of a critical value

266   (i.e. a threshold) of water inputs, zero or low runoff is observed below the critical value whereas

267   a strong linear correlation exists between the runoff response and the water inputs above the

268   threshold. The presence (or absence) of thresholds in the hydrological response of the North-

269   Watch catchments was visually assessed in two ways: the relationship between Q_{flow} and W_{input}

270   was used to detect runoff initiation thresholds, while the relationship between Q_{flow} and $W_{input} -$

271     $W_{storage}$ was used to detect effective input thresholds. Both types of thresholds were identified

272   based on the clearest slope change, or break in slope in input-output scatter plots. Three

273   metrics were then used to characterize catchment hydrological behaviour at each site:

274   -   The Spearman rank correlation coefficient $r_{Spearman}$ between the output variable (Q_{flow})

275   and the input variable (W_{input} or $W_{input} - W_{storage}$); it was computed to measure the

276   strength of the relationship between water inputs, dynamic storage deficits and runoff

277   response at the catchment outlet, and its statistical significance ($p < 0.05$) was assessed;

278   -   The threshold value, when it was identifiable from the input-output scatter plots; and

279   -   The coefficient of determination R^2 between input and output values above the

280   threshold value (when applicable).

281   The Spearman rank correlation coefficient $r_{Spearman}$ was determined for all data, whereas R^2

282   values were only computed for data subsets above the threshold. Since the hockey stick

283   conceptualization assumes a strong linear correlation between the runoff response and the

284   water inputs above the threshold, the R^2 is the Pearson correlation coefficient to the power of

285   2; in our study, it was strongly correlated to the slope of the best-fit regression line and a good

286   indication of the catchment efficiency to produce runoff.

287

288   ***2.3   Catchment controls***

289

290   Spearman rank correlation coefficients were also calculated between the three metrics

291   of catchment hydrological behaviour and a range of hydroclimatic and topographic catchment

292   properties (Table 3). This was done to investigate which catchment characteristics might explain

293   any differences in hydrological behaviour among the sites (research question (iii)). These

294   correlations between catchment characteristics and the three metrics of hydrological behaviour

295   are hereafter referred to as $corr_{catchment}$ to distinguish them from the Spearman rank correlation

296   coefficient ($r_{Spearman}$) that is used to measure the strength of the input-output relationships. The

297   topographic properties described in Table 3 were derived from Digital Terrain Models (DTMs)

298   with a pixel resolution of 10 m available for all nine sites. Briefly, each catchment's relief was

299   computed as the difference between the minimum and maximum elevation scaled by the

300   squared root of the catchment area. The terrain slope was estimated using both the D8 (Quinn

301   *et al.*, 1991) and the MD∞ (Seibert and McGlynn, 2007) flow direction algorithms. After surface

302   topography-driven flow paths were determined for each DTM pixel based on the direction of

303   steepest descent, four indices were derived: the elevation above the stream, the distance from

304   the stream, the average gradient along the flow path to the stream, and the ratio of the flow

305   path length to the flow path gradient which was used as a proxy for travel times (Gardner and

306   McGlynn, 2009). The downslope index (Hjerdt *et al.*, 2004) is defined as the gradient towards

307   the closest point at least 5 m (in altitude) below a certain point while the upslope area draining

308   through each pixel was calculated using the MD∞ algorithm (Seibert and McGlynn, 2007). Two

309   variants of the topographic wetness index (Beven and Kirkby, 1979) were considered both using

310   the upslope area per unit contour length divided (i) by the local slope in one case, and (ii) by the

311   downslope index in the other case. All DTM-based indices were aggregated into one value for

312   each catchment using the catchment-wide median value. Lastly, the median sub-catchment

313   area was computed as an indicator of catchment drainage structure (McGlynn and Seibert,

314   2003). For all catchments, the stream network was defined using a 5 ha accumulated area

315   threshold for stream initiation. The median of the local catchment areas of all stream pixels

316   upstream of the catchment outlet was then estimated (McGlynn and Seibert, 2003; McGlynn *et*

317   *al.*, 2003).

318

319

320   **3   Results**

321

322   ***3.1   Visual identification of thresholds***

323   Scatter plots of Q_{flow} versus W_{input} for rainfall-triggered and snowmelt-driven events are

324   presented in Figures 2 and 3, respectively, while scatter plots of Q_{flow} versus $W_{input} - W_{storage}$ are

325   presented in Figures 4 and 5. By working with two series of plots, the inclusion of a storage

326   proxy variable in the scatter plots in Figures 4 and 5 was evaluated with regard to its ability to

327   improve the strength of the relationships between input and output hydrological variables. For

328   rainfall-triggered events, inter-catchment differences could be observed in the relationship

329   between Q_{flow} and W_{input}; indeed, a linear plot was obtained for the Strontian site while a clearer

330   nonlinear curve was associated with the Sleepers River site and a large scatter was encountered

331   for the Dorset site (e.g., Figure 2). Some catchments also exhibited differences in hydrological

332   response between input types, as was the case for the Wolf Creek site where the overall scatter

333   pattern associated with rainfall-triggered events was significantly different from that associated

334   with snowmelt-driven events. Conversely, for the Hubbard Brook site there was no significant

335   difference between the scatter pattern associated with rainfall-triggered events and the pattern

336   associated with snowmelt-driven events. Regardless of the input type considered, higher

337   Spearman rank correlation coefficients ($r_{Spearman}$) were found when $W_{input} - W_{storage}$ (Figures 4-5)

338   rather than W_{input} (Figures 2-3) was the dependent variable; this reflects a slightly better

339   characterization of hydrological response in all nine North-Watch catchments when a proxy of

340   dynamic storage was used. For the Strontian catchment, for example, $r_{Spearman}$ values were 0.91

341   and 0.88 with and without consideration of the storage component, respectively (Figures 2 and

342   4). Nonlinear input-output relationships somehow reminiscent of the hockey stick shape were

343   dominant for all catchments, and the linear relationship observed for Strontian could be

344   equated to a hockey stick shape with a very small (near zero) threshold (Figure 4). At the end of

345   the visual assessment procedure, some patterns of hydrological response were characterized as

346   unclear (Table 4) and were often associated with $r_{Spearman}$ values below 0.6 (e.g., Mharcaidh and

347   Wolf Creek sites, Figure 2).

348

349   Table 4 summarizes the threshold values identified from a visual assessment of the

350   scatter plots; that visual assessment was highly subjective given absent, multiple or very subtle

351   breakpoints in most plots. Nevertheless, the identified threshold values were highly variable

352   between catchments: for rainfall events, they ranged from 50 to 100 mm (median value:

353   80 mm) when the storage deficit was not taken into account and from 36 to 80 mm (median

354   value: 55 mm) when the storage deficit was considered. Threshold values for snowmelt events

355   were noticeably higher: they ranged from 25 to 180 mm, with a median value of 120 mm when

356   the storage deficit was not considered and 85 mm when the storage deficit was considered

357   (Table 4).

358

359   ***3.2   Differences between rainfall-triggered and snowmelt-driven events***

360

361   The three metrics used to characterize catchment hydrological behaviour at each site

362   (i.e., $r_{Spearman}$, the threshold value, and the R^2 value above the threshold) are reported in Figures

363   2-4 and Tables 4-5. The threshold characterization of hydrological response was weaker for

364   snowmelt-driven events than it was for rainfall-triggered events. Indeed, for snowmelt-driven

365   events, the mean Spearman rank correlation coefficient among all sites was 0.59 without and

366   0.69 with consideration of the storage deficit (Figure 3 and Figure 5, respectively). In contrast,

367   the mean $r_{Spearman}$ for rainfall-triggered events among all sites was 0.70 without and increased to

368   0.78 with the storage deficit taken into account (Figure 2 and Figure 4, respectively). Apart from

369   the Girnock and the Mharcaidh catchments, for which snowmelt-driven threshold values were

370   systematically smaller than their rainfall-driven counterparts, snowmelt-driven thresholds were

371   usually larger than rainfall-driven ones by a factor of 1.3 to 3.4 (Table 4).

372

373   For rainfall-triggered events with or without consideration of the storage deficit, the

374   highest R^2 values above the threshold were found for Sleepers River, H.J. Andrews, Hubbard

375   brook, and Wolf Creek (Table 5). For snowmelt-driven events, however, R^2 values above the

376   threshold were generally low or not computed due to a non-identifiable threshold. One

377   exception was the Sleepers River site for which the R^2 above the threshold exceeded 0.6 for

378   both rainfall and snowmelt events, regardless of whether the storage deficit was considered or

379   not. All other catchments were associated with R^2 above the threshold of less than 0.5 for

380   snowmelt events (Table 5).

381

382   **_3.3_**   **_Catchment controls on hydrological behaviour_**

383

384   For each type of event, the catchments for which a threshold could not be

385   identified were excluded from the correlation analyses involving those metrics. In

386   spite of the small sample sizes (five to nine sites), some significant correlations

387   were observed between the three metrics of catchment hydrological behaviour

388   and some hydroclimatic and topographic catchment properties of the North-

389   Watch sites (Table 6). For rainfall-triggered events, regardless of whether the

390   storage deficit ($W_{storage}$) was considered, $r_{Spearman}$ was positively linked to the

391   median flow path distance to the stream ($corr_{catchment}$ = 0.81 and 0.83, p-value <

392   0.05). When $W_{storage}$ was not considered, $r_{Spearman}$ values were positively related

393   to catchment mean elevation ($corr_{catchment}$ = 0.71, p-value < 0.05) while the rainfall

394   threshold value was positively correlated with mean temperature

395 (corr$_{catchment}$ = 0.66, p-value < 0.05). When W$_{storage}$ was considered, however, the

396 strength of input-output relationships was especially high for high relief

397 catchments (corr$_{catchment}$ = 0.67, p-value < 0.05), and the rainfall threshold value

398 was positively correlated with the median sub-catchment area (corr$_{catchment}$ = 0.81,

399 p-value < 0.05). A few significant catchment controls were also identified for

400 metrics of hydrological behaviour for snowmelt-driven events; however, they

401 should be interpreted with caution given the more uncertain identification of

402 thresholds generally for those events. For instance, when W$_{storage}$ was not taken

403 into account, the water input (rain+snowmelt) threshold values were positively

404 correlated with the catchment mean elevation (corr$_{catchment}$ = 0.66, p-value < 0.05).

405 The strength of the input-output relationship was also positively correlated with

406 the median sub-catchment area (corr$_{catchment}$ = 0.88, p-value < 0.05). The R^2 of

407 input-output data above the threshold was correlated with a limited number of

408 catchment controls: statistically significant negative correlations were notably

409 present with the BFI and two slope segments of the catchments' flow duration

410 curves (Table 6).

411

412

413 **4   Discussion**

414

415 ***4.1   Hydrological insights from threshold detection in northern catchments***

416

417   *4.1.1   Storage deficit conceptualization*

418

419   This study sought to better understand one key aspect of catchment nonlinear

420   behaviour: the threshold precipitation-discharge response in northern catchments – with and

421   without snowmelt influence. Such thresholds reflect the integration of various levels of

422   catchment complexity (Zehe and Sivapalan, 2009), and conceptually they indicate when a

423   critical value in a hydrological state variable becomes exceeded and a rapid flow generation

424   mechanism responsible for event runoff is initiated. This contrasts with times when the same

425   hydrological state variable has a value below the threshold and the rapid flow-producing

426   mechanisms are switched off or are less active (O'Kane and Flynn, 2007). As the memory of

427   these switches is local in both space and time (O'Kane and Flynn, 2007), the challenge is to

428   understand the controls exerted on thresholds when predicting catchment-scale hydrologic

429   response across geographic regions.

430

431   For the range of mid- to high-latitude northern catchments considered here, the visual

432   identification of thresholds was useful for characterizing catchment hydrological behaviour with

433   and without the consideration of dynamic water storage dynamics (or antecedent wetness

434   conditions). That approach is similar to that adopted by Detty and McGuire (2010) who plotted

435   total quickflow against the sum of gross precipitation and an antecedent soil moisture index

436   (ASI). They found that the input-output relationships were stronger and easier to characterize

437   when the ASI was considered in addition to gross precipitation. We found that the strength of

438   the input-output relationships (i.e., the Spearman rank correlation coefficient $r_{Spearman}$) was

439   slightly greater when the difference $W_{input} - W_{storage}$ was used as the hydrological input variable

440   rather than W_{input}. It is, however, worth noting that the assumption of instantaneous (i.e., sub-

441   daily) routing might be incorrect, and hence the variable $W_{storage}$ might capture delays related to

442   both storage deficit satisfaction and catchment routing rather than delays related to storage

443   deficit satisfaction only.

444

445   This study dealt only with runoff initiation thresholds and input magnitude (i.e., effective

446   precipitation input) thresholds while input intensity thresholds were ignored. Indeed, storage

447   capacity (or storm amount) thresholds are often associated with saturation excess flow

448   mechanisms while rainfall intensity thresholds can be associated with infiltration excess flow

449   mechanisms, and their differentiation is not always straightforward (McGrath *et al.*, 2007). Past

450   work has been conducted in humid temperate, forested catchments with very high soil

451   infiltration capacities (Tromp-Van Meerveld and McDonnell, 2006a; Graham and McDonnell,

452   2010) where rainfall intensity was examined but had little effect on threshold values of runoff

453   production (as suggested by Hewlett and Hibbert, 1967). Nevertheless, some of the North-

454   Watch catchments (Krycklan, Sleepers River) develop conditions where infiltration excess runoff

455   could occur, such as runoff over frozen ground (Shanley and Chalmers, 1999; Laudon *et al.*,

456   2007); these effects could not be considered in the present study due to the lack of empirical

457   data necessary for all sites.

458

459   *4.1.2   Comparison of rainfall and snowmelt events*

460

461   No clear input-output pattern could be discerned for some catchments (Figures 2, 3 and

462   5, Table 4) and it was difficult to assess whether this reflected a process reality or a

463   data/methodological problem, especially regarding the degree-day method used to

464   approximate snowmelt inputs and or the criteria used to define rainfall-triggered and

465   snowmelt-driven events across all nine sites. It is likely that degree-day methods worked better

466   for warmer sites with snowpacks near 0°C, in contrast to colder sites (e.g., Krycklan, Wolf Creek)

467   where snowpack energetics have a greater influence on the hydrological cycle. Such differences

468   were not taken into account in the current study. The use of similar criteria for the definition of

469   hydrological events across all nine sites was also problematic as it occasionally resulted in very

470   large event precipitation amounts (e.g., up to 100 mm for single events at Dorset and up to 200

471   mm at Wolf Creek). These corresponded to compound hydrographs produced by a succession of

472   smaller events rather than single hydrograph peaks, and this raised the issue of how

473   hydrological events should be defined under contrasting conditions within inter-site

474   comparisons. While it may have been beneficial to use site-specific and event type-specific

475   (rainfall-triggered versus snowmelt-driven) criteria to divide precipitation and runoff events, the

476   objective here was only to transfer the methodology from temperate environments to higher-

477   latitude catchments. The use of daily rather than hourly or sub-hourly precipitation and

478   discharge values was also dictated by data availability and built upon a previous threshold

479   identification study at the catchment scale at the Maimai and HJ Andrews sites (i.e., Graham

480   and McDonnell, 2010)—the criteria used for the delineation of water input events in the current

481   study were similar to those used by Graham and McDonnell (2010) in their definition of storm

482   events. We acknowledge that the use of daily data, as well as the sole selection of hydrological

483 events with a minimum 1-day delay between the start of the water input event and the

484 initiation of runoff response, has likely biased the present analysis: since flashy events occurring

485 at the scale of hours were effectively discarded, the identified threshold dynamics are only be

486 applicable to longer duration events. Clearly more research is needed: provided the availability

487 of sub-daily weather and hydrometric data, a sensitivity analysis could be conducted to assess

488 the identifiability of hydrologic thresholds depending on event definition criteria.

489

490 　　　Although nonlinear behaviour associated with rainfall-triggered events was identified for

491 more than half of the nine investigated catchments (e.g., Figure 4), snowmelt-driven events

492 were problematic: the common conceptualization of nonlinear hydrological behaviour, namely

493 the hockey stick-like input-output relationship, appeared to work fairly well for rainfall-triggered

494 events but not for snowmelt-driven events. In some cases, the lack of clear input-output

495 relationships with snowmelt-driven events is likely a true reflection of different physical

496 processes. For instance, the Girnock and the Mharcaidh catchments were the only ones for

497 which snowmelt-driven threshold values were systematically smaller than their rainfall-driven

498 counterparts (Table 4), and this might be explained by the fact that the two Scottish catchments

499 have the smallest and most transient snowpacks, hence the lack of potential for snowpack

500 storage of early melt events. Also, snowfall usually occurs in the wettest winter months in the

501 two Scottish catchments when there is little available storage in the soil.

502

503 *4.1.3　Catchment controls*

504

505   Despite the useful case studies published to date, work until now has not explored any

506   tangible rules to up-scale or down-scale threshold values based on drainage basin properties.

507   The work reported here shows some statistically significant correlations between hydrological

508   behaviour and hydroclimatic or physical catchment properties (Table 6) that appear to have a

509   physical basis. For instance, it could be inferred from Table 6 that for rainfall-triggered events,

510   the greater the median flow path distance to the stream, the higher the $r_{Spearman}$ for rainfall-

511   triggered events— illustrating the strength of the hydrological input-output relationship. Flow

512   path distance to the stream is a good surrogate measure for hydrologic proximity, which is a

513   precursor to identifying which parts of the catchment are the most likely to be connected to the

514   channel and contribute to streamflow (Ali and Roy, 2010). It can be hypothesized that the

515   greater the median flow path distance to the stream, the more likely that remote catchment

516   areas will be connected to the stream when specific hydrological conditions are reached or

517   exceeded; hence the stronger the input-output relationship and the weaker the threshold

518   effect. When dynamic storage deficits were not considered, catchments with higher mean

519   temperatures were also associated with greater rainfall thresholds. This correlation needs to be

520   interpreted with caution since temperature and precipitation are highly correlated for the

521   North-Watch catchments (Carey *et al.*, 2010). In the absence of water limitation, one hypothesis

522   is that catchments in warmer climates are subjected to greater evaporation, hence the higher

523   critical water input value needed to generate significant quickflow. However, this hypothesis is

524   difficult to verify at the event scale. When storage deficits were considered, input-output

525   relationships were stronger in the high-elevation and mostly headwater sites, confirming the

526   tight coupling between input and outputs in smaller catchments. The rainfall threshold value

527 was correlated with the median sub-catchment area, which suggests that threshold dynamics

528 are not influenced by the basin's total drainage area but rather by the spatial organization and

529 topology of hydrological response units (Buttle, 2006). This is consistent with previous scaling

530 work which revealed that streamwater mean residence time was unrelated to basin area, but

531 rather strongly controlled by internal distributions of flowpath length and gradient or drainage

532 density (McGuire *et al.*, 2005) or by soil typology (Hrachowitz *et al.*, 2009). It is, however, worth

533 noting that the differences between controls on rain-only events and those on rain+snowmelt

534 events might be due to the small number of sites (n ≤ 9): slight changes in the ranking of

535 catchments according to their threshold values for different types of events can indeed lead to

536 significant differences in the computed Spearman rank corr$_{catchment}$ values.

537

538

539 **4.2   *How do our results compare to previous threshold studies?***

540

541 *4.2.1 Threshold types and values*

542

543   Some of the runoff initiation and effective precipitation input thresholds reported here

544 fell within the range of previously published data for rainfall-triggered events (Table 4) while

545 others were well above this range. Most threshold values reported in the literature at the

546 hillslope scale range from 55 mm or less (Whipkey, 1965; Mosley, 1979; Tani, 1997; Buttle *et al.*,

547 2004; Uchida *et al.*, 2005, 2006; Tromp-Van Meerveld and McDonnell, 2006a) whereas

548 catchment-scale studies (Graham and McDonnell, 2010; Penna *et al.*, 2011) have identified

549 input rainfall threshold values that cover a wider range than hillslope studies (e.g., 23 mm in a

550 1.9 km² headwater catchment in the Italian Alps; 8.5 mm in a 3.8 ha catchment at Maimai; 56+

551 mm at one of the North-Watch sites, HJ Andrews). Graham and McDonnell (2010) found that

552 the rainfall threshold could vary from 0 mm to 83 mm at HJ Andrews depending on antecedent

553 drainage within two nested 9 to 101.3 ha sub-watersheds. In previously published studies,

554 thresholds for hillslope runoff initiation have been found to be greater than runoff initiation

555 thresholds for the catchments these hillslopes reside in. This is likely due to the fact that

556 additional geomorphic features at the catchment scale (i.e., riparian zones) are closely linked to

557 the channel and show more immediate connection to catchment flow response. McGlynn and

558 McDonnell (2003) showed strong hysteresis in streamflow response to storm rainfall, whereby

559 rising groundwater levels in riparian zones occur before the rising limb of the storm hydrograph

560 and the threshold-like hillslope response precedes the falling limb of the storm hydrograph.

561 Thus, larger catchment-scale runoff initiation thresholds can also be attributed to the variable

562 buffering potential of the riparian zone (McGlynn and McDonnell, 2003). Similarly, Tetzlaff *et al*.

563 (2014) found that a peatland riparian zone in a sub-catchment of the Girnock provided runoff

564 responses to small events (>3mm), but only in larger events (>30mm) did surrounding hillslopes

565 connect and produce a non-linear increased runoff response. These findings are broadly

566 consistent with those of the present study.

567

568 *4.2.2   Shapes of nonlinear input-output relationships*

569

570   Previous studies have shown different shapes of nonlinear hydrological behaviours such

571   as the hockey stick (e.g., Weiler *et al.*, 2005; Tromp-Van Meerveld and McDonnell, 2006a; Detty

572   and McGuire, 2010), the Heaviside or step function (e.g., James and Roulet, 2007) or the sigmoid

573   function (e.g., Zehe *et al.*, 2007); nevertheless reasons behind those different shapes have not

574   been explained. At the hillslope scale, input-output relations from a range of hillslopes were

575   shown to fit the hockey stick shape with the only nuance that the slope of the relationship after

576   the threshold varied among the sites (Weiler *et al.*, 2005). In this paper, a single

577   conceptualization of nonlinear behaviour (the hockey stick) was applied to all nine catchments

578   for the sake of simplicity and site comparison and not necessarily because similarities in

579   underlying processes, connectivity structure between landscape units or storage capacities were

580   assumed across the nine sites (following the logic of Lehmann *et al.*, 2007). Surprisingly,

581   however, when strong nonlinear input-output relationships were present in the data, the

582   hockey stick conceptualization seemed appropriate to portray the catchments' hydrological

583   behaviour (e.g., Figure 4). When thresholds were identifiable, the R^2 (that is correlated to the

584   slope) of the input-output relationship above the threshold was specific to each catchment, as

585   found by Weiler *et al.* (2005) at the hillslope scale. As for the input-output patterns labelled as

586   "unclear" (Table 4), it was not possible to say from a visual assessment alone whether other

587   types of nonlinear functions would have fitted the data better and led to the identification of

588   hydrologic thresholds.

589

590   *4.2.3   Threshold identification methodology*

591

592   The largest methodological challenge in this study was the visual identification of

593   nonlinear behaviours from scatter plots, namely the identification of the critical input value as

594   the first point where the input-output "curve" departs from zero or a given minimum level in

595   Figures 2-5. While previous hydrologic studies showed rather clear input-output threshold

596   relationships due to less data points or clearer dynamics at the hillslope scale, such was not the

597   case here where we identified the clearest inflection point in each scatter plot – when it existed.

598   On most plots in Figures 2-5, however, data points tended to cluster in a band around the

599   inflection point, thus suggesting a range of possible threshold values. For each catchment, the

600   thresholds reported in Table 4 were the highest among the range of possible values. While this

601   methodological choice likely led to a bias towards higher threshold values, it was assumed that

602   this bias would be consistent across all sites and would not change the ranking of the

603   catchments when sorted according to ascending threshold values.

604

605   Detty and McGuire (2010), who worked at one of the North-Watch sites (Hubbard

606   Brook), implied that well identified precipitation input thresholds should be associated with an

607   almost zero slope below the threshold and an R^2 value close to 1 above the threshold. One can,

608   however, hypothesize that in the context of large and complex catchments, process-specific

609   thresholds likely combine to determine the overall switching "on and off" of runoff contributing

610   zones at the catchment scale, and this superimposition of process dynamics could lead to

611   piecewise (or hybrid) input-output functions. This is especially probable for cold-region

612   landscapes where snowmelt can significantly increase the amount of active source areas and

613   water inputs to the stream through a cascade of soil moisture storage thresholds, snowpack

614　water storage thresholds, and radiation intensity thresholds that influence the rate of ground

615　thaw. Such a superposition of storage and intensity thresholds would make any visual

616　assessment of precipitation-runoff response impossible and rather require a mathematically-

617　based detection method (Lintz *et al.*, 2011). Here it is suggested that in the case of complex

618　cold region catchments in particular, nonlinear and domain-dependent mathematical functions

619　should be examined with regards to their potential to account for multiple storage and/or

620　intensity thresholds driving the system over different possible ranges of inputs.

621

622

623　**5　　Conclusion**

624

625　　　The novel contributions of this study were to: (i) shift the focus from single humid

626　temperate catchments to a range of contrasting mid- to high-latitude catchments; (ii) test for

627　the existence of runoff initiation and effective precipitation thresholds in rainfall-driven versus

628　snowmelt-driven conditions; and (iii) investigate physiographic and hydroclimatic drivers behind

629　precipitation-runoff response. The work could be useful to up-scale or down-scale threshold

630　values based on drainage basin properties and hydroclimatic properties. Storm amount critical

631　values were quantified and out of the nine catchments investigated, one was characterized by a

632　linear input-output behaviour while the others were mainly associated with nonlinear

633　behaviours. The consideration of antecedent storage deficit slightly improved the ability to

634　characterize the different rainfall-runoff catchment dynamics. For rainfall-triggered events,

635　catchment hydroclimatic or physical characteristics such as the median flow path distance to

636 the stream, the mean annual air temperature or the median sub-catchment area were strong

637 predictors of either the strength of the hydrological input-output relationship or the actual

638 runoff initiation or effective precipitation threshold value identified for each site. The

639 characterization of snowmelt-runoff catchment dynamics was more difficult, however,

640 suggesting that the sole focus on input magnitude thresholds (i.e., storage thresholds) might be

641 insufficient to understand catchment behaviour when snowmelt constitutes a large portion of

642 the water input. Further studies are therefore needed to investigate the relative effects of

643 storage and intensity thresholds in northern regions where energy dynamics are critical in

644 runoff generation.

645

646

647 **Acknowledgements**

651

652   **References**

653   Ali, G. A. and A. G. Roy (2010), Shopping for Hydrologically Representative Connectivity Metrics

654   in a Humid Temperate Forested Catchment, *Water Resources Research, 46*, W12544,

655   doi:10.1029/2010WR009442.

656   Allan, C. J. and N. T. Roulet (1994), Runoff generation in zero-order Precambrian Shield

657   catchments: the stormflow response of a heterogeneous landscape, *Hydrological*

658   *Processes, 8*, 369–388.

659   Anderson, N. H. (1992), Influence of disturbance on insect communities in Pacific Northwest

660   streams, *Hydrobiologia, 248*, 79-92.

661   Bailey, A. S., J. W. Hornbeck, J. L. Campbell, and C. Eagar (2003), *Hydrometerological database*

662   *for Hubbard Brook Experimental Forest: 1955-2000*. USDA Forest Service, NE Research

663   Station General Technical Report NE-305.

664   Bates, C. G. and A. J. Henry (1928), Forest and streamflow experiment at Wagon Wheel Gap,

665   Colorado, *Monthly Weather Review Supplement, 30*, 1-79.

666   Bayfield, N. G. and A. J. Nolan (1998), Vegetation and soils of the Allt a'Mharcaidh catchment,

667   Cairngorm mountains, *Scottish Geographical Magazine, 114*, 18-21.

668   Beven, K. J. and M. J. Kirkby (1979), A physically based variable contributive area model of basin

669   hydrology, *Hydrological Sciences Bulletin, 24*, 43-69.

670   Buttle, J. M. (2006). Mapping first-order controls on streamflow from drainage basins: the T3

671   template, *Hydrological Processes, 20*, 3415–3422.

672   Buttle, J. M., P. J. Dillon, and G. R. Eerkes (2004), Hydrologic coupling of slopes, riparian zones

673   and streams: an example from the Canadian Shield, *Journal of Hydrology, 287*, 161-177.

674   Carey, S. K. and W. L. Quinton (2005), Evaluating Runoff Generation During Summer Using

675   Hydrometric, Stable Isotope and Hydrochemical Methods in a Discontinuous Permafrost

676   Alpine Catchment, *Hydrological Processes, 19*(1), 95-114.

677   Carey, S. K., D. Tetzlaff, J. Seibert, C. Soulsby, J. Buttle, H. Laudon, J. McDonnell, K. McGuire, D.

678   Caissie, J. Shanley, M. Kennedy, K. Devito, and J. W. Pomeroy (2010), Inter-Comparison

679   of Hydro-Climatic Regimes Across Northern Catchments: Synchronicity, Resistance and

680   Resilience, *Hydrological Processes, 24*(24), 3591-3602.

681   Carey, S. K., D. Tetzlaff, J. Buttle, H. Laudon, J. McDonnell, K. McGuire, J. Seibert, C. Soulsby, and

682   J. Shanley (2013), Use of colormaps and wavelet coherence to discern short and longer-

683   term climate influences on streamflow variability in northern catchments, *Water*

684   *Resources Research, 49*, 1-14, doi: 10.1002/wrcr.20469.

685   Detty, J. M. and K. J. McGuire (2010), Threshold Changes in Storm Runoff Generation at a Till-

686   Mantled Headwater Catchment, *Water Resources Research, 46*, W07525,

687   doi:10.1029/2009WR008102.

688   Dickinson, W. T. and H. Whiteley (1970), Watershed areas contributing to runoff, in *Results of*

689   *Research on Representative and Experimental Basins (Proceedings of the Wellington*

690   *Symposium)*, IAHS Publication Number 96: Oxfordshire, United Kingdom; 12-26.

691   Eimers, M. C., S. A. Watmough, and J. M. Buttle (2008), Long-term trends in dissolved organic

692   carbon concentration: a cautionary note, *Biogeochemistry, 87*, 71-81.

693   Gardner, K. K. and B. L. McGlynn (2009), Seasonality in spatial variability and influence of land

694   use/land cover and watershed characteristics on stream water nitrate concentrations in

695   a developing watershed in the Rocky Mountain West, *Water Resources Research, 45*,

696   doi:10.1029/2008wr007029.

697 Goodyear, D.R. (1997), Hillslope-riparian interactions in a forested Canadian Shield basin, MSc

698   thesis, Watershed Ecosystems Graduate Program, Trent University, Peterborough,

699   Ontario, Canada, pp. 111.

700 Graham, C. B. and J. J. McDonnell (2010), Hillslope Threshold Response to Rainfall: (2)

701   Development and Use of a Macroscale Model, *Journal of Hydrology, 393*(1-2), 77-93.

702 Hamon, W. R. (1961), Estimating potential evaporation, *Journal of Hydraulics Division,*

703   *Proceedings of the American Society of Civil Engineers, 871*, 107–120.

704 Hewlett, J. D. and A. R. Hibbert (1967), Factors affecting the response of small watersheds to

705   precipitation in humid areas, in *Proceedings of the International Symposium on Forest*

706   *Hydrology*, Pergamon: New York; 275-290.

707 Hjerdt, K. N., J. J. McDonnell, J. Seibert, and A. Rodhe (2004), A New Topographic Index to

708   Quantify Downslope Controls on Local Drainage, *Water Resources Research, 40*(5),

709   W05602, doi:10.1029/2004WR003130.

710 Hrachowitz, M., C. Soulsby, D. Tetzlaff, J. J. C. Dawson, and I. A. Malcolm (2009), Regionalization

711   of Transit Time Estimates in Montane Catchments by Integrating Landscape Controls,

712   *Water Resources Research, 45*, W05421, doi:10.1029/2008WR007496.

713 James, A. L. and N. T. Roulet (2007), Investigating Hydrologic Connectivity and Its Association

714   With Threshold Change in Runoff Response in a Temperate Forested Watershed,

715   *Hydrological Processes, 21*(25), 3391-3408.

716 Juston, J., J. Seibert, and P. O. Johansson (2009). Temporal sampling strategies and uncertainty

717   in calibrating a conceptual hydrological model for a small boreal catchment. *Hydrological*

718   *Processes, 23(21)*, 3093-3109.

719   Kane, D.L. and D. Yang (2004), Northern Research Basins Water Balance, *IAHS Publ. No. 290*,

720   IAHS Press, Wallingford.

721   Laudon, H., V. Sjoblom, I. Buffam, J. Seibert, and M. Morth (2007), The Role of Catchment Scale

722   and Landscape Characteristics for Runoff Generation of Boreal Streams, *Journal of*

723   *Hydrology, 344*, 198-209.

724   Laudon, H., I. Taberman, A. Ågren, M. Futter, M. Ottosson-Löfvenius, and K. Bishop (2013a), The

725   Krycklan Catchment Study—A flagship infrastructure for hydrology, biogeochemistry, and

726   climate research in the boreal landscape, *Water Resources Research, 49*,

727   doi:10.1002/wrcr.20520.

728   Laudon, H., D. Tetzlaff, C. Soulsby, S., J. Seibert, J. Buttle, J. Shanley, J. McDonnell, and K.

729   McGuire (2013b), Change in winter climate will affect dissolved organic carbon and water

730   fluxes in mid-to-high latitude catchments, *Hydrological Processes, 27*, 700-709.

731   Lehmann, P., C. Hinz, G. McGrath, H. J. Tromp-Van Meerveld, and J. J. McDonnell (2007), Rainfall

732   Threshold for Hillslope Outflow: an Emergent Property of Flow Pathway Connectivity,

733   *Hydrology and Earth System Sciences, 11*(2), 1047-1063.

734   Likens, G. E. and F. H. Bormann (1990), *Biogeochemistry of a forested ecosystem*, New York:

735   Springer-Verlag.

736   Lintz, H. E., B. McCune, A. N. Gray, and K. A. McCulloh (2011), Quantifying Ecological Thresholds

737   From Response Surfaces, *Ecological Modelling, 222*(3), 427-436.

738   Lyon, S. W., M. T. Walter, P. Gerard-Marchant, and T. S. Steenhuis (2004), Using a Topographic

739   Index to Distribute Variable Source Area Runoff Predicted With the SCS Curve-Number

740　　　　　　Equation , *Hydrological Processes, 18*(15), 2757-2771.

741　McCartney, S. E., S. K. Carey, and J. W. Pomeroy (2006), Intra-basin variability of snowmelt

742　　　　　　water balance calculations in a subarctic catchment, *Hydrological Processes, 20*, 1001-

743　　　　　　1016.

744　McGlynn, B. L. and J. J. McDonnell (2003), Quantifying the Relative Contributions of Riparian and

745　　　　　　Hillslope Zones to Catchment Runoff, *Water Resources Research, 39*(11), 1310,

746　　　　　　doi:10.1029/2003WR002091.

747　McGlynn, B., J. McDonnell, M. Stewart, and J. Seibert (2003), On the Relationships Between

748　　　　　　Catchment Scale and Streamwater Mean Residence Time, *Hydrological Processes, 17*(1),

749　　　　　　175-181.

750　McGlynn, B. L. and J. Seibert (2003), Distributed Assessment of Contributing Area and Riparian

751　　　　　　Buffering Along Stream Networks, *Water Resources Research, 39*(4), 1082,

752　　　　　　doi:1010.1029/2002WR001521.

753　McGrath, G. S., C. Hinz, and M. Sivapalan (2007), Temporal Dynamics of Hydrological Threshold

754　　　　　　Events, *Hydrology and Earth System Sciences, 11*(2), 923-938.

755　McGuire, K. J., J. J. McDonnell, M. Weiler, C. Kendall, B. L. McGlynn, J. M. Welker, and J. Seibert

756　　　　　　(2005), The Role of Topography on Catchment-Scale Water Residence Time, *Water*

757　　　　　　*Resources Research, 41*(5), W05002, doi:10.1029/2004WR003657.

758　McGuire, K. J. and J. J. McDonnell (2010), Hydrological Connectivity of Hillslopes and Streams:

759　　　　　　Characteristic Time Scales and Nonlinearities, *Water Resources Research, 46*, W10543,

760　　　　　　doi:10.1029/2010WR009341.

761　Mosley, M. P. (1979), Streamflow generation in a forested watershed, *Water Resources*

762   *Research, 15*, 795-806.

763   O'Kane, J. P. and D. Flynn (2007), Thresholds, Switches and Hysteresis in Hydrology From the

764   Pedon to the Catchment Scale: a Non-Linear Systems Theory, *Hydrology and Earth*

765   *System Sciences, 11*(1), 443-459.

766   Penna, D., H. J. Tromp-Van Meerveld, A. Gobbi, M. Borga, and G. Dalla Fontana (2011), The

767   Influence of Soil Moisture on Threshold Runoff Generation Processes in an Alpine

768   Headwater Catchment, *Hydrology and Earth System Sciences, 15*(3), 689-702.

769   Phillips, J. D. (2003), Sources of Nonlinearity and Complexity in Geomorphic Systems, *Progress in*

770   *Physical Geography, 27*(1), 1-23.

771   Phillips, J. D. (2006), Evolutionary Geomorphology: Thresholds and Nonlinearity in Landform

772   Response to Environmental Change, *Hydrology and Earth System Sciences, 10*(5), 731-

773   742.

774   Quinn, P., K. Beven, P. Chevallier, and O. Planchon (1991), The Prediction of Hillslope Flow Paths

775   for Distributed Hydrological Modeling Using Digital Terrain Models, *Hydrological*

776   *Processes, 5*(1), 59-79.

777   Quinton, W. L. and P. Marsh (1999), A Conceptual Framework for Runoff Generation in a

778   Permafrost Environment, *Hydrological Processes*, *13*, 2563-2581.

779   Robins, N. S. (1990), *Hydrogeology of Scotland*, London, UK: HMSO.

780   Seibert, J. and B. McGlynn (2007), A new triangular multiple flow direction algorithm for

781   computing upslope areas from gridded digital elevation models, *Water Resources*

782   *Research, 43*, W04501, doi:10.1029/2006WR005128.

783   Shanley, J. B. and A. Chalmers (1999), The effect of frozen soil on snowmelt runoff at Sleepers

784   River, Vermont, *Hydrological Processes, 13*, 1843-1857.

785   Sivakumar, B. (2005), Hydrologic Modeling and Forecasting: Role of Thresholds, *Environmental*

786   *Modelling & Software, 20*(5), 515-519.

787   Sloto, R. A. and M. Y. Crouse (1996), *HYSEP: A computer program for streamflow hydrograph*

788   *separation and analysis*. U.S. Geological Survey Water-Resources Investigations Report

789   96-4040.

790   Spence, C. (2007), On the Relation Between Dynamic Storage and Runoff: a Discussion on

791   Thresholds, Efficiency, and Function, *Water Resources Research, 43*(12), W12416,

792   doi:10.1029/2006WR005645.

793   Spence, C. and M. K. Woo (2002), Hydrology of subarctic Canadian Shield: bedrock upland,

794   *Journal of Hydrology, 262*, 111–127.

795   Spence, C. and M. K. Woo (2003), Hydrology of subarctic Canadian Shield: Soil filled valleys,

796   *Journal of Hydrology, 279*, 151-166.

797   Spence, C. and M. K. Woo (2006), Hydrology of subarctic Canadian Shield: heterogeneous

798   headwater basins, *Journal of Hydrology, 317*, 138–154.

799   Steenhuis, T. S., M. Winchell, J. Rossing, J. A. Zollweg, and M. F. Walter (1995), SCS runoff

800   equation revisited for variable source runoff areas, *ASCE Journal of Irrigation and*

801   *Drainage Engineering, 121*, 234-238.

802   Tani, M. (1997), Runoff Generation Processes Estimated From Hydrological Observations on a

803   Steep Forested Hillslope With a Thin Soil Layer, *Journal of Hydrology, 200*(1-4), 84-109.

804   Tetzlaff, D., C. Soulsby, S. Waldron, I. A. Malcolm, P. J. Bacon, S. M. Dunn, A. Lilly, and A. F.

805   Youngson (2007), Conceptualization of Runoff Processes Using a Geographical

806   Information System and Tracers in a Nested Mesoscale Catchment, *Hydrological*

807   *Processes, 21*(10), 1289-1307.

808   Tetzlaff, D., S. Carey and C. Spence (2013a), Preface - Catchments in the future North:

809   interdisciplinary science for sustainable management in the 21st Century, *Hydrological*

810   *Processes*, *27*, 635–639.

811   Tetzlaff, D., C. Soulsby, J. Buttle, R. Capell, S. Carey, L. Kruitbos, H. Laudon, J. McDonnell, K.

812   McGuire, J. Seibert, and J. Shanley (2013b), Catchments on the cusp? Structural and

813   functional change in northern ecohydrology, *Hydrological Processes, 27*, 766-774.

814   Tetzlaff, D., C. Birkel, J. Dick, J. Geris, and C. Soulsby (2014), Storage dynamics in

815   hydropedological units control hillslope connectivity, runoff generation and the

816   evolution of catchment transit time distributions. *Water Resources Research*,

817   doi: 10.1002/2013WR014147

818   Tromp-Van Meerveld, H. J. and J. J. McDonnell (2006a), Threshold Relations in Subsurface

819   Stormflow: 1. A 147-Storm Analysis of the Panola Hillslope, *Water Resources Research,*

820   *42*(2), W02410, doi:10.1029/2004WR003778.

821   Tromp-Van Meerveld, H. J. and J. J. McDonnell (2006b), Threshold Relations in Subsurface

822   Stormflow: 2. The Fill and Spill Hypothesis, *Water Resources Research, 42*(2), W02411,

823   doi:10.1029/2004WR003800.

824   Uchida, T., I. Tromp-van Meerveld, and J. J. McDonnell (2005), The role of lateral pipe flow in

825   hillslope runoff response: an intercomparison of non-linear hillslope response, *Journal of*

826   *Hydrology, 311*, 117–133.

827   Uchida, T., J. J. McDonnell, and Y. Asano (2006), Functional Intercomparison of Hillslopes and

828   Small Catchments by Examining Water Source, Flowpath and Mean Residence Time, *Journal*

829   *of Hydrology, 327*(3-4), 627-642.

830   Weiler, M., J. McDonnell, I. Tromp-Van Meerveld, and T. Uchida (2005), Subsurface stormflow,

831   in *Encyclopedia of Hydrological Sciences*, edited by Anderson, M. G., John Wiley & Sons:

832   Chichester, UK; 1719-1732.

833   Whipkey, R. Z. (1965), Subsurface stormflow from forested slopes, *International Association of*

834   *Scientific Hydrology Bulletin, 10*(2), 74-85.

835   Zehe, E., H. Elsenbeer, F. Lindenmaier, K. Schulz, and G. Blöschl (2007), Patterns of predictability

836   in hydrological threshold systems, *Water Resources Research, 43*, W07434,

837   doi:10.1029/2006WR005589.

838   Zehe, E. and M. Sivapalan (2009), Threshold behaviour in hydrological systems as (human) geo-

839   ecosystems: manifestations, controls, implications, *Hydrology and Earth System*

840   *Sciences, 13*, 1273–1297.

841   **Table 1.** Selected characteristics of the North-Watch catchments. Evaporation and storage values are derived from annual water
842   balance estimates (Carey *et al.*, 2010). Q5 and Q95 are the flow values that are exceeded 95% and 5% of the time, respectively.
843

Catchment name	Gauging site	Coded name	Area (km²)	Mean elevation (m)	Relief (m)	Mean annual temperature (deg C)	Mean annual precipitation (mm)	Percentage of snow (%)	Mean annual evaporation (mm)	Mean annual runoff (mm)	Q5 (mm)	Q95 (mm)	Mean annual storage change (mm)
Strontian	Polloch	STR	8	340	740	9.08	2632	4	417	2213	30.08	0.1	206
Mharcaidh	Site 1	MHA	10	704	779	5.7	1222	20	326	873	10.56	0.39	146
Girnock	Littlemill	GIR	30	405	620	6.73	1059	10	453	603	8.96	0.02	175
Hubbard Brook	W3	HUB	0.41	642	210	6.41	1381	25	497	882	17.58	0.01	255
Sleepers River	W9	SLE	0.41	604	167	4.66	1256	25	510	743	10.11	0.01	336
H.J. Andrews	Mack Creek	HJA	5.81	1200	860	9.22	2158	40	412	1744	26.54	0.29	561
Wolf Creek	Granger	WOL	7.6	1700	750	-2.15	478	45	127	352	4.84	0.07	141
Dorset	Harp Lake 5	DOR	1.9	373	93	4.94	980	28	401	577	11.16	0	263
Krycklan	S7	KRY	0.5	280	72	2.41	651	40	323	327	6.42	0	191

844

845　**Table 2.** Data used for the identification of hydrological events for the North-Watch sites.

846

Catchment name	Length of daily rainfall-runoff record used	Total number of events identified	Number of rainfall-only events	Number of snowmelt-driven events
Strontian	From 2-Jul-89 until 12-Dec-96	161	154	7
Mharcaidh	From 1-Jan-90 until 30-Jul-94	109	66	43
Girnock	From 1-Jan-72 until 18-Mar-94	543	370	173
Hubbard Brook	From 1-Oct-58 until 30-Sep-07	1080	663	417
Sleepers River	From 1-Oct-91 until 20-May-01	206	112	94
H.J. Andrews	From 17-Oct-98 until 21-Sep-04	105	61	44
Wolf Creek	From 8-Apr-98 until 4-Oct-08	82	52	29
Dorset	From 1-Nov-76 until 29-Apr-02	498	304	194
Krycklan	From 5-Oct-90 until 31-Dec-07	269	158	111

847

848   **Table 3.** Catchment characteristics tested against the three metrics of hydrological behaviour
849   (i.e., $r_{Spearman}$ between hydrological inputs and outputs, threshold value, and R^2 values for data
850   above the threshold value).

851

Variable name	Description
MeanTemperature	Mean annual daily temperature (deg C)
MeanPrecipitation	Mean annual total precipitation (mm)
PrctSnow	Mean percentage of total annual precipitation that falls as snow
MeanEvaporation	Mean annual total evaporation (mm); computations involved using the potential evaporation formula of Hamon (1961) and deriving actual values using a correction factor (Carey *et al.*, 2010).
MeanStorage	Mean annual storage (mm) derived using annual water balance estimates (Carey *et al.*, 2010)
Area	Catchment drainage area (km²)
MeanElevation	Mean elevation value (m) computed over the whole catchment area
Relief	Catchment relief (m)
BFI	Baseflow index - the long-term ratio of total baseflow to total streamflow
FDCS_lowflow	Slope of the flow duration curve computed between the 70th and the 100th percentiles
FDCS_intermediateflow	Slope of the flow duration curve computed between the 30th and the 70th percentiles
FDCS_highflow	Slope of the flow duration curve computed between the 0th and the 30th percentiles
ElevationAboveStream	Median elevation above the stream (m)
DistanceFromStream	Median flow path distance to the stream (m)
GradientToStream	Median gradient to the stream (m)
TransitTimeProxy	Median value of the ratio of flowpath length to flowpath gradient
D8Gradient	Median terrain slope computed using the D8 flow algorithm
DinfGradient	Median terrain slope computed using the D∞ flow algorithm
d5	Downslope index; median value of the gradient towards the closest point which is at least 5 m (in altitude) below a target catchment pixel
SubcatchmentArea	Median sub-catchment area (km²)
UpslopeArea	Median upslope area (km²)
TWI	Median value of ln(a/tan β) where a is the upslope area per unit contour length and tan β is the D8 gradient for each catchment pixel
TWId5	Same as the TWI except that the downslope index gradient (5 m) is used as a slope surrogate instead of the D8 gradient

852 **Table 4.** Visually identified thresholds for the North-Watch sites. A threshold value is reported
853 when a nonlinear hydrological response can be discerned in Figures 2 to 5; the term "linear" is
854 used when a linear hydrological response (i.e., no inflection point) is detected, and the term
855 "unclear" is used when no definite hockey stick pattern can be observed. For linear responses, a
856 low threshold value of 1 mm (corresponding to the minimum precipitation event size) was used
857 for subsequent correlation analyses.

858

859

	Threshold values (when they exist)			
	Hydrologic response		Hydrologic response and storage deficit	
	Rainfall-triggered (Figure 2)	Snowmelt-driven (Figure 3)	Rainfall-triggered (Figure 4)	Snowmelt-driven (Figure 5)
Dorset	~ 50 mm	~ 170 mm	~ 50 mm	~ 150 mm
Girnock	~ 80 mm	~ 60 mm	~ 60 mm	~ 50 mm
H.J. Andrews	~ 80 mm	Unclear	~ 60 mm	Unclear
Hubbard Brook	~ 90 mm	~ 120 mm	~ 80 mm	~ 110 mm
Krycklan	~ 50 mm	Unclear	~ 40 mm	~ 60 mm
Mharcaidh	Unclear	~ 30 mm	~ 36 mm	~ 25 mm
Sleepers	~ 100 mm	~ 180 mm	~ 80 mm	~ 180 mm
Strontian	Linear (1 mm)	Linear (1 mm)	Linear (1 mm)	Linear (1 mm)
Wolf Creek	Unclear	Unclear	~ 40 mm	Unclear

860

861 **Table 5.** R^2 values above the threshold for the nine North-Watch catchments. Asterisks (*) flag
862 catchments and event types for which no statistically significant R^2 values above the threshold
863 could be computed either due to the absence of a threshold or due to insufficient data (fewer
864 than three event points above the threshold).

865

	R^2 values above the knot			
	Hydrologic response		Hydrologic response and storage deficit	
	Rainfall-triggered	Snowmelt-driven	Rainfall-triggered	Snowmelt-driven
Dorset	0.12	0.15	0.15	0.21
Girnock	0.59	0.39	0.56	0.47
H.J. Andrews	0.66	*	0.70	*
Hubbard Brook	0.54	0.40	0.61	0.48
Krycklan	0.53	*	0.39	0.18
Mharcaidh	*	0.14	0.14	0.13
Sleepers	0.90	0.79	0.65	0.7
Strontian	*	*	*	*
Wolf Creek	*	*	0.81	*

866

867 **Table 6**. Spearman rank correlation ($corr_{catchment}$) between the three metrics of hydrological
868 behaviour and catchment properties. Reported Spearman rank correlation coefficients are
869 significant at the 95% statistical level.
870

	Spearman r				Threshold value				R² above the threshold value			
	Hydrologic response		Hydrologic response and storage deficit		Hydrologic response		Hydrologic response and storage deficit		Hydrologic response		Hydrologic response and storage deficit	
	Rainfall-triggered	Snowmelt-driven	Rainfall-triggered	Snowmelt-driven	Rainfall-triggered	Snowmelt-driven	Rainfall-triggered	Snowmelt-driven	Rainfall-triggered	Snowmelt-driven	Rainfall-triggered	Snowmelt-driven
MeanTemperature					0.66							0.73
MeanPrecipitation												
PrctSnow												
MeanEvaporation												
MeanStorage												
Area												
MeanElevation	0.71					0.66						
Relief			0.67									
BFI												-0.72
FDCS_lowflow									-0.66			
FDCS_intermediateflow									-0.66			
FDCS_highflow												-0.72
ElevationAboveStream												
DistanceFromStream	0.81		0.83									
GradientToStream												
TransitTimeProxy												
D8Gradient												
DinfGradient												
d5												
SubcatchmentArea				0.88			0.81					
UpslopeArea												
TWI												
TWId5												

871

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
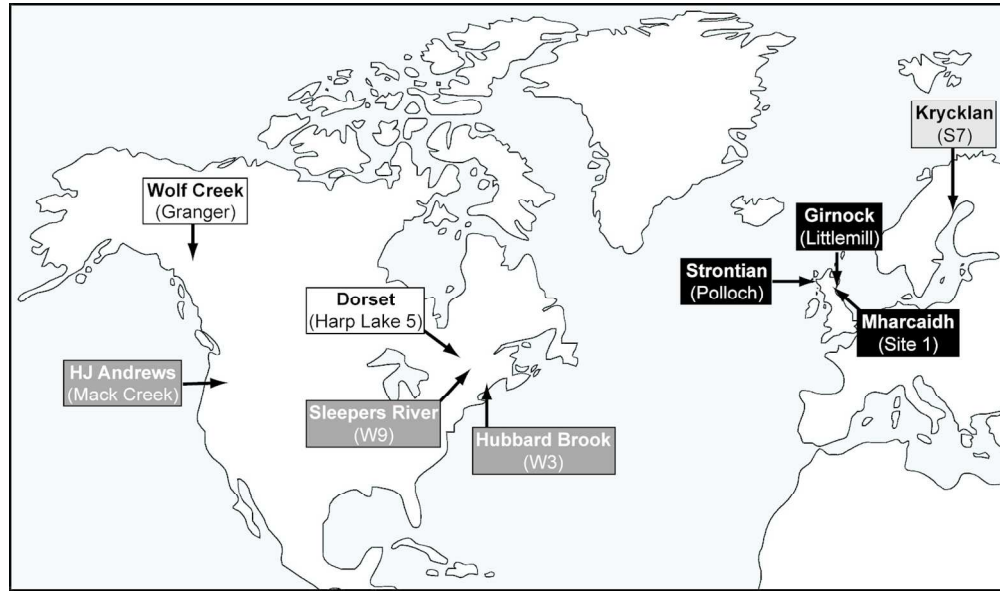46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 1. Location of the nine North-Watch catchments. White, light grey, dark grey and black rectangles signal Canadian, Swedish, US and Scottish catchments, respectively. Catchment names are reported in bold while specific site names (stream gauges) are mentioned in brackets.
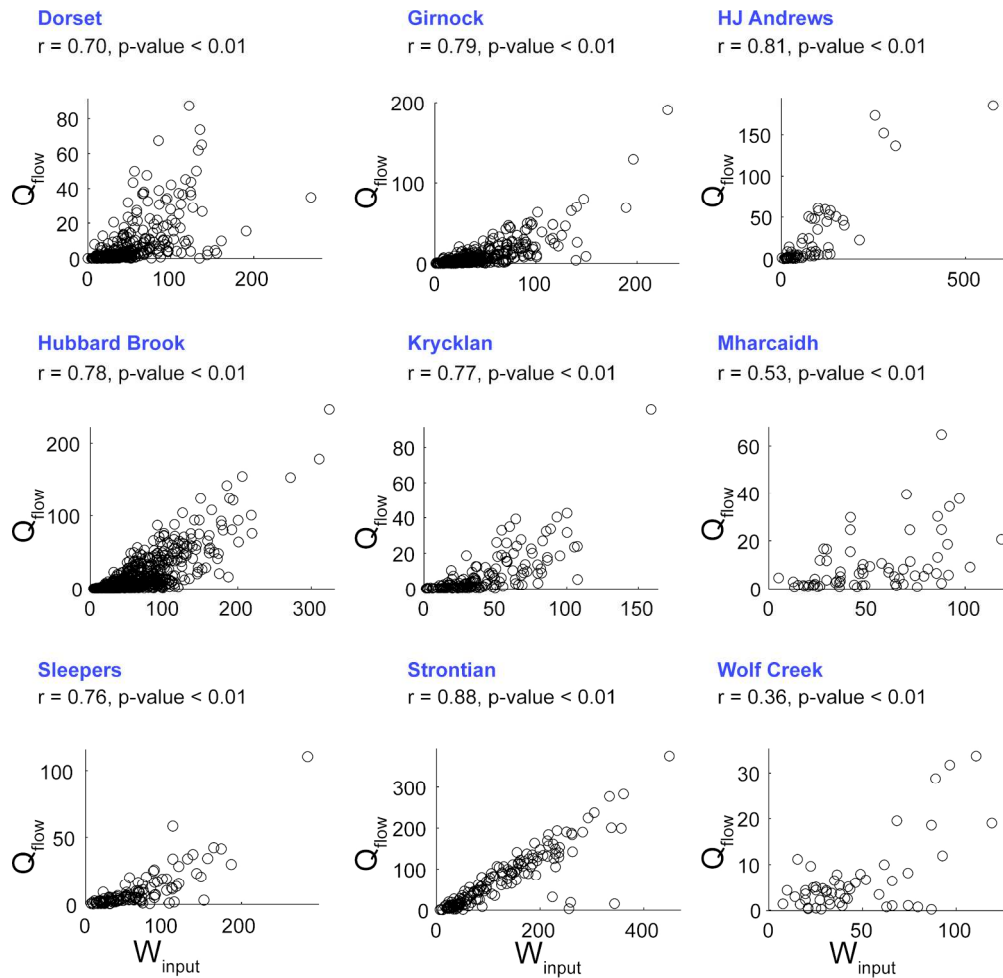132x78mm (300 x 300 DPI)

Figure 2. Total quickflow (Qflow, mm) vs. total water input (Winput, mm) in the nine North-Watch catchments for rainfall-triggered events. The Spearman rank correlation coefficient (rSpearman, abbreviated as "r") and its associated p-value are reported.
187x181mm (300 x 300 DPI)

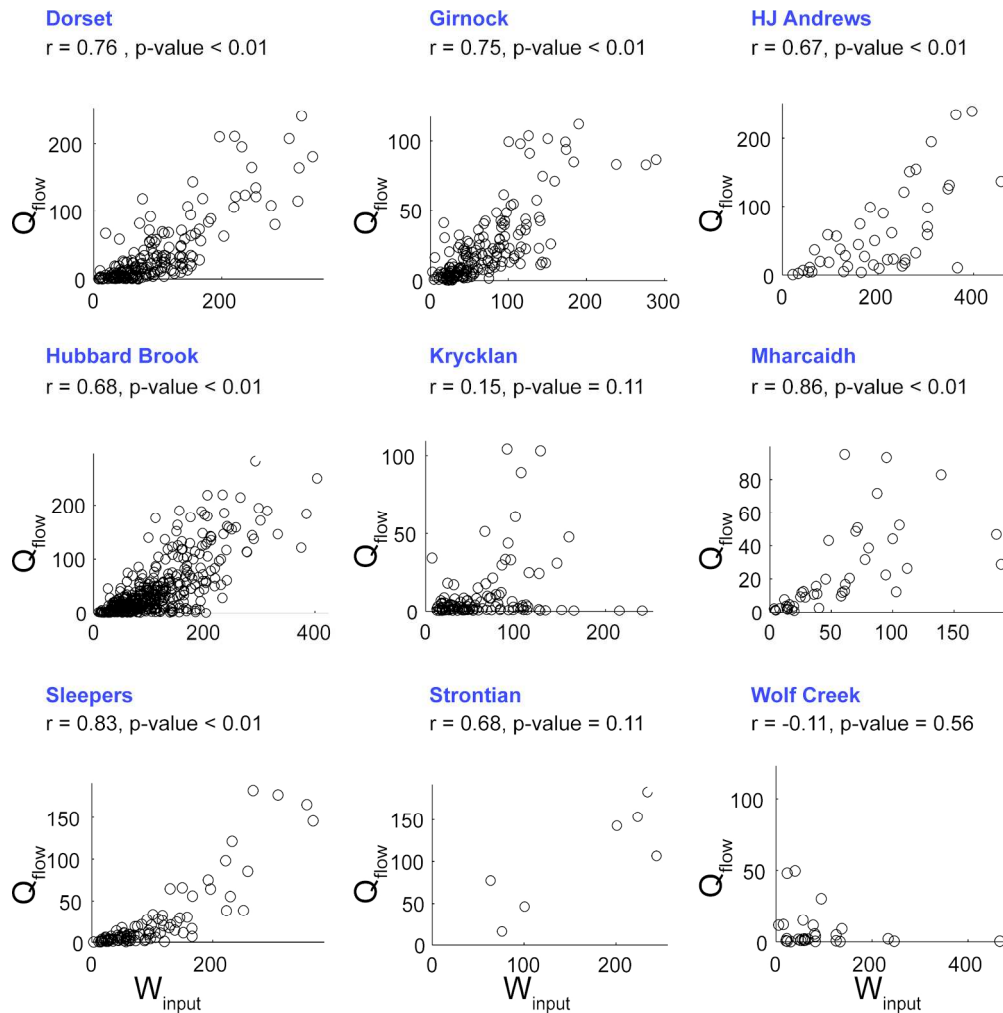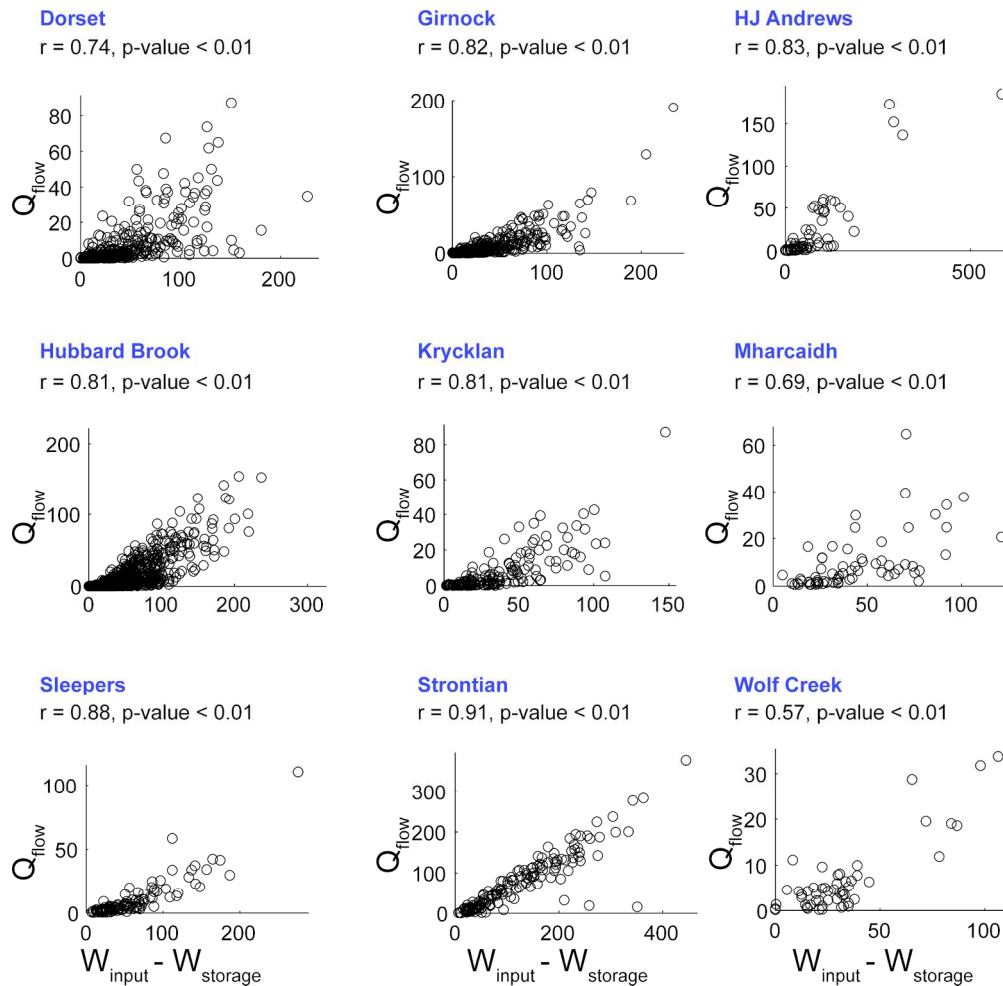Figure 3. Total quickflow (Qflow, mm) vs. total water input (Winput, mm) in the nine North-Watch catchments for snowmelt-driven events. The Spearman rank correlation coefficient (rSpearman, abbreviated as "r") and its associated p-value are reported.
180x181mm (300 x 300 DPI)

Figure 4. Total quickflow (Qflow, mm) vs. effective water input (Winput -Wstorage, mm) in the nine North-Watch catchments for rainfall-triggered events. The Spearman rank correlation coefficient (rSpearman, abbreviated as "r") and its associated p-value are reported.
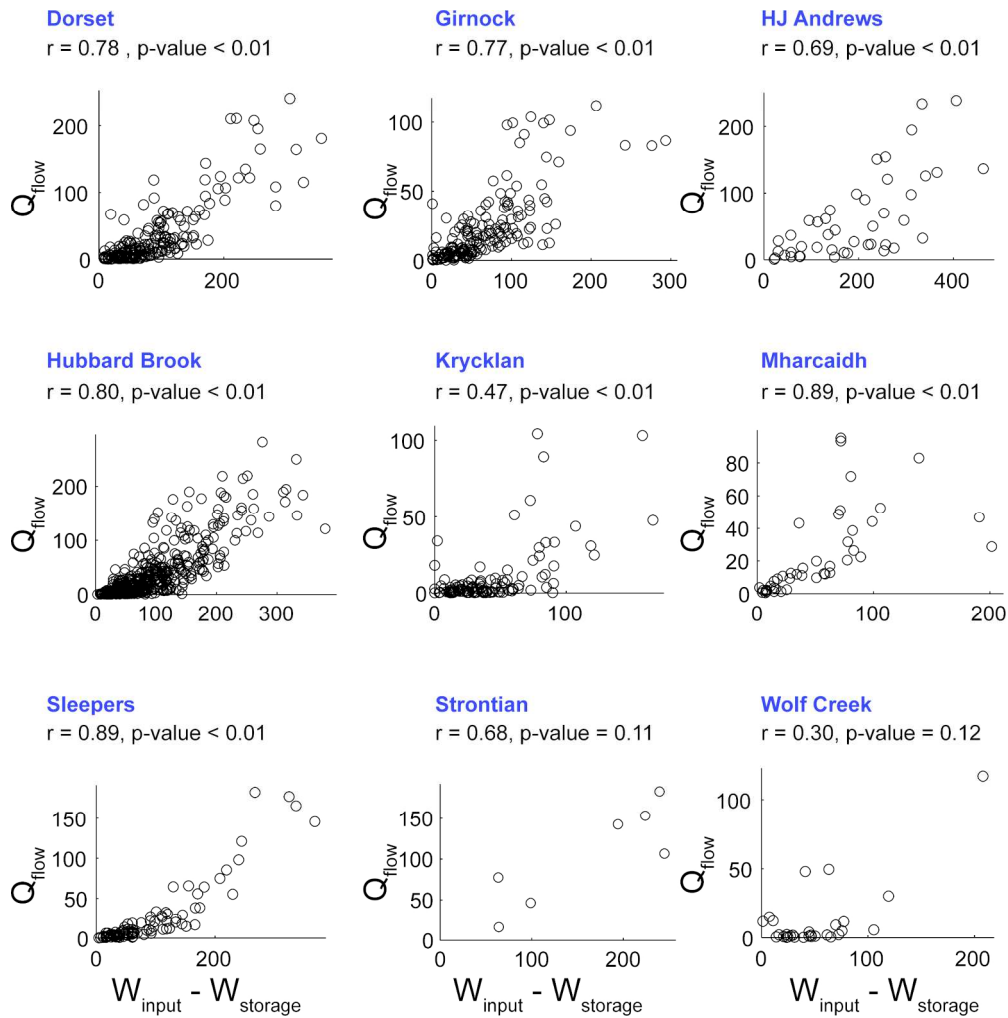183x178mm (300 x 300 DPI)

Figure 5. Total quickflow (Qflow, mm) vs. effective water input (Winput -Wstorage, mm) in the nine North-Watch catchments for snowmelt-driven events. The Spearman rank correlation coefficient (rSpearman, abbreviated as "r") and its associated p-value are reported.
178x178mm (300 x 300 DPI)