1   # Flood risk reduction and flow buffering as ecosystem services:

2   # a flow persistence indicator for watershed health

4   Meine van Noordwijk[1,2], Lisa Tanika[1], Betha Lusiana[1]

5   [1]{World Agroforestry Centre (ICRAF), SE Asia program, Bogor, Indonesia}

6   [2]{Wageningen University, Plant Production Systems, Wageningen, the Netherlands}

7   Correspondence to: Meine van Noordwijk (m.vannoordwijk@cgiar.org)

1 **Abstract**

2 Flood damage depends on location and adaptation of human presence and activity to

3 inherent variability of river flow. Reduced predictability of river flow is a common

4 sign of degrading watersheds associated with increased flooding risk and reduced dry-

5 season flows. The dimensionless FlowPer parameter (F_p), representing predictability,

6 is key to a parsimonious recursive model of river flow, $Q_t = F_p Q_{t-1} + (1-F_p)(P_t - E_{tx})$,

7 with Q, P and E expressed in mm d^{-1}. F_p varies between 0 and 1, and can be derived

8 from a time-series of measured (or modeled) river flow data. The spatially averaged

9 precipitation term P_t and preceding cumulative evapotranspiration since previous rain

10 E_{tx} are treated as constrained but unknown, stochastic variables. A decrease in F_p from

11 0.9 to 0.8 means peak flow doubling from 10 to 20% of peak rainfall (minus its

12 accompanying E_{tx}) and, in a numerical example, an increase in expected flood

13 duration by 3 days. We compared F_p estimates from four meso-scale watersheds in

14 Indonesia and Thailand, with varying climate, geology and land cover history, at a

15 decadal time scale. Wet-season (3-monthly) F_p values are lower than dry-season

16 values in climates with pronounced seasonality. A wet-season F_p value above 0.7 was

17 achievable in forest-agroforestry mosaic case studies. Interannual variability in F_p is

18 large relative to effects of land cover change; multiple years of paired-plot data are

19 needed to reject no-change null-hypotheses. While empirical evidence at scale is

20 understandably scarce, F_p trends over time serve as a holistic scale-dependent

21 performance indicator of degrading/recovering watershed health.

22 **1 Introduction**

23 Degradation of watersheds and its consequences for river flow regime and flooding intensity

24 are a widespread concern (Brauman et al., 2007; Bishop and Pagiola, 2012; Winsemius et al.,

25 2013). Current watershed rehabilitation programs that focus on increasing tree cover in upper

26 watersheds are only partly aligned with current scientific evidence of effects of large-scale

27 tree planting on streamflow (Ghimire et al., 2014; Malmer et al., 2010; Palmer, 2009; van

28 Noordwijk et al., 2007, 2015; Verbist et al 2010). The relationship between floods and change

29 in forest quality and quantity, and the availability of evidence for such a relationship at

30 various scales has been widely discussed over the past decades (Andréassian, 2004;

31 Bruijnzeel, 2004; Bradshaw et al., 2007; van Dijk et al., 2009). The ratio between peak and

32 average flow decreases between from headwater streams to main rivers in a predictable

1 manner; while mean annual discharge scales with $(area)^{1.0}$, maximum river flow scales with

2 $(area)^{0.7}$ on average (Rodríguez-Iturbe and Rinaldo, 2001; van Noordwijk et al., 1998). The

3 determinants of peak flows are thus scale-dependent, with space-time correlations in rainfall

4 interacting with subcatchment-level flow buffering in peakflows at any point along the river.

5 Whether and where peakflows lead to flooding depends on the capacity of the rivers to pass

6 on peakflows towards downstream lakes or the sea, assisted by riparian buffer areas with

7 sufficient storage capacity (Baldasarre et al., 2013). Well-studied effects of forest conversion

8 on peak flows in small upper stream catchments (Alila et al., 2009) do not necessarily

9 translate to flooding downstream. As summarized by Beck et al. (2013) meso- to macroscale

10 catchment studies (>1 and >10 000 km^2, respectively) in the tropics, subtropics, and warm

11 temperate regions have mostly failed to demonstrate a clear relationship between river flow

12 and change in forest area. Lack of evidence cannot be firmly interpreted as evidence for lack

13 of effect, however. A recent econometric study for Peninsular Malaysia by Tan-Soo et al.

14 (2014) concluded that, after appropriate corrections for space-time correlates in the data-set

15 for 31 meso- and macroscale basins (554-28,643 km^2), conversion of inland rain forest to

16 monocultural plantations of oil palm or rubber increased the number of flooding days

17 reported, but not the number of flood events, while conversion of wetland forests to urban

18 areas reduced downstream flood duration. This study may be the first credible empirical

19 evidence at this scale. The difference between results for flood duration and flood frequency

20 and the result for draining wetland forests warrant further scrutiny. Consistency of these

21 findings with river flow models based on a water balance and likely pathways of water under

22 the influence of change in land cover and land use has yet to be shown. Two recent studies for

23 Southern China confirm the conventional perspective that deforestation increases high flows,

24 but are contrasting in effects of reforestation. Zhou et al. (2010) analyzed a 50-year data set

25 for Guangdong Province in China and concluded that forest recovery had not changed the

26 annual water yield (or its underpinning water balance terms precipitation and

27 evapotranspiration), but had a statistically significant positive effect on dry season (low)

28 flows. Liu et al. (2015), however, found for the Meijiang watershed (6983 km2) in

29 subtropical China that while historical deforestation had decreased the magnitudes of low

30 flows (daily flows $\leqq$ Q95%) by 30.1%, low flows were not significantly improved by

31 reforestation. They concluded that recovery of low flows by reforestation may take much

32 longer time than expected probably because of severe soil erosion and resultant loss of soil

33 infiltration capacity after deforestation.

1   The statistical challenges of attribution of cause and effect in such data-sets are considerable

2   with land use/land cover interacting with spatially and temporally variable rainfall, geological

3   configuration and the fact that land use is not changing in random fashion or following any

4   pre-randomized design (Alila et al., 2009; Rudel et al., 2005). Hydrologic analysis across 12

5   catchments in Puerto Rico by Beck et al. (2013) did not find significant relationships between

6   the change in forest cover or urban area, and change in various flow characteristics, despite

7   indications that regrowing forests increased evapotranspiration. Yet, the concept of a

8   'regulating function' on river flow regime for forests and other semi-natural ecosystems is

9   widespread. The considerable human and economic costs of flooding at locations and times

10 beyond where this is expected make the presumed 'regulating function' on flood reduction of

11 high value (Brauman et al., 2007) – if only we could be sure that the effect is real, beyond the

12 local scales (< 10 km^2) of paired catchments where ample direct empirical proof exists

13 (Bruijnzeel, 1990, 2004). Here we will explore a simple recursive model of river flow (van

14 Noordwijk et al., 2011) that (i) is focused on (loss of) predictability, (ii) can account for the

15 types of results obtained by the cited recent Malaysian study (Tan-Soo et al., 2014), and (iii)

16 may constitute a suitable performance indicator of watershed 'health' through time,

17 combining statistical properties of the local rainfall regime, land cover effects on soil structure

18 and any engineering modifications of water flow (Ma et al., 2014).

19   ⇨ Fig. 1

20 Figure 1 is compatible with a common dissection of risk as the product of hazard, exposure

21 and vulnerability. Extreme discharge events plus river-level engineering co-determine hazard,

22 while exposure depends on topographic position interacting with human presence, and

23 vulnerability can be modified by engineering at a finer scale. A recent study (Jongman et al.,

24 2015) found that human fatalities and material losses between 1980 and 2010 expressed as a

25 share of the exposed population and gross domestic product were decreasing with rising

26 income. Yet, the planning needed to avoid extensive damage requires quantification of the

27 risk of higher than usual discharges, especially at the upper tail end of the flow frequency

28 distribution.

29 The statistical scarcity of 'extreme events' and the challenge of data collection where they do

30 occur, make it hard to rely on empirical data as such. Existing data on flood frequency and

31 duration, as well as human and economic damage are influenced by topography, human

32 population density and economic activity, interacting with engineered infrastructure (steps 5-9

33 in Fig. 1), as well as the extreme rainfall events that are their proximate cause. Common

1   hydrological analysis of flood frequency (called 1 in 10-, 1 in 100-, 1 in 1000-year flood

2   events, for example) doess not separately attribute flood magnitude to rainfall and land use

3   properties, and analysis of likely change in flood frequencies in the context of climate change

4   adaptation has been challenging (Milly et al., 2002; Ma et al., 2014). There is a lack of simple

5   performance indicators for watershed health   (step 3 in Fig. 1) that align with local

6   observations of river behavior and concerns about its change and that can reconcile local,

7   public/policy and scientific knowledge, thereby helping negotiated change in watershed

8   management (Leimona et al., 2015). The behavior of rivers depends on many climatic (step 4

9   in Figure 1) and terrain factors (step 1 in Figure 1) that make it a challenge to differentiate

10   between anthropogenically induced ecosystem structural and soil degradation (step 0) and

11   intrinsic variability (Fig. 1). Hydrologic models tend to focus on predicting hydrographs and

12   are usually tested on data-sets from limited locations. Despite many decades of hydrologic

13   modeling, current hydrologic theory, models and empirical methods have been found to be

14   largely inadequate for sound predictions in ungauged basins (Hrachowitz et al., 2013). Efforts

15   to resolve this through harmonization of modelling strategies have so far failed. Existing

16   models differ in the number of explanatory variables and parameters they use, but are

17   generally dependent on empirical data of rainfall that are available for specific measurement

18   points but not at the spatial resolution that is required for a close match between measured and

19   modeled river flow. Spatially explicit models have conceptual appeal (Ma et al., 2010) but

20   have too many degrees of freedom and too many opportunities for getting right answers for

21   wrong reasons if used for empirical calibration (Beven, 2011). Parsimonious, parameter-

22   sparse models are appropriate for the level of evidence available to constrain them, but these

23   parameters are themselves implicitly influenced by many aspects of existing and changing

24   features of the watershed, making it hard to use such models for scenario studies of

25   interacting land use and climate change. Here we present a more direct approach deriving a

26   metric of flow predictability that can bridge local concerns and concepts to quantified

27   hydrologic function: the 'flow persistence' parameter (step 3 in Figure 1).

28   In this contribution to the debate on forests and floods we will first define the metric 'flow

29   persistence' in the context of temporal autocorrelation of river flow and derive a way to

30   estimate its numerical value. We will then apply the algorithm to river flow data for a number

31   of contrasting meso-scale watersheds, representing variation in rainfall and land cover, and

32   and test the internal consistency of results based on historical data: one located in the humid

33   tropics of Indonesia, and one in the unimodal subhumid tropics of northern Thailand. As a

1 next step we show how projected changes in rainfall patterns (frequency, intensity, temporal

2 and spatial autocorrelation) are expected to interact with changes in land cover, soil

3 infiltration behaviour and landscape-level buffering elements such as wetlands and

4 impoundments, on the regularity of river flow, as captured by the flow persistence metric.

5 Possible applications of the flow persistence metric to questions on low flows are left for a

6 later analysis. In the discussion we will consider the new flow persistence metric in terms of

7 three groups of criteria (Clark et al., 2011; Lusiana et al., 2011; Leimona et al., 2015) based

8 on salience (1,2), credibility (3,4) and legitimacy (5-7):

9   1. Does flow persistence relate to important aspects of watershed behavior?

10   2. Does it's quantification help to select management actions?

11   3. Is there consistency of numerical results?

12   4. How sensitive is it to noise in data sources?

13   5. Does it match local knowledge?

14   6. Can it be used to empower local stakeholders of watershed management?

15   7. Can it inform local risk management?

16 **2 Flow persistence as a suitable hydrological metric: theory**

17 2.1 Basic equations

18 One of the easiest-to-observe aspects of a river is its day-to-day fluctuation in waterlevel,

19 related to the volumetric flow (discharge) via rating curves (Maidment, 1992). Without

20 knowing details of upstream rainfall and the pathways the rain takes to reach the river,

21 observation of the daily fluctuations in waterlevel allows important inferences to be made. It

22 is also of direct utility: sudden rises can lead to floods without sufficient warning, while rapid

23 decline makes water utilization difficult. Indeed, a common local description of watershed

24 degradation is that rivers become more 'flashy' and less predictable, having lost a buffer or

25 'sponge' effect (Joshi et al., 2004; Ranieri et al., 2004; Rahayu et al., 2013). The probably

26 simplest model of river flow at time t, Q_t, is that it is similar to that of the day before (Q_{t-1}), to

27 the degree F_p, a dimensionless parameter called 'flow persistence' (van Noordwijk et al.,

28 2011) plus an additional stochastic term ε:

1    $Q_t = F_p Q_{t-1} + \varepsilon$ [1].

2   Q_t is for this analysis expressd in mm d^{-1}, which means that measurements in m^3 s^{-1} need to be

3   divided by the relevant catchment area, with appropriate unit conversion. If river flow were

4   constant, it would be perfectly predictable, i.e. F_p would be 1.0 and ε zero; in contrast, an F_p-

5   value equal to zero and ε directly reflecting erratic rainfall represents the lowest possible

6   predictability.

7   The F_p parameter is conceptually identical to the 'recession constant' commonly used in

8   hydrological models, typically assessed during an extended dry period when the ε term is

9   negligible and streamflow consists of baseflow only (Tallaksen, 1995); empirical deviations

10   from a straight line in a plot of the logarithm of Q against time are common and point to

11   multiple rather than a single groundwater pool that contributes to base flow. With increasing

12   size of a catchment area it is increasingly likely that there indeed are multiple, partly

13   independent groundwater contributions.

14   As we will demonstrate, it is possible to derive F_p even when ε is not negligible. In climates

15   without distinct dry season this is essential; elsewhere it allows a comparison of apparent F_p

16   between wet and dry parts of the hydrologic year. A decrease over the years of F_p indicates

17   'watershed degradation' (i.e. greater contrast between high and low flows), and an increase

18   'improvement' or 'rehabilitation' (i.e. more stable flows).

19   If we consider the sum of river flow over a sufficiently long period, we can expect ΣQ_t to

20   closely approximate ΣQ_{t-1}, and thus

21   $\Sigma Q_t = F_p \Sigma Q_{t-1} + \Sigma \varepsilon$ [2].

22   From this relationship we obtain a first way of estimating the F_p value if a complete

23   hydrograph is available:

24   $F_p = 1 - \Sigma \varepsilon / \Sigma Q_t$ [3].

25   Rearranging Eq.(3) we obtain

26   $\Sigma \varepsilon = (1 - F_p) \Sigma Q_t$ [4].

27   The F_p term is equivalent with one of several ways to separate baseflow from peakflows. The

28   $\Sigma \varepsilon$ term reflects the sum of peak flows in mm, while $F_p \Sigma Q_t$ reflects the sum of base flow, also

29   in mm. For $F_p = 1$ (the theoretical maximum) we conclude that all ε must be zero, and all flow

30   is 'base flow'. The stochastic ε can be interpreted in terms of what hydrologists call 'effective

1   rainfall' (i.e. rainfall minus on-site evapotranspiration, assessed over a preceding time period

2   tx since previous rain event):

3   $Q_t = F_p\, Q_{t-1} + (1-F_p)(P_{tx} - E_{tx})$ [5].

4   Where P_{tx} is the (spatially weighted) precipitation (assuming no snow or ice) in mm d^{-1} and

5   E_{tx}, also in mm d^{-1}, is the preceding evapotranspiration that allowed for infiltration during this

6   rainfall event (*i.e.* evapotranspiration since the previous soil-replenishing rainfall that induced

7   empty pore space in the soil for infiltration and retention). More complex attributions are

8   possible, aligning with the groundwater replenishing bypassflow and the water isotopic

9   fractionation involved in evaporation (Evaristo et al., 2015).

10 The multiplication of effective rainfall times $(1-F_p)$ can be checked by considering the

11 geometric series $(1-F_p)$, $(1-F_p)\, F_p$, $(1-F_p)\, F_p^2$, …, $(1-F_p)\, F_p^n$ which adds up to $(1-F_p)(1 - F_p^n)/(1-$

12 $F_p)$ or $1 - F_p^n$. This approaches 1 for large n, suggesting that all of the water attributed to time

13 t, *i.e.* $P_t - E_{tx}$, will eventually emerge as river flow. For $F_p = 0$ all of $(P_t - E_{tx})$ emerges on the

14 first day, and riverflow is as unpredictable as precipitation itself. For $F_p = 1$ all of $(P_t - E_{tx})$

15 contributes to the stable daily flow rate. For declining F_p, $(1 > F_p > 0)$, river flow gradually

16 becomes less predictable, because a greater part of the stochastic precipitation term

17 contributes to variable rather than evened-out river flow.

18 Taking long term summations of the right- and left- hand sides of Eq.(5) we obtain:

19 $\Sigma Q_t = \Sigma(F_p\, Q_{t-1} + (1-F_p)(P_t - E_{tx})) = F_p\, \Sigma\, Q_{t-1} + (1-F_p)(\,\Sigma\, P_t - \Sigma\, E_{tx}))$ [6].

20 Which is consistent with the basic water budget, $\Sigma Q = \Sigma P - \Sigma E$, at time scales that changes in

21 soil water buffer stocks can be ignored. As such the total annual, and hence the mean daily

22 river flow are independent of F_p. This does not preclude that processes of watershed

23 degradation or restoration that affect the partitioning of P over Q and E also affect F_p.

24 **2.2 Low flows**

25 The lowest flow expected in an annual cycle is $Q_x\, F_p^{Nmax}$ where Q_x is flow on the first day

26 without rain and N_{max} the longest series of dry days. Taken at face value, a decrease in F_p has

27 a strong effect on low-flows, with a flow of 10% of Q_x reached after 45, 22, 14, 10, 8 and 6

28 days for $F_p = 0.95$, 0.9, 0.85, 0.8, 0.75 and 0.7, respectively. However, the groundwater

29 reservoir that is drained, equalling the cumulative dry season flow if the dry period is

30 sufficiently long, is $Q_x/(1-F_p)$. If F_p decreases to F_{px} but the groundwater reservoir (Res =

1 $Q_x/(1-F_p)$) is not affected, initial flows in the dry period will be higher ($Q_x F_{px}^i (1-F_{px})$ Res $>$

2 $Q_x F_p^i (1-F_p)$ Res for $i < \log((1-F_{px})/(1-F_p))/\log(F_p/F_{px})$). It thus matters how low flows are

3 evaluated: from the perspective of the lowest level reached, or as cumulative flow. The

4 combination of climate, geology and land form are the primary determinants of cumulative

5 low flows, but if land cover reduces the recharge of groundwater there may be impacts on dry

6 season flow, that are not directly reflected in Fp.

7 If a single F_p value would account for both dry and wet season, the effects of changing F_p on

8 low flows may well be more pronounced than those on flood risk. Tests are needed of the

9 dependence of F_p on Q (see below). Analysis of the way an aggregate F_p depends on the

10 dominant flow pathways provides a basis for differentiating F_p within a hydrologic year.

12 **2.3 Flow-pathway dependent flow persistence**

13 A further interpretation of Eq.(1) can be that three pathways of water through a landscape

14 contribute to river flow (Barnes, 1939): groundwater release with $F_{p,g}$ values close to 1.0,

15 overland flow with $F_{p,o}$ values close to 0, and interflow with intermediate $F_{p,i}$ values.

16 $Q_t = F_{p,g} Q_{t-1,g} + F_{p,i} Q_{t-1,i} + F_{p,o} Q_{t-1,o} + \varepsilon$ [7],

17 $F_p = (F_{p,g} Q_{t-1,g} + F_{p,i} Q_{t-1,i} + F_{p,o} Q_{t-1,o})/Q_{t-1}$ [8].

18 On this basis a decline or increase in overall weighted average F_p can be interpreted as

19 indicator of a shift of dominant runoff pathways through time within the watershed.

20 Similarly, a second interpretation of F_p emerges based on the fractions of total river flow that

21 are based on groundwater, overland flow and interflow pathways:

22 $F_p = F_{p,g} (\Sigma Q_{t,g} / \Sigma Q_t) + F_{p,o} (\Sigma Q_{t,o} / \Sigma Q_t) + F_{p,i} (\Sigma Q_{t,i} / \Sigma Q_t)$ [9].

23 Beyond the type of degradation of the watershed that, mostly through soil compaction, leads

24 to enhanced infiltration-excess (or Hortonian) overland flow (Delfs et al., 2009), saturated

25 conditions throughout the soil profile may also induce overland flow, especially near valley

26 bottoms (Bonell, 1993; Bruijnzeel, 2004). Thus, the value of $F_{p,o}$ can be substantially above

27 zero if the rainfall has a significant temporal autocorrelation, with heavy rainfall on

28 subsequent days being more likely than would be expected from general rainfall frequencies.

29 If rainfall following a wet day is more likely to occur than following a dry day, as is

30 commonly observed in Markov chain analysis of rainfall patterns (Jones and Thornton, 1997;

31 Bardossy and Plate, 1991), the overland flow component of total flow will also have a partial

1   temporal autocorrelation, adding to the overall predictability of river flow. In a hypothetical

2   climate with evenly distributed rainfall, we can expect F_p to be 1.0 even if there is no

3   infiltration and the only pathway available is overland flow. Even with rainfall that is variable

4   at any point of observation but has low spatial correlation it is possible to obtain F_p values of

5   (close to) 1.0 in a situation with (mostly) overland flow (Ranieri at al., 2004).

6   **2.4 Numerical example**

7   Figure 2 provides an example of the way a change in F_p values (based on Eq. 1) influences the

8   visual pattern of river flow for a unimodal rainfall regime with a well-developed dry season.

9   The increasing 'spikedness' of the graph as F_p is lowered indicates reduced predictability of

10   flow on any given day during the wet season on the basis of the flow on the preceding day. A

11   bi-plot of river flow on subsequent days for the same simulations (Fig. 3) shows two main

12   effects of reducing the F_p value: the scatter increases, and the slope of the lower envelope

13   containing the swarm of points is lowered (as it equals F_p). Both of these changes can provide

14   entry points for an algorithm to estimate F_p from empirical time series, provided the basic

15   assumptions of the simple model apply and the data are of acceptable quality (see Section 3

16   below). For the numerical example shown in Fig. 2, the maximum daily flow doubled from 50

17   to 100 mm when the F_p value decreased from a value close to 1 (0.98) to nearly 0.

18   ⇨ Fig. 2

19   ⇨ Fig. 3

20   **2.5 Flow persistence as a simple flood risk indicator**

21   For numerical examples (implemented in a spreadsheet model) flow on each day can be

22   derived as:

23   $Q_t = \Sigma_j^t F_p^{t-j} (1-F_p)\, p_j\, P_j$ [10].

24   Where p_j reflects the occurrence of rain on day j (reflecting a truncated sine distribution for

25   seasonal trends) and P_j is the rain depth (drawn from a uniform distribution). From this model

26   the effects of F_p (and hence of changes in F_p) on maximum daily flow rates, plus maximum

27   flow totals assessed over a 2-5 d period, was obtained in a Monte Carlo process (without

28   Markov autocorrelation of rainfall in the default case – see below). Relative flood protection

29   was calculated as the difference between peak flows (assessed for 1-5 d duration after a 1 year

30   'warm-up' period) for a given F_p versus those for $F_p = 0$, relative to those at $F_p = 0$.

1 In further analyzing this numerical example, we evaluated the maximum flow by

2 accumulating over a 1-5 d period (in a moving average routine) and compared the maximum

3 obtained for each F_p with what, for the same Monte Carlo realization, was obtained for F_p of

4 zero. This way a relative flood protection, expressed as reduction of peak flow, could be

5 related to F_p (Fig. 4A). Relative flood protection decreased to less than 10% at F_p values of

6 around 0.5, with slightly weaker flood protection when the assessment period was increased

7 from 1 to 5 days (between 1 and 3 d it decreased by 6.2%, from 3 to 5 d by a further 1.3%).

8 Two counteracting effects are at play here: a lower F_p means that a larger fraction $(1-F_p)$ of

9 the effective rainfall contributes to river flow, but the increased flow is less persistent. In the

10 example the flood protection in situations where the rainfall during 1 or 2 d causes the peak is

11 slightly stronger than where the cumulative rainfall over 3-5 d causes floods, as typically

12 occurs downstream.

13 As we expect peak flow to be proportional to $(1-F_p)$ times peak rainfall amounts, the effect of

14 a change in F_p not only depends on the change in F_p that we are considering, but also on its

15 initial value, with greater F_p values leading to more rapid increases in high flows (Fig. 4B).

16 However, flood duration rather responds to changes in F_p in a curvilinear manner, as flow

17 persistence implies flood persistence (once flooding occurs), but the greater the flow

18 persistence the less likely such a flooding threshold is passed (Fig. 4C). The combined effect

19 may be restricted to about 3 d of increase in flood duration for the parameter values used in

20 the default example, but for different parametrization of the stochastic ε other results might be

21 obtained.

22   ⇨ Fig. 4

23 3 Methods

24 3.1 An algorithm for deriving F_p from a time series of stream flow data

25 Equation (3) provides a first method to derive F_p from empirical data if these cover a full

26 hydrologic year. In situations where there is no complete hydrograph and/or in situations

27 where we want to quantify F_p for shorter time periods (e.g. to characterise intraseasonal flow

28 patterns) and the change in the storage term of the water budget equation cannot be ignored,

29 we need an algorithm for estimating F_p from a series of daily Q_t observations.

30 Where rainfall has clear seasonality, it is attractive and indeed common practice to derive a

31 groundwater recession rate from a semi-logarithmic plot of Q against time (Tallaksen, 1995).

32 As we can assume for such periods that ε = 0, we obtain $F_p = Q_t / Q_{t-1}$, under these

1 circumstances. We cannot be sure, however, that this $F_{p,g}$ estimate also applies in the rainy

2 season, because overall wet-season F_p will include contributions by $F_{p,o}$ and $F_{p,i}$ as well

3 (compare Eq. 9). In locations without a distinct dry season, we need an alternative method.

4 A biplot of Q_t against Q_{t-1} (as in Fig. 3) during times of flow recession will lead to a scatter of

5 points above a line with slope F_p, with points above the line reflecting the contributions of ε

6 >0, while the points that plot on the F_p line itself represent $\varepsilon = 0$ mm d^{-1}. There is no

7 independent source of information on the frequency at which $\varepsilon = 0$, nor what the statistical

8 distribution of ε values is if it is non-zero. Calculating back from the Q_t series we can obtain

9 an estimate of Q_{add} as the realization of the stochastic ε for any given estimate of F_p, and

10 select the most plausible value. For high F_p estimates there will be many negative Q_{add} values,

11 for low F_p estimates all Q_{add} values will be larger. An algorithm to derive a plausible F_p

12 estimate can thus make use of the corresponding distribution of 'apparent Q_{add}' values as

13 estimates of ε ($Q_{add} = Q_t - F_p Q_{t-1}$). While ε, and thus in theory Q_{add} cannot be negative, small

14 negative Q_{add} estimates are likely when using real-world data with their inherent errors. The

15 FlowPer F_p algorithm (van Noordwijk et al., 2011) derives the distribution of $Q_{add,Fptry}$

16 estimates for a range of $F_{p,try}$ values (Fig. 5B) and selects the value $F_{p,try}$ that minimizes the

17 variance Var($Q_{add,Fptry}$) (or its standard deviation) (Fig. 5C). It is implemented in a spreadsheet

18 workbook that can be downloaded from the ICRAF website (****).

19 ➜Fig. 5

20 A consistency test is needed that the high-end Q_t values relate to Q_{t+1} in the same was as do

21 low or medium Q_t values. Visual inspection of Q_{t+1} versus Q_t, with the derived F_p value,

22 provides a qualitative view of the validity of this assumption.

3.2 GenRiver model for effects of land cover on river flow

24 The GenRiver model (van Noordwijk et al., 2011) is based on a simple water balance concept

25 with a daily timestep and a flexible spatial subdivision of a watershed that influences the

26 routing of water and employs spatially explicit rainfall. Land cover affects rainfall

27 interception losses as well as soil macroporosity (bulk density) modifying infiltration rates.

28 Any land-cover change scenarios are interpolated annually between measured time-series

29 data. The model may use measured rainfall data, or use a rainfall generator that involves

30 Markov chain temporal autocorrelation (rain persistence). The model itself, a manual and

31 application case studies are freely available (**weblink**; van Noordwijk et al., 2011).

3.3 Empirical data-sets

Table 1 provides summary characteristics of four meso-scale watersheds used for testing the F_p algorithm and application of the GenRiver model. Basic site-specific parameterization is given in Table 2 and land-cover specific default parameters in Table 3, while Table 4 describes the six scenarios of land-use change that were evaluated in terms of their hydrological impacts.

⇨ Table 1
⇨ Table 2
⇨ Table 3
⇨ Table 4

3.4 Bootstrapping

We used a bootstrap approach to estimate the minimum number of observation (or yearly data) required for a pair-wise comparison test between two time-series of stream flow data (representing 2 scenarios of land use) to be distinguishable from a null-hypothesis of no effect. We built a simple macro in R (R Core Team, 2015) using the following steps:

(i)   Take a sample of size n from both time-series data with replacement, N times,

(ii) Apply the Kolmogorov-Smirnov test, and record the P-value,

(iii) Perform (i) and (ii) for different size of n

(iv) Tabulate the p-value from various n, and determine the value of n when the p-value reached equal to or less than 0.025. The associated n represents the minimum number of observations required. Appendix 1 provides an example of the macro in R.

4   Results

4.1 Empirical data of flow persistence as basis for model parameterization

Overall the estimates from modeled and observed data are related with 16% deviating more than 0.1 and 3% more than 0.15. The flow persistence estimates derived from the wettest three-month period are about 0.2 lower than those derived for the driest period, when baseflow dominates (Fig. 6). If we can expect $F_{p,i}$ and $F_{p,o}$ to be approximately 0.5 and 0, this difference between wet and dry periods implies a 40% contribution of interflow in the wet season, a 20% contribution of overland flow or any combination of the two effects.

⇨ Fig. 6

4.2 F_p effects for scenarios of land cover change

1   ⇨ Fig. 7

2   Among the four watersheds there is consistency in that the 'forest' scenario has the highest,

3   and the 'degraded lands' the lowest F_p value (Fig. 7), but there are remarkable differences as

4   well: in Cidanau the interannual variation in F_p is clearly larger than land cover effects, while

5   in the Way Besai the spread in land use scenarios is larger than interannual variability. In

6   Cidanau a peat swamp between most of the catchment and measuring point buffers most of

7   landcover related variation in flow, but not the interannual variability. Considering the

8   frequency distributions of F_p values over a 20 year period, we see one watershed (Way Besai)

9   where the forest stands out from all others, and one (Bialo) where the degraded lands are

10   separate from the others. Given the degree of overlap of the frequency distributions, it is clear

11   that multiple years of empirical observations will be needed before a change can be affirmed.

12   Figure 8 shows the frequency distributions of expected effect sizes on Fp of a comparison of

13   any land cover with either forest or degraded lands. Table 5 translates this information to the

14   number of years that a paired plot (in the absence of measurement error) would have to be

15   maintained to reject a null-hypothesis of no effect, at p=0.05. As the frequency distributions

16   of F_p differences of paired catchments do not match a normal distribution, a Kolmorov-

17   Smirnov test can be used to assess the probability that a no-difference null hypothesis can

18   yield the difference found. By bootstrapping within the years where simulations supported by

19   observed rainfall data exist, we found for the Way Besai catchment, for example, that 20

20   years of data would be needed to assert (at $P = 0.05$) that the ReFor scenario differs from

21   AgFor, and 16 years that it differs from Actual and 11 years that it differs from Degrade. In

22   practice, that means that empirical evidence that survives statistical tests will not emerge,

23   even though effects on watershed health are real.

24   ⇨ Fig. 8
25   ⇨ Table 5

26   At process-level the increase in 'overland flow' in response to soil compaction due to land

27   cover change has a clear and statistically significant relationship with decreasing Fp values in

28   all catchments (Fig 8A), but both year-to-year variation within a catchment and differences

29   between catchments influence the results as well, leading to considerable spread in the biplot.

30   Contrary to expectations, the disappearance of 'interflow' by soil compaction is not reflected

31   in measurable change in Fp value. The temporal difference between overland and interflow

32   (one or a few days) gets easily blurred in the river response that integrates over multiple

33   streams with variation in delivery times; the difference between overland- or interflow and

1 baseflow is much more pronounced. Apparently, according to our model, the high

2 macroporosity of forest soils that allows interflow and may be the 'sponge' effect attributed to

3 forest, delays delivery to rivers by one or a few days, with little effect on the flow volumes at

4 locations downstream where flow of multiple days accumulates. The difference between

5 overland- or interflow and baseflow in time-to-river of rainfall peaks is much more

6 pronounced..

7   ⇨ Fig. 9

8 Tree cover has two contradicting effects on baseflow: it reduces the surplus of rainfall over

9 evapotranspiration (annual water yield) by increased evapotranspiration (especially where

10 evergreen trees are involved), but it potentially increases soil macroporosity that supports

11 infiltration and interflow, with relatively little effect on waterholding capacity measured as

12 'field capacity' (after runoff and interflow have removed excess water). Fig. 6 shows that the

13 total volume of baseflow differs more between sites and their rainfall pattern than it varies

14 with tree cover. Between years total evapotranspiration and baseflow totals are positively

15 correlated (see supplementary information), but for a given rainfall there is a tradeoff. Overall

16 these results support the conclusion that generic effects of deforestation on decreased flow

17 persistence, and of (agro)/(re)-forestation on increased flow persistence are small relative to

18 interannual variability due to specific rainfall patterns, and that it will be hard for any

19 empirical data process to pick-up such effects, even if they are qualitatively aligned with valid

20 process-based models.

21 **5 Discussion**

22 In view of our results the lack of robust evidence in the literature of effects of change in forest

23 and tree cover on flood occurrence may not be a surprise; effects are subtle and most data sets

24 contain considerable noise. Yet, such effects are consistent with current process and scaling

25 knowledge of watersheds. The key strength of our flow persistence parameter, that it can be

26 derived from observing river flow at a single point along the river, without knowledge of

27 rainfall events and catchment conditions, is also its major weakness. If rainfall data exist, and

28 especially rainfall data that apply to each subcatchment, the Q_{add} term doesn't have to be

29 treated as a random variable and event-specific information on the flow pathways may be

30 inferred for a more precise account of the hydrograph. But for the vast majority of rivers in

31 the tropics, advances in remotely sensed rainfall data are needed to achieve that situation and

32 F_p may be all that is available to inform public debates on the relation between forests and

1   floods. We will discuss the flow persistence metric against criteria based on salience,

2   credibility and legitimacy. Key *salience* aspects are "Does flow persistence relate to important

3   aspects of watershed behavior?" and "Does it help to select management actions?". Figures 2

4   and 6 show that most of the effects of a decreasing F_p value on peak discharge (which is the

5   basis for downstream flooding) occur between F_p values of 1 and 0.7, with the relative flood

6   protection value reduced to 10% when F_p reaches 0.5. As indicated in Fig. 1, peak discharge

7   is only one of the factors contributing to flood risk in terms of human casualties and physical

8   damage. The F_p value has an inverse effect on the fraction of recent rainfall that becomes river

9   flow, but the effect on peak flows is less, as higher F_p values imply higher base flow. The way

10   these counteracting effects balance out depends on details of the local rainfall pattern

11   (including its Markov chain temporal autocorrelation), as well as the downstream topography

12   and risk of people being at the wrong time at a given place, but the F_p value is en efficient

13   way of summarizing complex land use mosaics and upstream topography in its effect on river

14   flow. The difference between wet-season and dry-season F_p deserves further analysis. In

15   climates with a real rainless dry-season, dry season F_p is dominated by the groundwater

16   release fraction of the watershed, regardless of land cover, while in wet season it depends on

17   the mix (weighted average) of flow pathways. The degree to which F_p can be influenced by

18   land cover needs to be assessed for each landscape and land cover combination, including the

19   locally relevant forest and forest derived land classes, with their effects on interception, soil

20   infiltration and time pattern of transpiration. The F_p value can summarize results of models

21   that explore land use change scenarios in local context. To select the specific management

22   actions that will maintain or increase F_p a locally calibrated land use/hydrology model is

23   needed, such as GenRiver or SWAT (Yen et al., 2015). The empirical data summarized here

24   for (sub)humid tropical sites in Indonesia and Thailand show that values of F_p above 0.9 are

25   scarce in the case studies provided, but values above 0.8 were found, or inferred by the model,

26   for forested landscapes. Agroforestry landscapes generally presented F_p values above 0.7,

27   while open-field agriculture or degraded soils led to F_p values of 0.5 or lower. Despite

28   differences in local context, it seems feasible to relate typical F_p values to the overall

29   condition of a watershed.

30   Key *credibility* questions are "Consistency of numerical results?" and "How sensitive are

31   results to noisy data sources?". Intra-annual variability of F_p values was around 0.2 in our

32   results, interannual variability in either annual or seasonal F_p was generally in the 0.1 range,

33   while the difference between observed and simulated flow data as basis for F_p calculations

1   was mostly less than 0.1. With current methods, it seems that effects of land cover change on

2   flow persistence that shift the F_p value by about 0.1 are the limit of what can be asserted from

3   empirical data (with shifts of that order in a single year a warning sign rather than a firmly

4   established change). When derived from observed river flow data F_p is suitable for monitoring

5   change (degradation, restoration) and can be a serious candidate for monitoring performance

6   in outcome-based ecosystem service management contracts. Where further uncertainty is

7   introduced by the use of modeled rather than measured river flow, the lack of fit of models

8   similar to the ones we used here would mean that scenario results are indicative of directions

9   of change rather than a precision tool for fine-tuning combinations of engineering and land

10 cover change as part of integrated watershed management.

11 *Legitimacy* aspects are "Does it match local knowledge?" and "Can it be used to empower

12 local stakeholders of watershed management?" and "Can it inform risk management?". As the

13 F_p parameter captures the predictability of river flow that is a key aspect of degradation

14 according to local knowledge systems, its results are much easier to convey than full

15 hydrographs or excedance probabilities of flood levels. By focusing on observable effects at

16 river level, rather than prescriptive recipes for land cover ("reforestation"), the F_p parameter

17 can be used to more effectively compare the combined effects of land cover change, changes

18 in the riparian wetlands and engineered water storage reservoirs, in their effect on flow

19 buffering. It is a candidate for shifting environmental service reward contracts from input to

20 outcome based monitoring (van Noordwijk et al., 2012). As such it can be used as part of a

21 negotiation support approach to natural resources management in which leveling off on

22 knowledge and joint fact finding in blame attribution are key steps to negotiated solutions that

23 are legitimate and seen to be so (van Noordwijk et al., 2013; Leimona et al., 2015).

24 Quantification of F_p can help assess tactical management options (Burt et al., 2014) as in a

25 recent suggestion to minimize negative downstream impacts of forestry operations on stream

26 flow by avoiding land clearing and planting operations in locally wet La Niña years. But the

27 most challenging aspect of the management of flood, as any other environmental risk, is that

28 the frequency of disasters is too low to intuitively influence human behavior where short-term

29 risk taking benefits are attractive. Wider social pressure is needed for investment in watershed

30 health (as a type of insurance premium) to be mainstreamed, as individuals waiting to see

31 evidence of necessity are too late to respond. In terms of flooding risk, actions to restore or

32 retain watershed health can be similarly justified as insurance premium. It remains to be seen

33 whether or not the transparency of the F_p metric and its intuitive appeal are sufficient to make

1 the case in public debate when opportunity costs of foregoing reductions in flow buffering by

2 profitable land use are to be compensated and shared (Burt et al., 2014).

3 In conclusion, the F_p metric allows efficient summaries of complex landscape processes into a

4 single parameter that summarizes the effects of landscape management. It integrates changes

5 in tree cover (deforestation, reforestation, agroforestation) at the level that these influence

6 river flow. Flow persistence is the result of rainfall persistence and the temporal delay

7 provided by the pathway water takes through the soil and the river system. High flow

8 persistence indicates a reliable water supply, while minimizing peak flow events. Wider tests

9 of the F_p metric as boundary object in science-practice-policy boundary chains (Kirchoff et al

10 2015; Leimona et al., 2015) are needed.

Data availability

12 Table 6 specifies the rainfall and river flow data we used for the four basins and specifies the

13 links to detailed descriptions.

14 ⇨ Table 6

Author contributions

17 MvN designed method and paper, LT handled the case study data and modeling and BL

18 contributed statistical analysis; all contributed and approved the final manuscript

Acknowledgements

20 This research is part of the Forests, Trees and Agroforestry research program of the CGIAR.

21 Several colleagues contributed to the development and early tests of the F_p method. Thanks

22 are due to Eike Luedeling, Sonya Dewi and Sampurno Bruijnzeel for comments on an earlier

23 version of the manuscript.

References

26 Alila, Y., Kura, P.K., Schnorbus, M., and Hudson, R.: Forests and floods: A new paradigm

27 sheds light on age-old controversies. Water Resour Res 45: W08416, 2009.

28 Andréassian, V.: Waters and forests: from historical controversy to scientific debate. Journal

29 of Hydrology 291: 1–27, 2004.

1 Baldassarre, G.D., Kooy, M., Kemerink, J.S. and Brandimarte, L.,: Towards understanding
2   the dynamic behaviour of floodplains as human-water systems. *Hydrology and Earth*
3   *System Sciences 17*(8), 3235-3244, 2013.

4 Bardossy, A. and Plate, E.J.: Modeling daily rainfall using a semi-Markov representation of
5   circulation pattern occurrence. *Journal of Hydrology*, *122*(1), pp.33-47, 1991.

6 Barnes, B.S.: The structure of discharge-recession curves. *Eos, Transactions American*
7   *Geophysical Union*, *20*(4), 721-725, 1939.

8 Beck, H. E., Bruijnzeel, L. A., Van Dijk, A. I. J. M., McVicar, T. R., Scatena, F. N., and
9   Schellekens, J.: The impact of forest regeneration on streamflow in 12 mesoscale humid
10   tropical catchments. *Hydrology and Earth System Sciences 17*(7), 2613-2635, 2013.

11 Beven, K.J.,: *Rainfall-runoff modelling: the primer*. John Wiley & Sons, 2011

12 Bishop, J. and Pagiola, S. (Eds.): Selling forest environmental services: market-based
13   mechanisms for conservation and development. Taylor & Francis, Abingdon (UK), 2012.

14 Bonell, M.: Progress in the understanding of runoff generation dynamics in forests. *Journal of*
15   *Hydrology*, *150*(2), pp.217-275, 1993.

16 Bradshaw, C.J.A., Sodhi, N.S., Peh, K.S.H., and Brook, B.W.: Global evidence that
17   deforestation amplifies flood risk and severity in the developing world. Global Change
18   Biol. 13: 2379–2395, 2007.

19 Brauman, K.A., Daily, G.C., Duarte, T.K.E., and Mooney, H.A.: The nature and value of
20   ecosystem services: an overview highlighting hydrologic services. Annu. Rev. Environ.
21   Resour. 32: 67-98, 2007.

22 Bruijnzeel, L.A.: Hydrological functions of tropical forests: not seeing the soil for the trees.
23   Agr. Ecosyst. Environ. 104: 185–228, 2004.

24 Burt, T.P., Howden, N.J.K., McDonnell, J.J., Jones, J.A., Hancock, G.R.: Seeing the climate
25   through the trees: observing climate and forestry impacts on streamflow using a 60-year
26   record. Hydrological Processes. DOI: 10.1002/hyp.10406, 2014.

27 Clark, W. C., Tomich, T. P., van Noordwijk, M., Guston, D., Catacutan, D., Dickson, N. M.,
28   and McNie, E.: Boundary work for sustainable development: natural resource
29   management at the Consultative Group on International Agricultural Research (CGIAR).
30   Proc. Nat. Acad. Sci., doi:10.1073/pnas.0900231108, 2011.

1 Delfs, J.O., Park, C.H., and Kolditz, O.: A sensitivity analysis of Hortonian flow. Advances in
2   Water Resources 32(9): 1386-1395, 2009.

3 Evaristo, J., Jasechko, S. and McDonnell, J.J.,: Global separation of plant transpiration from
4   groundwater and streamflow. *Nature*, *525*(7567), pp.91-94, 2015.

5 Ghimire, C.P., Bruijnzeel, L.A., Lubczynski, M.W., and Bonell, M.: Negative trade-off
6   between changes in vegetation water use and infiltration recovery after reforesting
7   degraded pasture land in the Nepalese Lesser Himalaya. Hydrology and Earth System
8   Sciences 18(12): 4933-4949, 2014.

9 Hrachowitz, M., Savenije, H.H.G., Blöschl, G., McDonnell, J.J., Sivapalan, M., Pomeroy,
10   J.W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J E., Gelfan, A.,
11   Gupta, H.V., Hughes, D. A., Hut, R.W., Montanari, A., Pande, S., Tetzlaff, D., Troch,
12   P.A., Uhlenbrook, S., Wagener, T., Winsemius, H.C., Woods, R.A., Zehe, E. and
13   Cudennec, C.: A decade of Predictions in Ungauged Basins (PUB)—a
14   review. Hydrological sciences journal, 58(6), 1198-1255, 2013.

15 Jones, P.G. and Thornton, P.K., 1997. Spatial and temporal variability of rainfall related to a
16   third-order Markov model. *Agricultural and Forest Meteorology*, *86*(1), pp.127-138.

17 Jongman, B., Winsemius, H. C., Aerts, J. C., de Perez, E. C., van Aalst, M. K., Kron, W., and
18   Ward, P. J.: Declining vulnerability to river floods and the global benefits of
19   adaptation. Proceedings of the National Academy of Sciences, 112(18), E2271-E2280,
20   2015.

21 Joshi, L., Schalenbourg, W., Johansson, L., Khasanah, N., Stefanus, E., Fagerström, M.H. and
22   van Noordwijk, M.,: Soil and water movement: combining local ecological knowledge
23   with that of modellers when scaling up from plot to landscape level. In: van Noordwijk,
24   M., Cadisch, G. and Ong, C.K. (Eds.*) Belowground Interactions in Tropical
25   Agroecosystems*, CAB International, Wallingford (UK). pp. 349-364, 2004.

26 Kirchhoff, C.J., Esselman, R., and Brown, D.: Boundary Organizations to Boundary Chains:
27   Prospects for Advancing Climate Science Application. Climate Risk Management
28   doi:10.1016/j.crm.2015.04.001, 2015.

29 Leimona, B., Lusiana, B., van Noordwijk, M., Mulyoutami, E., Ekadinata, A., and
30   Amaruzama, S.: Boundary work: knowledge co-production for negotiating payment for
31   watershed services in Indonesia. Ecosystems Services 15, 45–62, 2015.

1 Liu, W., Wei, X., Fan, H., Guo, X., Liu, Y., Zhang, M. and Li, Q.,: Response of flow regimes
2   to deforestation and reforestation in a rain-dominated large watershed of subtropical
3   China. *Hydrological Processes* 10459, 2015

4 Lusiana, B., van Noordwijk, M., Suyamto, D., Joshi, L., and Cadisch, G.: Users' perspectives
5   on validity of a simulation model for natural resource management. International Journal
6   of Agricultural Sustainability 9(2): 364-378, 2011.

7 Ma, X., Xu, J., van Noordwijk, M.: Sensitivity of streamflow from a Himalayan catchment to
8   plausible changes in land-cover and climate. Hydrological Processes 24: 1379–1390,
9   2010.

10 Ma, X., Lu, X., van Noordwijk, M., Li, J.T., and Xu, J.C.: Attribution of climate change,
11   vegetation restoration, and engineering measures to the reduction of suspended sediment
12   in the Kejie catchment, southwest China. Hydrol. Earth Syst. Sci. 18: 1979–1994, 2014.

13 Maidment, D.R.: *Handbook of hydrology*. McGraw-Hill Inc., 1992.

14 Malmer, A., Murdiyarso, D., Bruijnzeel L.A., and Ilstedt, U.: Carbon sequestration in tropical
15   forests and water: a critical look at the basis for commonly used generalizations. *Global*
16   *Change Biology 16*(2): 599-604, 2010.

17 Milly, P.C.D., Wetherald, R., Dunne, K.A., and Delworth, T.L.: Increasing risk of great
18   floods in a changing climate. Nature 415(6871): 514-517, 2002.

19 Palmer, M.A.: Reforming watershed restoration: science in need of application and
20   applications in need of science. Estuaries and coasts 32(1): 1-17, 2009.

21 R Core Team R: A language and environment for statistical computing. R Foundation for
22   Statistical Computing, Vienna, Austria. URL http://www.R-project.org/, 2015

23 Rahayu, S., Widodo, R.H., van Noordwijk, M., Suryadi, I. and Verbist, B.: *Water monitoring*
24   *in watersheds*. Bogor, Indonesia. World Agroforestry Centre (ICRAF) SEA Regional
25   Program. 104 p., 2013

26 Ranieri, S.B.L., Stirzaker, R., Suprayogo, D., Purwanto, E., de Willigen, P. and van
27   Noordwijk, M.,: Managing movements of water, solutes and soil: from plot to landscape
28   scale. In: van Noordwijk, M., Cadisch, G. and Ong, C.K. (Eds.) *Belowground*
29   *Interactions in Tropical Agroecosystems*, CAB International, Wallingford (UK). pp. 329-
30   347, 2004.

1 Rodríguez-Iturbe, I., and Rinaldo, A.: Fractal river basins: chance and self-organization.
2   Cambridge University Press, Cambridge, 2001.

3 Rudel, T. K., Coomes, O. T., Moran, E., Achard, F., Angelsen, A., Xu, J., and Lambin, E.:
4   Forest transitions: towards a global understanding of land use change. Global
5   Environmental Change 15(1): 23-31, 2005.

6 Tallaksen, L.M.: A review of baseflow recession analysis. J Hydrol 165, 349-370, 1995.

7 Tan-Soo, J.S., Adnan, N., Ahmad, I., Pattanayak, S.K., and Vincent, J.R.: Econometric
8   Evidence on Forest Ecosystem Services: Deforestation and Flooding in
9   Malaysia. Environmental and Resource Economics, on-line:
10   http://link.springer.com/article/10.1007/s10640-014-9834-4, 2014.

11 van Dijk, A.I., van Noordwijk, M., Calder, I.R., Bruijnzeel, L.A., Schellekens, J., and
12   Chappell, N.A.: Forest-flood relation still tenuous – comment on 'Global evidence that
13   deforestation amplifies flood risk and severity in the developing world'. Global Change
14   Biology 15: 110-115, 2009.

15 van Noordwijk, M., van Roode, M., McCallie, E.L. and Lusiana, B.,: Erosion and
16   sedimentation as multiscale, fractal processes: implications for models, experiments and
17   the real world. In: F. Penning de Vries, F. Agus and J. Kerr (Eds.) Soil Erosion at Multiple
18   Scales, Principles and Methods for Assessing Causes and Impacts.. CAB International,
19   Wallingford. pp 223-253, 1998.

20 van Noordwijk, M., Agus, F., Verbist, B., Hairiah, K. and Tomich, T.P.: Managing Watershed
21   Services. In: Scherr, S.J. and McNeely, J.A. (Eds) Ecoagriculture Landscapes. Farming
22   with Nature: The Science and Practice of Ecoagriculture. Island Press, Washington DC,
23   pp 191 – 212, 2007.

24 van Noordwijk, M., Widodo, R.H., Farida, A., Suyamto, D., Lusiana, B., Tanika, L., and
25   Khasanah, N..: GenRiver and FlowPer: Generic River and Flow Persistence Models. User
26   Manual Version 2.0. Bogor, Indonesia: World Agroforestry Centre (ICRAF) Southeast
27   Asia Regional Program. 119 p, 2011.

28 van Noordwijk, M., Leimona, B., Jindal, R., Villamor, G.B., Vardhan, M., Namirembe, S.,
29   Catacutan, D., Kerr, J., Minang, P.A., and Tomich, T.P.: Payments for Environmental
30   Services: evolution towards efficient and fair incentives for multifunctional landscapes.
31   Annu. Rev. Environ. Resour. 37: 389-420, 2012.

1 van Noordwijk, M., Lusiana, B., Leimona, B., Dewi, S. and Wulandari, D.: Negotiation-
2   support toolkit for learning landscapes. Bogor, Indonesia. World Agroforestry Centre
3   (ICRAF) Southeast Asia Regional Program, 2013.

4 van Noordwijk, M., Leimona, B., Xing, M., Tanika, L., Namirembe, S., and Suprayogo, D.,:
5   Water-focused landscape management. Climate-Smart Landscapes: Multifunctionality In
6   Practice. eds Minang PA et al.. Nairobi, Kenya: World Agroforestry Centre (ICRAF), pp
7   179-192, 2015.

8 Verbist, B., Poesen, J., van Noordwijk, M. Widianto, Suprayogo, D., Agus, F., and Deckers,
9   J.: Factors affecting soil loss at plot scale and sediment yield at catchment scale in a
10   tropical volcanic agroforestry landscape. Catena 80: 34-46, 2010.

11 Winsemius, H.C., van   Beek, L.P.H., Jongman, B., Ward, P.J., and   Bouwman, A.:   A
12   framework for global river flood risk assessments. Hydrol Earth Syst Sci 17:1871–1892,
13   2013.

14 Yen, H., White, M.J., Jeong, J., Arabi, M. and Arnold, J.G.: Evaluation of alternative surface
15   runoff accounting procedures using the SWAT model.International Journal of Agricultural
16   and Biological Engineering, 8(1), 2015.

17 Zhou, G., Wei, X., Luo, Y., Zhang, M., Li, Y., Qiao, Y., Liu, H. and Wang, C.,: Forest
18   recovery and river discharge at the regional scale of Guangdong Province, China. *Water
19   Resources Research*, *46*(9) 008829, 2010.

1   Table 1. Basic physiographic characteristics of the four study watersheds

Parameter	Bialo	Cidanau	Mae Chaem	Way Besai
Location	South Sulawesi, Indonesia	West Java, Indonesia	Northern Thailand	Lampung, Sumatera, Indonesia
Area (km^2)	111.7	241.6	3891.7	414.4
Elevation (m a.s.l.)	0 - 2874	30 – 1778	475-2560	720-1831
Flow pattern	Parallel	Parallel (with two main river flow that meet in the downstream area)	Parallel	Radial
Dominant land cover type	Mixed garden (cocoa and clove)	Mixed garden	Forest (evergreen, deciduous and pine)	Coffee (monoculture and multistrata)
Mean annual rainfall, mm	1695	2573	1027	2474
Mean annual runoff, mm	947	917	259	1673
Major soils	Inceptisols	Inceptisols	Ultisols, Entisols	Andisols
% Natural forest	13	3.1 (forest and swamp forest)	84 (deciduous, evergreen, pine)	3.6

1   Table 2. Parameters of the GenRiver model used for the four site specific simulations (van

2   Noordwijk et al., 2011 for definitions of terms; sequence of parameters follows the pathway

3   of water)

Parameter	Definition	Unit	Bialo	Cidanau	Mae Chaem	Way Besai
RainIntensMean	Average rainfall intensity	mm d^{-1}	30	30	3	30
RainIntensCoefVar	Coefficient of variation of rainfall intensity	mm d^{-1}	0.8	0.3	0.5	0.3
RainInterceptDripRt	Max rain interception drip rate	mm d^{-1}	80	10	10	10
RainMaxIntDripDur	Rain interception drip duration	hr	0.8	0.5	0.5	0.5
InterceptEffectontrans	Rain interception effect on transpiration	-	0.35	0.8	0.3	0.8
MaxInfRate	Maximum infiltration capacity	mm d^{-1}	580	800	150	720
MaxInfSubsoil	Maximum infiltration sub soil capacity	mm d^{-1}	80	120	150	120
PerFracMultiplier	Daily soil water drainage as fraction of groundwater release fraction	-	0.35	0.13	0.1	0.1
MaxDynGrWatStore	Dynamic groundwater storage capacity	mm	100	100	300	300
GWReleaseFracVar	Groundwater release fraction, applied to all subcatchments	-	0.15	0.03	0.05	0.1
Tortuosity	Stream shape factor	-	0.4	0.4	0.6	0.45
Dispersal Factor	Drainage density	-	0.3	0.4	0.3	0.45
River Velocity	River flow velocity	m s^{-1}	0.4	0.7	0.35	0.5

2   Table 3. GenRiver defaults for land-use specific parameter values, used for all four

3   watersheds (BD/BDref indicates the bulk density relative to that for an agricultural soil

4   pedotransfer function; see van Noordwijk et al., 2011)

Land cover Type	Potential interception (mm/d)	Relative drought threshold	BD/BDref
Forest[1]	3.0 - 4.0	0.4 - 0.5	0.8 - 1.1
Agroforestry[2]	2.0 - 3.0	0.5 - 0.6	0.95 - 1.05
Monoculture tree[3]	1.0	0.55	1.08
Annual crops	1.0 - 3.0	0.6 - 0.7	1.1 - 1.5
Horticulture	1.0	0.7	1.07
Rice field[4]	1.0 - 3.0	0.9	1.1 - 1.2
Settlement	0.05	0.01	1.3
Shrub and grass	2.0 - 3.0	0.6	1.0 - 1.07
Cleared land	1.0 - 1.5	0.3 - 0.4	1.1 - 1.2

6   Note: 1. Forest: primary forest, secondary forest, swamp forest, evergreen forest, deciduous forest

7   2. Agroforestry: mixed garden, coffee, cocoa, clove

8   3. Monoculture : coffee

9   4. Rice field: irrigation and rainfed

1   Table 4. Land use scenarios explored for four watersheds

Scenario	Description
NatFor	Full natural forest, hypothetical reference scenario
ReFor	Reforestation, replanting shrub, cleared land, grass land and some agricultural area with forest
AgFor	Agroforestry scenario, maintaining agroforestry areas and converting shrub, cleared land, grass land and some of agricultural area into agroforestry
Actual	Baseline scenario, based on the actual condition of land cover change during the modeled time period
Agric	Agriculture scenario, converting some of tree based plantations, cleared land, shrub and grass land into rice fields or dry land agriculture, while maintain existing forest
Degrade	No change in already degraded areas, while converting most of forest and agroforestry area into rice fields and dry land agriculture

1   Table 5. Number of years of observations on flow persistence required to reject the null-

2   hypothesis of 'no land use effect' at p-value = 0.05 using Kolmogorov-Smirnov test. The

3   probability of the test statistic in the first significant number is provided between brackets and

4   where the number of observations exceeds the time series available, results are given in *italics*

A. Natural Forest as reference

Way Besai (N=32)

	ReFor	AgFor	Actual	Agric
ReFor		20 (0.035)	16 (0.037)	13 (0.046)
AgFor			n.s.	n.s.
Actual				n.s.
Agric				
Degrade				

Bialo (N=18)

	ReFor	AgFor	Actual	Agric
ReFor		n.s.	n.s.	*37 (0.04)*
AgFor			n.s.	n.s.
Actual				n.s.
Agric				
Degrade				

Cidanau (N=20)

	ReFor	AgFor	Actual	Agric
ReFor		n.s.	n.s.	*32 (0.037)*
AgFor			n.s.	n.s.
Actual				n.s.
Agric				
Degrade				

Mae Chaem (N=15)

	ReFor	Actual	Agric	Degrade
ReFor		n.s.	*23 (0.049)*	*18 (0.050)*
Actual			*45 (0.037)*	*33 (0.041)*
Agric				*33 (0.041)*
Degrade				

B. Degraded scenario as reference

Way Besai (N=32)

	NatFor	ReFor	AgFor	Actual	Agric
NatFor		n.s.	17 (0.042)	13 (0.046)	7 (0.023)
ReFor			21 (0.037)	19 (0.026)	7 (0.023)
AgFor				n.s.	28 (0.046)
Actual					30 (0.029)
Agric					

Bialo (N=18)

	NatFor	ReFor	AgFor	Actual	Agric
NatFor		n.s.	n.s.	*41 (0.047)*	*19 (0.026)*
ReFor			n.s.	n.s.	*32 (0.037)*
AgFor				n.s.	n.s.
Actual					n.s.
Agric					

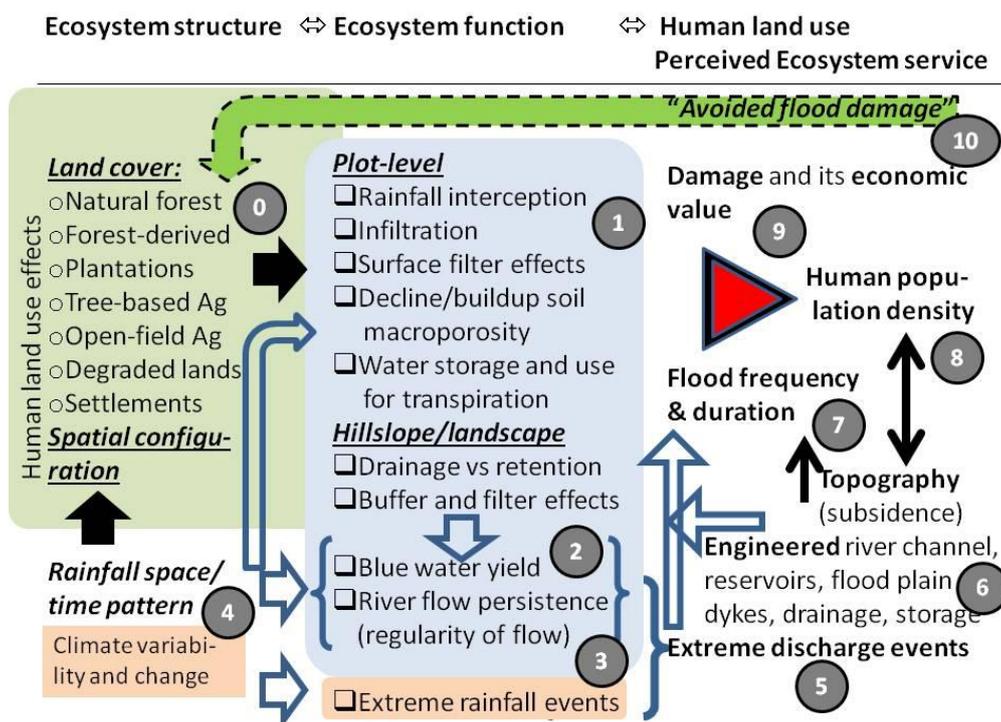
Cidanau (N=20)

	NatFor	ReFor	AgFor	Actual	Agric
NatFor		n.s.	n.s.	*33 (0.041)*	8 (0.034)
ReFor			n.s.	n.s.	15 (0.028)
AgFor				n.s.	n.s.
Actual					*25 (0.031)*
Agric					

Mae Chaem (N=15)

	NatFor	ReFor	Actual	Agric
NatFor		n.s.	25 (0.031)	12 (0.037)
ReFor			n.s.	*18 (0.050)*
Actual				*18 (0.050)*
Agric				

1   Table 6.5. Data availability

	Bialo	Cidanau	Mae Chaem	Way Besai
Rainfall data	1989-2009, Source: BWS Sulawesi and PUSAIR; Average rainfall data from the stations Moti, Bulo-bulo, Seka and Onto	1998-2008, source: BMKG	1998-2002, source: WRD55, MTD22, RYP48, GMT13, WRD 52	1976-2007, Source: BMKG, PU and PLN (interpolation of 8 rainfall stations using Thiessen polygon)
River flow data	1993-2010, source; BWS Sulawesi and PUSAIR	2000-2009, source: KTI	1954-2003, source: ICHARM	1976-1998, source: PU and PUSAIR
Reference of detailed report	Bialo	http://worldagroforestry.org/regions/southeast_asia/publications?do=view_pub_detail&pub_no=PO0292-13	http://worldagroforestry.org/regions/southeast_asia/publications?do=view_pub_detail&pub_no=MN0048-11	http://worldagroforestry.org/regions/southeast_asia/publications?do=view_pub_detail&pub_no=MN0048-11

2   Figure 1. Steps in a causal pathway that relates ecosystem structure to function, human land

3   use and a perceived ecosystem service of 'avoided flood damage'; blue (open) arrows refer

4   to water flow, black (solid) arrows to influences; plot-level processing of incoming rainfall

5   (1) influences the total blue-water yield (2) and its temporal pattern (3), in dependence of

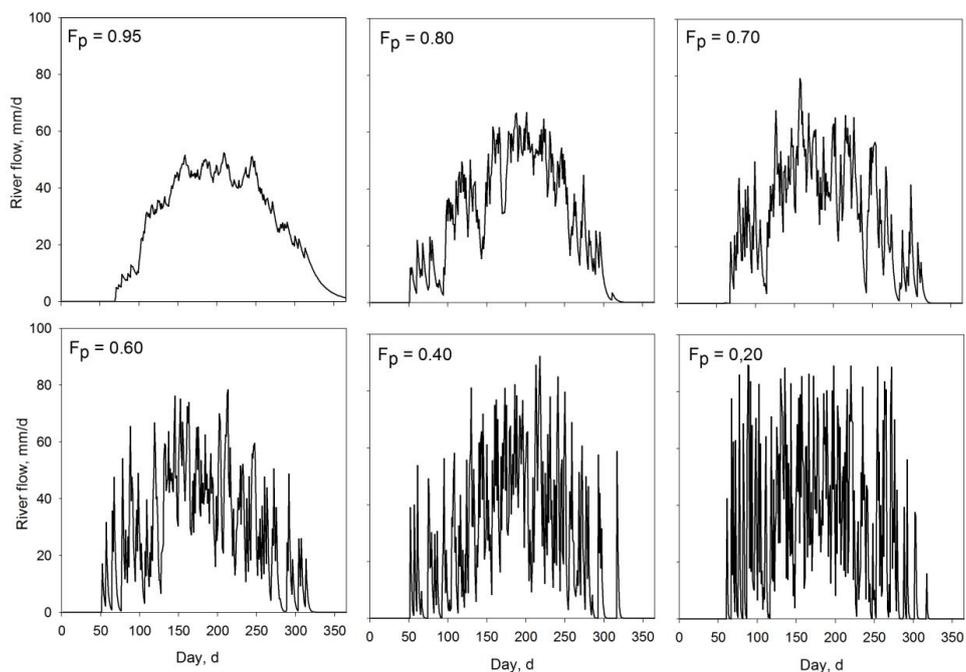
6   the time-space pattern of rainfall (4); extreme discharge events (5), jointly with the

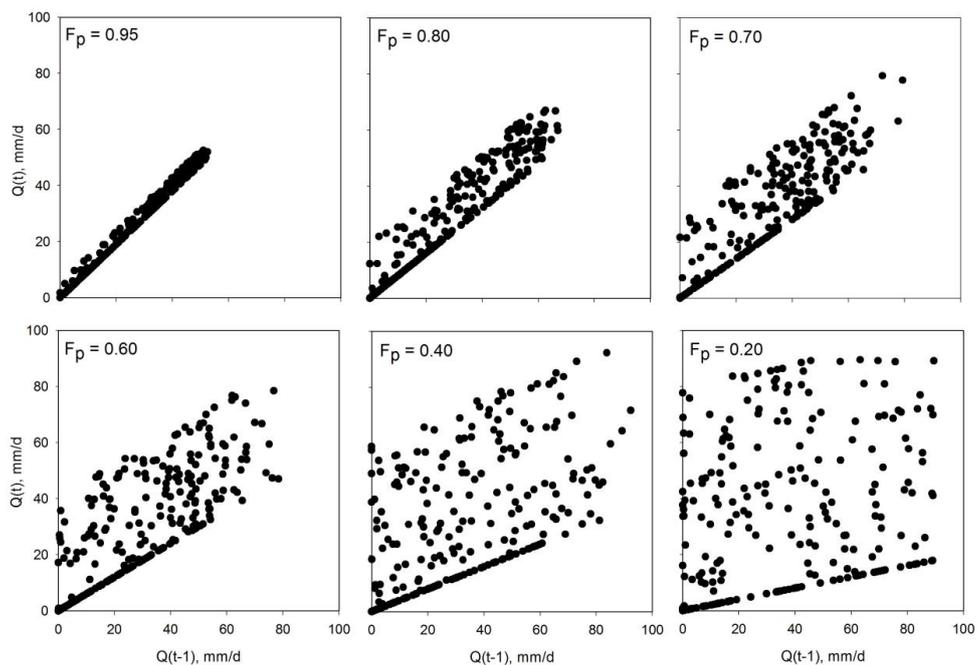
7   (engineered) river channel (6), and topography determine flood frequency and duration (7);

8   human population density and activity (8) together with flood characteristics determines

9   victims, damage and its economic consequence (9); attributing 'avoided flood damage'

10   (10) to land cover (0) and its influences on step 1 is thus complex, especially as *ceteris*

11   *paribus* assumptions do not generally hold and interactions are common

2   Figure 2. Example of daily river flow for a unimodal rainfall regime with clear dry season, in

3   response to change in the flow persistence parameter F_p

2   Figure 3. Biplots of Q(t) versus Q(t-1) for the same simulations as figure 2

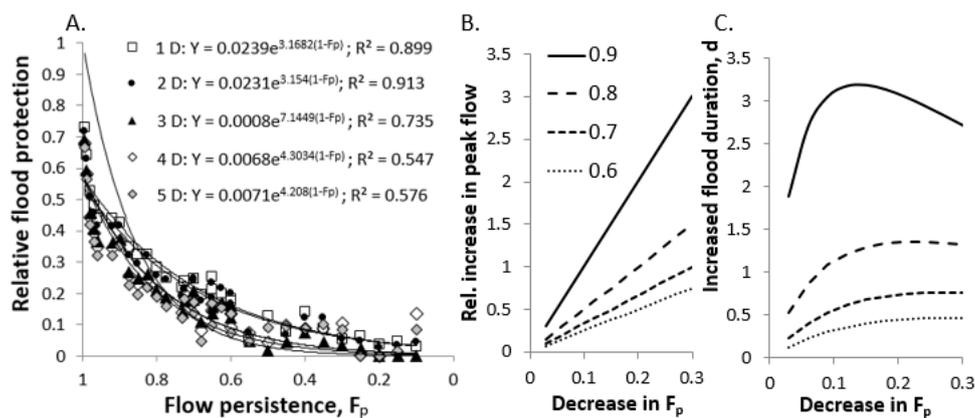
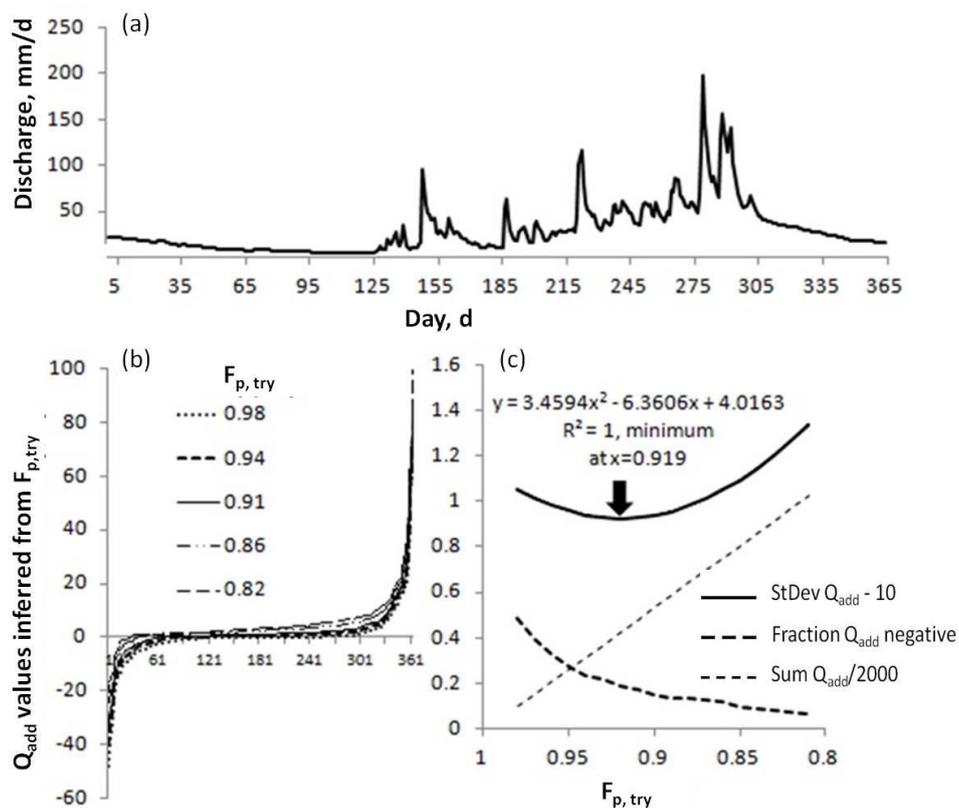
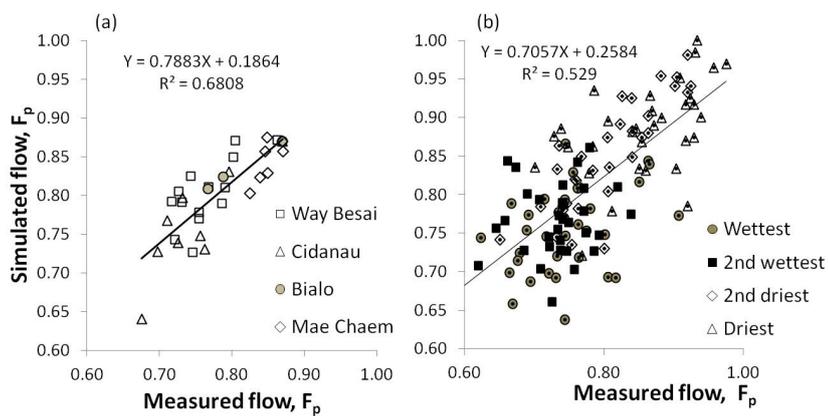



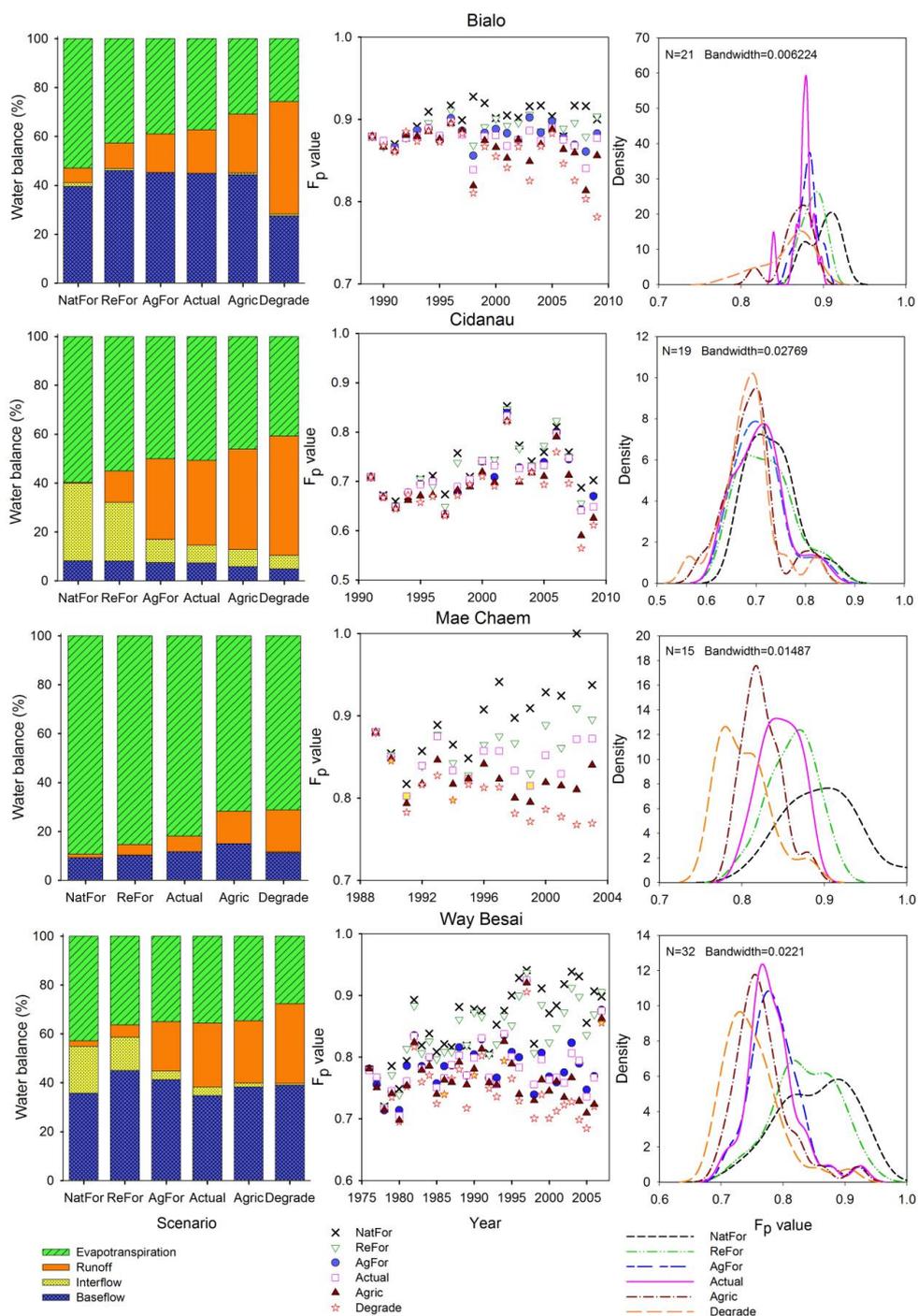
Figure 4. A. Effects of flow persistence on the relative flood protection (decrease in maximum flow measured over a 1 – 5 d period relative to a case with $F_p = 0$ (a few small negative points were replaced by small positive values to allow the exponential fit); B and C. effects of a decrease in flow persistence on the volume of water involved in peak flows (B; relative to the volume at F_p is 0.6 – 0.9) and in the duration (in d) of floods (C)

2   Figure 5. Example of the derivation of best fitting $F_{p,try}$ value for an example hydrograph (A)

3   on the basis of the inferred Q_{add} distribution (cumulative frequency in B), and three

4   properties of this distribution (C): its sum, frequency of negative values and standard

5   deviation; the $F_{p,try}$ minimum of the latter is derived from the parameters of a fitted

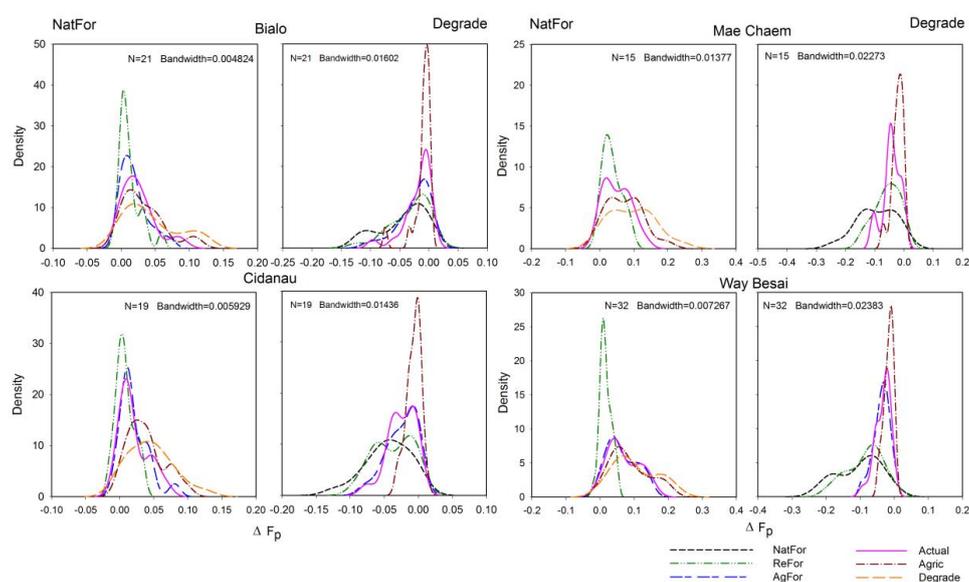
6   quadratic equation

2   Figure 6. Inter- (A) and intra- (B) annual variation in the F_p parameter derived from empirical

3   versus modeled flow: for the four test sites on annual basis (A) or three-monthly basis (B)

1 Figure 7. Effects of land cover change scenarios (Table 1) on the flow persistence value in

2 four watersheds, modelled in GenRiver[21] over a 20-year time-period, based on actual

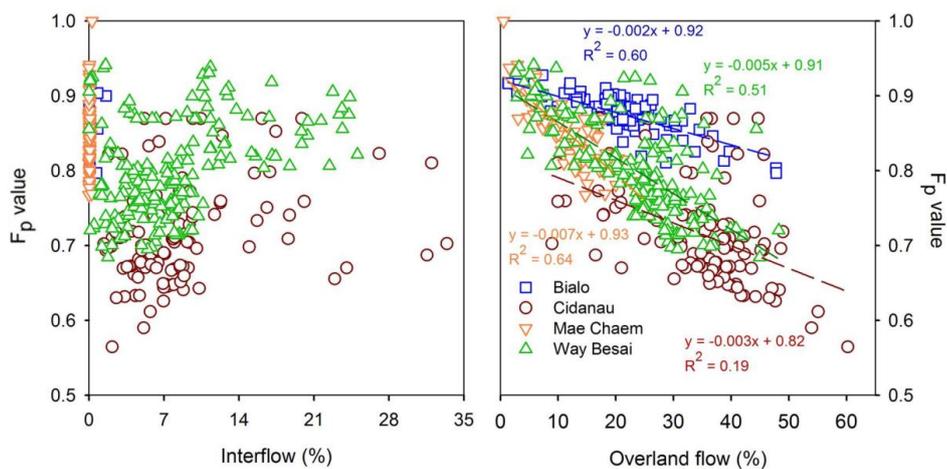
3 rainfall records; the left side panels show average water balance for each land cover

4 scenario, the middle panels the Fp values per year and land use, the right-side panels the

5 derived frequency distributions (best fitting Weibull distribution)

8 Figure 8. Frequency distribution of expected difference in F_p in 'paired plot' comparisons

9 where land cover is the only variable; left panels: all scenarios compared to 'reforestation',

10 righ panel: all scenarios compared to degradation; graphs are based on a kernel density

11 estimation (smoothing) approach

2 Figure 9. Correlations of F_p with fractions of rainfall that take overland flow and interflow

3  pathways through the watershed, across all years and land use scenarios of Fig. 7

1   Appendix 1. Example of a macro in R to estimate number of observation required using

2   bootstrap approach.

```
4    #The bootstrap procedure is to calculate the minimum sample size (number of observation) required
5    #for a significant land use effect on Fp
6    #bialo1 is a dataset contains delta Fp values for two different from Bialo watershed
7
8    #read data
9    bialo1 <- read.table("bialo1.csv", header=TRUE, sep=",")
10
11   #name each parameter
12   BL1 <- bialo1$ReFor
13   BL5 <- bialo1$Degrade
14
15   N = 1000 #number replication
16
17   n <- c(5:50) #the various sample size
18
19   J <- 46 #the number of sample size being tested (~ number of actual year observed in the dataset)
20
21   P15= matrix(ncol=J, nrow=R) #variable for storing p-value
22   P15Q3 <- numeric(J) #for storing p-Value at 97.5 quantile
23
24   for (j in 1:J) #estimating for different n
25
26   #bootstrap sampling
27   {
28   for (i in 1:N)
29   {
30   #sampling data
31   S1=sample(BL1, n[j], replace = T)
32   S5=sample(BL5, n[j], replace = T)
33
34   #Kolmogorov-Smirnov test for equal distribution and get the p-Value
35   KS15 <- ks.test(S1, S5, alt = c("two.sided"), exact = F) P15[i,j] <- KS15$p.value
36   }
37
38   #Confidence interval of CI
39   P15Q3[j] <- quantile(P15[,j], 0.975)
40
41   }
42
43   #saving P value data and CI
44
45   write.table(P15, file = "pValue15.txt") write.table(P15Q3, file = "P15Q3.txt")
```