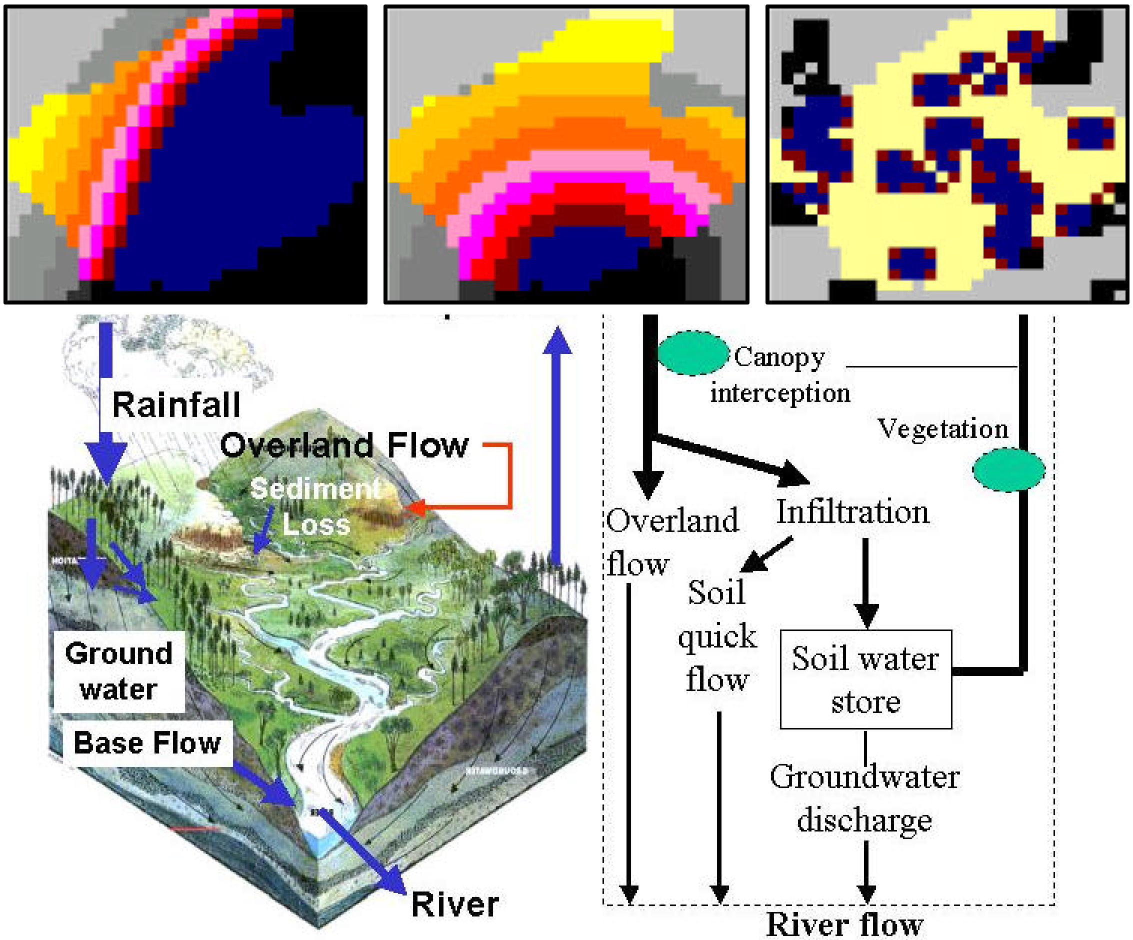
GenRiver and SpatRain model description

Meine van Noordwijk, Ai Farida, Desi Suyamto and Betha Lusiana
World Agroforestry Centre, ICRAF-SE Asia



World Agroforestry Centre
TRANSFORMING LIVES AND LANDSCAPES

Contents

1. Introduction

1.1 Modeling watershed functions

The term 'watershed functions' is often used in a rather loose way, suggesting that its various aspects (dimensions) change in a similar way when we make comparisons across climatic zones, land forms and human-induced land cover change. In reality, however, changes in total quantity of water may not be of the same relative magnitude (or even sign) as changes in quality or regularity of flow, and a differentiation among the 'functions' is needed. The 'functionality' of various aspects of river flow depends on the perspective, however, and thus may differ between various stakeholders. So, we may want to restrict ourselves to the hydrological 'consequences' of a watershed, and leave the value judgements of 'functions' to a later step in the analysis. The three main outcomes of current interest are:

- Quantity or total water yield
- Evenness of flow, which implies high flows in the 'dry' season and an absence of strong peak flows in the wet season
- Quality of water, with respect to its use as drinking water, other domestic uses, industrial use, irrigation or as habitat for fish and other water organisms

The behaviour of streams and rivers in these respects can be seen as the consequence of :

1. ***Site properties that 'come with the territory'***
   - local rainfall regime (and its temporal autocorrelation or tendency for wet days to follow wet days)
   - slope
   - soil depth and texture, determining the potential water storage, transport and retention
   - underlying landscape and geology that determines potential storage and release of groundwater
   - inherent properties of the riverbed

2. ***Scale***
   - size of the catchment (upstream of the observer/stakeholder) relative to the spatial autocorrelation of rainfall

3. ***Land use that directly depend on human activities***
   - infiltration and supply to groundwater as potentially influenced by soil structure that itself depends on vegetation and land use
   - vegetative aspects of the properties of the riverbed (and temporary storage) that dominate pulse transmission
   - irrigated agriculture and horticulture based on extractions from rivers

4. **Engineering structures**
   - canalisation of streams and rivers, increasing the rate of drainage
   - regulating structures in the river
   - impediments to rapid drainage in the form of dams and reservoirs

Where much of the public debate attributes most of the changes in 'watershed functions' to a change in forest cover (deforestation or reforestation), we need tools to account for the interactions of all four aspects mentioned here, to help us in assessing the causality of changes and the opportunities for interventions.

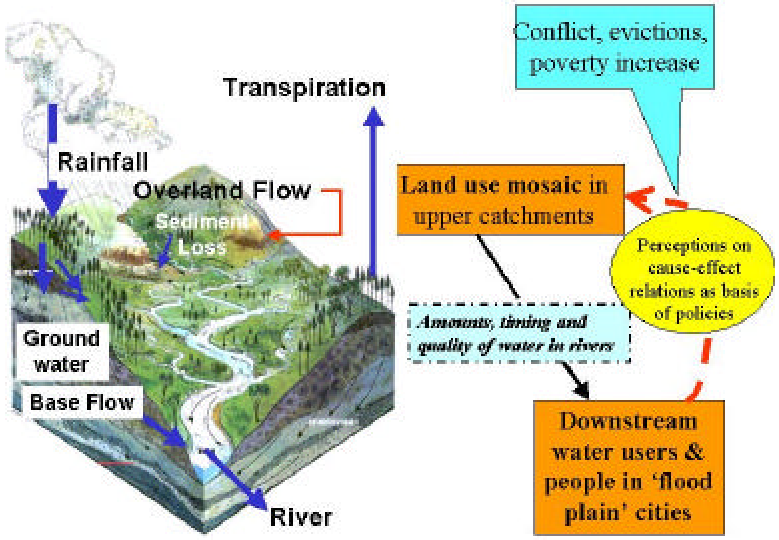


Figure 1. The biophysical relations between rainfall, land use in upper catchments and river flow to downstream areas are subject to discussions between down-stream and upland people whose perceptions on the cause-effect relations are reflected in policies that may aggravate poverty and conflict

Various approaches exist for modelling watershed functions, ranging from directly data-driven (empirical) approaches to models based on concepts of a water balance, soil physics and hydrology. Models differ by temporal and spatial scale: detailed description of rainfall and infiltration may require a minute (or even seconds) time step, especially on slopes where water will become surface runoff if it cannot infiltrate within seconds of reaching the soil surface. At the other end of the spectrum we may find empirical equations relating annual water yield of a catchment to annual rainfall (or precipitation in climate zones where snowfall and ice rains are significant). For some Indonesian catchments, for example, an empirical equation (Rizaldi Boer, *pers. comm.*) was derived as:

$$Q = 0.94\ P - 1000\ \text{mm year}^{-1}$$

with Q as river flow and P as precipitation both in mm year $^{-1}$. A tentative interpretation of these coefficients is that 6% of rainfall is lost through interception and direct evaporation from wet leaf surfaces and/or a rainfall-dependent increase in plant transpiration, and that the basic value for annual evapotranspiration is 1000 mm year $^{-1}$. Both these parameters, the interception loss, and the evapotranspiration will vary with the temporal distribution of rainfall and the land cover type, but the intercept is unlikely to change by more than 50% of the values given (so the intercept in unlikely to be more than 1500 or less than 500 mm year^{-1}), while the slope is probably confined to the range 0.8 – 1. The simple model may thus be fairly robust, but it is not sensitive to changes in land use or land cover (these could shift the parameters from the indicated values), and cannot be directly 'downscaled' to shorter periods of time (as it does not consider changes in storage terms). More sophisticated models will need to be explicit in the basic value for evapotranspiration of different types of land cover, and the degree to which these land covers induce direct evaporative losses.

Four classes of land cover can be distinguished from the perspective of evapotranspiration :

- ***open water*** bodies
  where water loss is determined by the relative humidity of the air and the presence of a stagnant boundary layer of air that reduces the transport of water vapour

- ***open soil***
  which may have a rate of evaporation similar to open water bodies when the surface is wet, but where evaporation may rapidly become limited by the rate of transport to the soil surface; soil cover with a litter layer provides a stagnant air zone, further reducing transport opportunities and mixing with the atmosphere

- ***seasonally green vegetation***
  most plants are able to provide their leaves (evaporating surfaces) with the amount of water that is needed for evaporation similar to an open water surface, during most of the rainy season; during periodic dry spells, plant transpiration is likely to drop below the value of open water, but stay above that of open soil

- ***evergeen vegetation***
  such as evergreen trees (e.g. pines, eucalypts, trees such as grevillea), irrigated rice paddies or vegetable crops will have a rate of transpiration equal to that of open water, or higher if lateral flows of dry air drive the evapotranspiration per unit area to higher levels

If we take for granted that effects of local land use on total annual rainfall are small, the main effect on total water yield of a catchment area is a change in the rate of evapotranspiration, or the return flow of water molecules to the atmosphere. In a simple equation: $Q = P – E - ?S$ or the total water yield (surface rivers Q_r + subsurface lateral flows Q_s + groundwater flows Q_g) equals precipitation (rainfall plus snow and ice, which in most parts of the tropic can be ignored) minus evapotranspiration minus the changes in storage terms of water in the catrchment. If the time frame for evaluation is sufficiently long relative to the variability of rainfall (e.g. one year for predictable humid climates but multiple years for more erratic drier areas), the ?S term can be ignored.

Efforts of land users that will reduce evapotranspiration and thus increase total water yield may thus be found in ***not*** planting evergeen trees (especially fast growing ones), or ***not*** irrigating rice paddies or vegetable crops in the dry season.

By expressing the rainfall and river flow in mm $year^{-1}$ we essentially use volume of water per unit area as the basis for calculations; if we consider larger areas, where both rainfall and evapotranspiration vary with space, we will need to make an effort to adjust the average value to maintain validity of the equation. For annual water yield, however, an area-based approach to scaling is valid, and values per unit area can be used to estimate values for any scale through multiplication with area. For properties such as ‘evenness of flow’ or probability of flooding, the relation with the scale of consideration is more complex, and a greater sensitivity to both the mean value of land cover fractions as well as the spatial organization of the landscape is probably needed.

If a greater model sensitivity to land use change is important for the question we try to answer or if we are interested in phenomena operating at shorter time scales than a year, we need to take into account the intermediate processes that determine the access to and use of water stored in the soil and the upper groundwater, as well as the rates of transport and temporary storage of water in the river network. The basic framework for a patch or plot level water balance (Figure 1) is well accepted, so the various models differ in the details of the time course of describing canopy interception and throughfall, and the way lateral flows over the surface and through the soil are described. As most plot level studies exclude surface inflows, there is a tendency to focus on surface runoff rather than run-on or net transport.

A number of existing models address only a single scale, be it a plot or a catchment as a whole (Table 1). Other models use a grid-cell approach with interactions between 'cells' leading to emergent behaviour at the catchment scale. A third category of models addresses the cross-scale questions in a more direct way by being specific on how properties change with the temporal and spatial scale of consideration. The ***GenRiver model*** was made for data-scarce situations and is therefore based on 'first principles', as these may be considered the safest bet for a wide range of applications (acknowledging that directly empirical models may have greater precision within the tested range). The model includes an attempt to relate across spatial scales (Fig. 3).

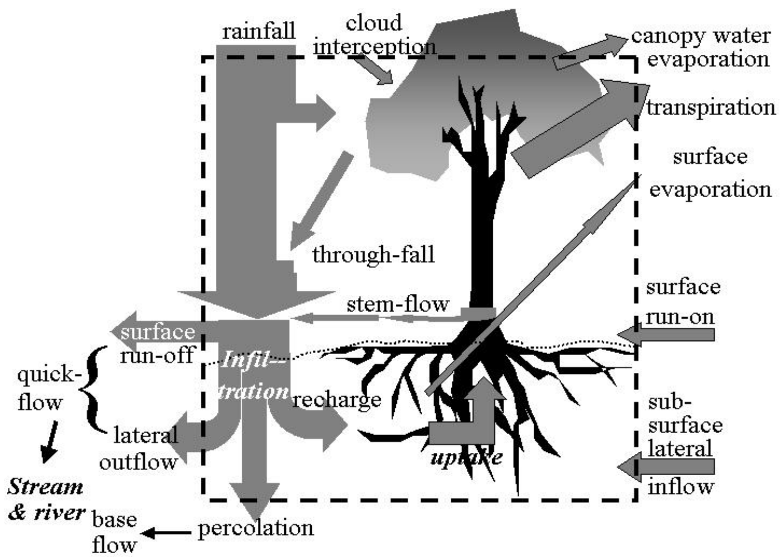


Figure 2. Water balance at patch level that forms the core of the GenRiver model

Table 1. Models concepts of river flow can be classified as:

	Single scale	Spatially explicit, multiple entities at the same scale	Across scale
Empirical, catchment specific	Hydrograph analysis, Runoff fraction at plot level, USLE, 'Parsimonious' catchment models	Spatial correlation of rainfall, USLE applied to GIS grid data	Nested hydrograph analyses, Sediment delivery ratio
Based on water balance and generic principles of soil physics and hydrology	Plot-level water balance, Catchment-level water yield model	GIS: raster or polygon based	Nested models with explicit scaling rules

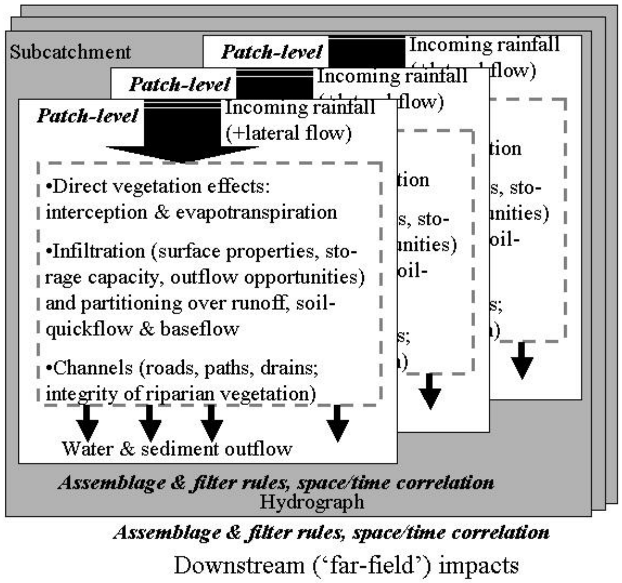


Figure 3. Models for watershed functions at catchment scale need to combine explicit rules for effects of land use on interception, infiltration and transport to the stream network at 'patch' scale, with assemblage and filter rules that reflect the river network and the changes that this can cause to the overall flow

1.2 Two alternative explanations for steady river flow

Everybody is probably familiar with the 'mental model' of the forests as a sponge, that receive rainfall and gradually feed it to the stream. The concept, clearly formulated in the 1920's in Indonesia, but is essence much older than that, has been seriously questioned in the 1930's (see appendix 1) and internationally in the last two decades (Calder, 2002). The validity of the concept is especially questionable for the humid tropics, where the sponge will be continuously wet and not able to absorb much of the incoming rainfall. Yet, the 'sponge' concept still leads to specific expectations that only 'forest' can play this role. If we accept that some forms of 'non-forest' can maintain infiltration rates, the 'local buffering' perspective still leads to strong concerns against any land use intervention that reduces the residence time of water in the system, on its way from rainfall to the river.

There is, however, an alternative explanation for even river flow patterns, that gets much less attention: spatial heterogeneity of rainfall. Simply said: if today it rains here and tomorrow there, the river that receives water from both areas may have a fairly steady flow, despite a poor buffering in either areas (see figure 1). If this second model dominates, changes in river flow may be due to a change in the spatial correlation of rainfall, not to land use change in any of the subcatchments per se.

A distinction between these two types of explanation for patterns in river flow is thus essential to evaluate the likely impact of current land use change in forested areas and the types of interventions that may be effective or not. The relative importance of the two explanations clearly depends on the scale of consideration. In small subcatchments there is hardly any space for the second explanation, and the first must dominate.

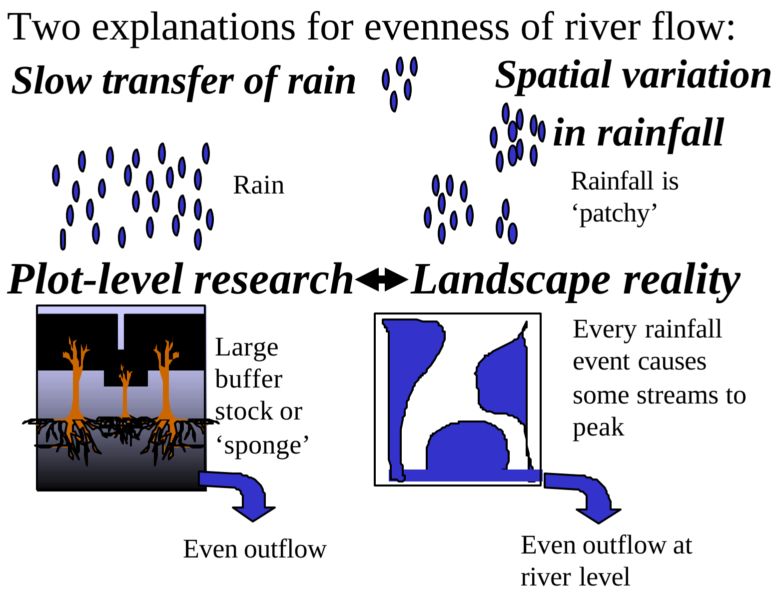


Figure 4. Two alternative models for steady river flow: the 'sponge' and 'patchy rain' version that are likely to dominate at the scale of plot level research (left) and at landscape scale (right)

In areas of several hundreds of square kilometers or at subcontinental scale, the second reason is likely to dominate. So, somewhere at intermediate scale the two may break even – can we assess where this occurs? Unfortunately, most past research was done in small plots and in 'scaling up' the possible impact of the second explanation was not recognized… In summarizing data in land use impacts on river flow (Kiersch and Tognetti, 2002) no cases were reported with measurable impacts of land use change on river flow of areas larger than 100 km^2.

Table 2. Well-documented impacts of land use change by basin size (Kiersch and Tognetti, 2002); x = Measured impact; - = No well-documented impact

Impact Type	Basin size [km]						
	0.1	1	10	10^2	10^3	10^4	10^5
Thermal regime	x	x	-	-	-	-	-
Pathogens	x	x	x	-	-	-	-
Average flow	x	x	x	x	-	-	-
Peak flow	x	x	x	x	-	-	-
Base flow	x	x	x	x	-	-	-
Groundwater recharge	x	x	x	x	-	-	-
Organic matter	x	x	x	x	-	-	-
Sediment load	x	x	x	x	-	-	-
Nutrients	x	x	x	x	x	-	-
Salinity	x	x	x	x	x	x	x
Pesticides	x	x	x	x	x	x	x
Heavy metals	x	x	x	x	x	x	x

The Genriver and SpatRain models were first designed to answer this rather specific question: how does spatial variability of rainfall influence the 'evenness' of river flow that is often attributed to forests as dominant land cover?, or 'explanation 2'. We first of all need a representation of rainfall with spatial patterns that are intermediate between uncorrelated random and fully coupled. We then need to link this to a model that includes the 'sponge' in its essential form, so that we can compare the relative importance of both processes. The two tools described here, SpatRain and GenRiver were developed for such a purpose. We will briefly outline the conceptual basis of both, describe the model implementation and parameter sensitivity, and then proceed with the analysis of the relative impacts of land use change on riverflow in catchments with spatially heterogeneous rainfall.

1.3 Quantification of 'buffering' of riverflow by watershed areas

A basic concept in 'watershed functions' is '***evenness of riverflow***', indicating low peak flows' and high 'base flows'. The variation in river debit between different rivers, however, is largely due to variation in rainfall, and it is no trivial task to separate this climatic effect (that we assume to be independent of local land use change, for the time being at least) from the impacts of land use change. The following definition of '***buffering***' can allow us to make this separation.

An efficient way of presenting the input and output of a watershed area in a single graph, is to look at the exceedance probabilities for daily rainfall, daily evapotranspiration and daily riverflow. If a sufficiently long time period is considered (at least 1 year), changes in storage in soil, groundwater and surface water may be negligible and the areas to the left of the curves for rainfall and evapotranspiration + riverflow should be approximately equal. The point of intersection has to have an X-value that equals the mean daily rainfall. The intersection would be at an exceedance probability of 0.5 if rainfall distribution were symmetrical and there would be no dry days – in reality skewness of rainfall distribution plus the fraction of days without rain cause the point of intersection to have a value on the Y-axis that is above 0.5.

In an 'asphalted' watershed, the riverflow curve may be expected to coincide with the rainfall curve and there is no buffering. In an ideally buffered situation the riverflow may be constant and equal to the mean at every day of the year. In between these two extremes we'll find real watersheds with a partial 'buffering'.

1.4 Target properties of the model

The model was developed with the following target properties. The model should be:

- based on solid principles of the plot-level water balance and the way this is influenced by ***land use change***, through vegetation and changes in soil structure over time preferably compatible in approach to the WaNuLCAS model that operates at higher spatial resolution of soil zones and layers for mixed copping and Agroforestry situations
- handling processes at less-than-hourly time scale where infiltration is concerned and at daily time scales for stream and river flow
- applicable to multiple subcatchments that together form a catchment and that receive rainfall events partially correlated (so in between the assumptions of 'homogeneity' and 'statistical independence')
- applicable to any land form and digital elevation model (DEM) at 'parameter' level, rather than by modifying model structure
- able to predict river flow (hydrograph) at multiple points of interest
- transparent in structure (assumptions) and easy to operate

2. Genriver model overview

The model was initially designed as a 'simple' (few parameter) model that still has a link to process-based models, and that can be gradually spatially differentiated, as the need arises.

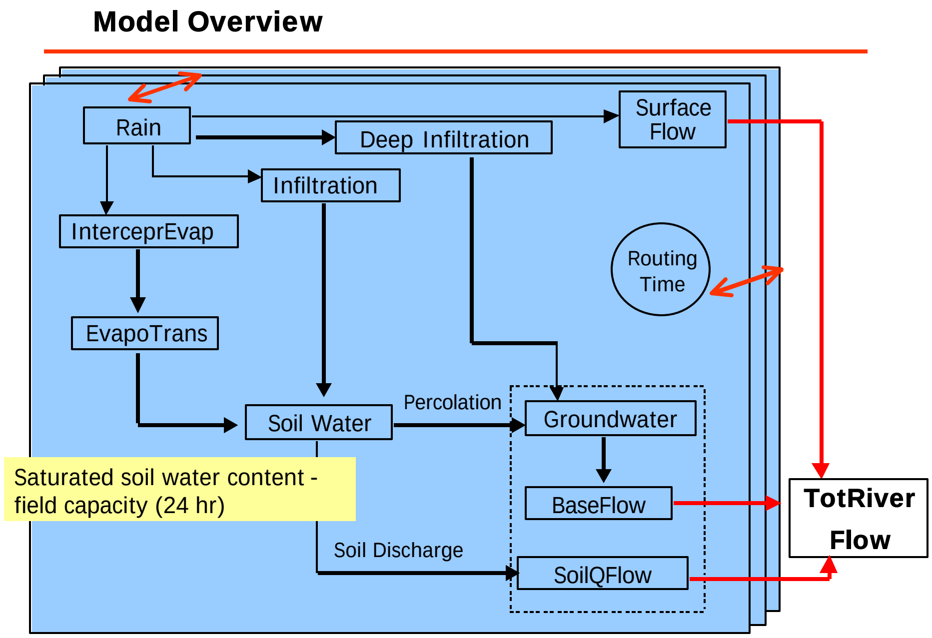


Figure 5. Overview of the GenRiver model; the multiple subcatchments that make up the catchment as a whole can differ in basic soil properties, land cover fractions that affect interception, soil structure (infiltration rate) and seasonal pattern of water use by the vegetation. The subcatchment will also typically differ in 'routing time' or in the time it takes the streams and river to reach the observation point of main interest

The core of the model is a 'patch level representation of a daily water balance, driven by local rainfall and modified by the land cover and soil properties of the patch. The patch can contribute to three types of stream flow: surface-quick flow on the day of the rainfall event, soil-quick flow on the next day and base flow, via the gradual release of groundwater.

Table 3. The overall water balance of the model , summed over space and time

In	**Out**
P = precipitation (Rainfall)	E = Evapotranspiration
- ? s = Changes in soil and groundwater storage	Q = River debit (summed over base flow, soil quick flow and surface quick flow)
- ? r =Changes in the volume of water in streams and rivers	? = Error (***unaccounted for***) term (difference between all in & out terms)

For the long-term behaviour the changes in soil and groundwater storage, as well as changes in the volume of streams and rivers will be negligible, while the error term should be negligible at all times if the model is correctly implemented.

On shorter time scales, however, the changes in storage in soil, groundwater (? s), streams and rivers (? r) are critically important for the variability in (daily) river flow as reflected in the 'hydrograph' (Fig. 6a).

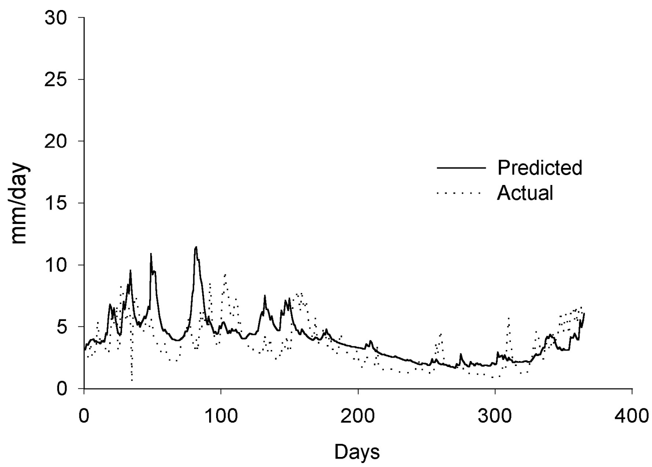


Figure 6a. An example of a hydrograph as output of the model for one year of simulation

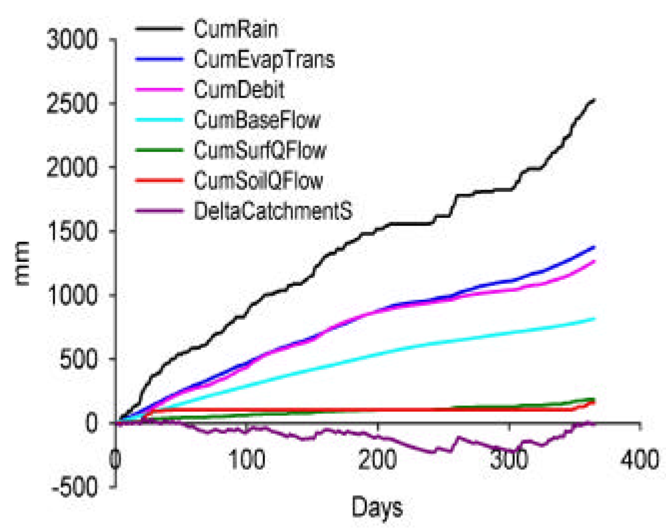


Figure 6b. Cumulative water balance as output of the model with cumulative river flow(CumDebit), approximately equal to cumulative rainfall (CumRainfall) minus cumulative evapotranpiration (CumEvapoT rans). The dynamic change in catchment storage (DeltaCatchmentStorage) account for differerence between these cumulative terms. Cumulative baseflow (CumBaseFlow), surface quick flow (CumSurfQFlow) and soil quick flow or 'inter flow' (CumSoilQFlow) can be calculated in this model.

Many models for river flow, especially for drier areas, focus on the overland flow directly after rainfall (Quickflow) but do not account for the 'slow flows', that derive from water that infiltrates into the soil but can take a range of pathways, with various residence times, to reach the streams and rivers, depending on land form, geology and extractions along the way. To keep things simple, GenRiver distinguishes only two steps in this: a soil quick flow (or 'inter flow') that is considered to reach the streams a day after the rainfall event, and a 'slow flow' that forms a fraction of the available store of groundwater (leading to an exponential decline of the groundwater store with time and a linear relation ship between the logarithm of the discharge and time in the absence of rainfall).

3. Brief description of GenRiver and component processes

A ***river*** is treated as a summation of ***stream***s, each originating in a ***subcatchment*** with its own daily rainfall, yearly land cover fractions and constant total area and distance to the river outflow (or measurement) point. Interactions between streams in their contribution to the river are considered to be negligible (i.e. there is no 'backflow' problem). Spatial patterns in daily rainfall events are translated into average daily rainfall in each subcatchment in a separate module. The ***subcatchment*** model represents interception, infiltration into soil, rapid percolation into subsoil, surface flow of water and rapid lateral subsurface flow into streams with parameters that can vary between land cover classes.

3.1 Daily rainfall amount

Rainfall at subcatchment level is implemented as daily amounts from long time records for each subcatchment, stored in an Excel spreadsheet. These data can be derived from actual records, or from a 'random generator' that takes temporal patterns (as done in Markov chain models, such as used by Jones et al., for temporal autocorrelation) into account, as well as the spatial correlation of rainfall at any point in time. See also the SpatRain description, below.

3.2 Rainfall intensity and time for infiltration

Rainfall duration is estimated from the daily amount, and a rainfall intensity estimate for the given day (mm $hour^{-1}$) that is derived from a mean value, a coefficient of variation and a random number. Rainfall duration determines the 'time available for infiltration', but this can be modified by canopy interception of rainfall followed by the duration of the 'dripping' phase (default 0.5 hour).

3.3 Interception

Storage capacity for intercepted water is treated as a linear function of leaf + branch area index of the land cover, with the option of modifiers for surface properties that determine the thickness of the water film. Interception-evaporation has priority over plant transpirational demand.

3.4 Surface infiltration/runoff

Infiltration is calculated as the minimum of

- the daily infiltration capacity times the fraction of a day that is available for infiltration (the latter reflects rainfall intensity as well as the local storage capacity of the soil surface)
- the amount that can be held by the soil at saturation minus the amount already present ***plus*** the amount that can enter the groundwater within a day (which in itself is the minimum of the potential daily transport rate and the difference between maximum storage capacity of groundwater and the current amount)

If the first constraint is active, the model generates 'infiltration limited runoff', in the second case 'saturation overland flow'.

3.5 Evapotranspiration

Total evapotranspiration is driven by potential evapotranspiration (Penman type) and (partially) met by :

- Intercepted water
- Land cover, with a drought-limitation proportional to soil water content relative to field capacity below a (vegetation dependent) threshold
- Soil surface evaporation (not explicit – to be included in the land cover /vegetation properties for transpiration)
- Weekly multiplier on potential daily evapotranspiration, reflecting overall phenology
- potential relative evapotranspiration per land cover type (per month).

3.6 Soil water redistribution

During a rain event the soil may get saturated, but within one day it is supposed to drain till 'field capacity' (with an operational definition of the soil water content 24 hours after a heavy rainfall event). The difference between saturation and field capacity can be either :

- Used for transpiration (but canopy intercepted rainfall takes priority to meet the demand)
- Drain to the groundwater reserve, calculated as the minimum of the amount that can be transported downwards and the fraction of soil water that will drain on any given day
- Drain to the rivers as 'soil quick flow': any water left above field capacity by the two preceding processes

3.7 Groundwater release to streams (baseflow)

Surface quick flow, soil quick flow and base flow all feed into streams. For each subcatchment a 'Routing' function determines the time delay before the water passesby a defined measuring point (currently the outflow from the catchment).

3.8 Distance (routing distance)

Distance from the mid point of each subcatchment to any number of observation point This parameter will derive the routing time for each subcatchment to each of the observation point, while excluding subcatchment downstream of the observation point.

Table 4. The model input is thus driven by the following parameters

Acronym	Definition	Dimension [default value]
RainIntensMean	Average intensity of rainfall	mm $hour^{-1}$ (30)
RainIntensCoefVar	Coefficient of variation of this intensity	[] (0.3)
InterceptPerClass (j)	Interception storage capacity per land cover class	mm
LandCoverFreq (i,t)	Land cover class frequency per unit i	[]
MaxInfRate (i)	Maximum infiltration capacity per unit i	mm day^{-1} (1000)
RelativeDrought Threshold (j)	Drought-limitation to transpiration per land cover class, as fraction of field capacity	[]
FieldCapacity (i)	Field capacity of the soil (soil water content 1 day after 'soalking' rain	mm (250)
SoilSatminusFC (i)	Difference between saturation water storage capacity and field capacity of the soil	mm (100)
MaxDynGrWatStore (i)	Dynamic groundwater storage capacity	mm (350)
Routing Time (i)	Routing time	day
PerFracMultiplier	Daily soil water drainage as fraction of groundwater release fraction	[] (0.5)
GWReleaseFracVar	An option to have a constant groundwater release fraction for each subcathment or using single value for	[] (1)

	the whole catchment	
GWReleaseFracConst (i)	Daily groundwater release fraction	[] (0.03)
InitRelSoil	Initial soil water content relative to field capacity	[] (1)
InitRelGroundwater	Initial groundwater store relative to maximum value	[] (1)

Note : index **t** refers to time dependent input, **i** to subcatchment and **j** to land cover classes

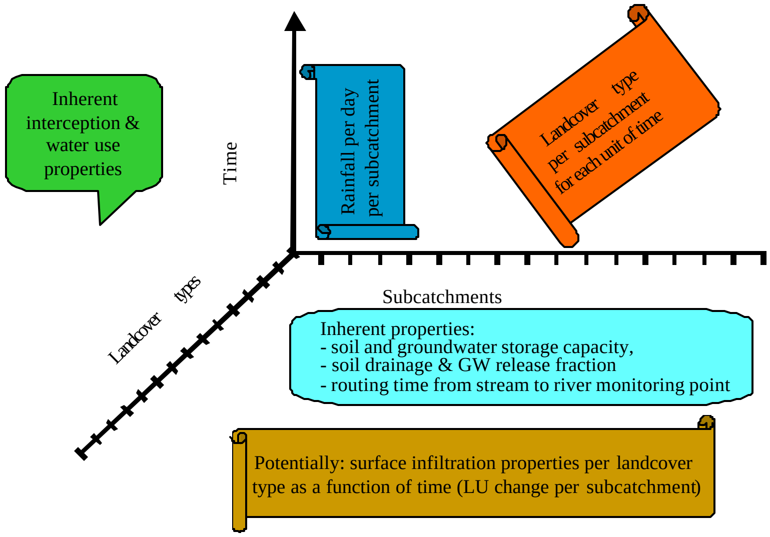


Figure 7. Array dimensions is used in the model

In the model a number of 'array' dimensions is used. Space-related properties are described with the 'subcatchment' array, that is (potentially) used to specify soil and groundwater properties as well as the 'routing' time for stream flow to reach the point where the river is monitored. A number of land cover classes can be distinguished (as in the 'FALLOW' model, but the GenRiver model operates on daily time steps, while FALLOW has a yearly time step), with different properties for in terception of rainfall and water use by transpiration (both at the potential level and where a drought threshold is passed). Time is the main dimension of change, of course, and rainfall is the main time-dependent input, that can be made specific to each subcatchment (but not differentiating land cover classes). Surface infiltration properties are currently described as a constant, but they may well have to be made dependent on subcatchment. A relationship between this property and the land cover class is required that describes the change in this property as a function of time since land cover change.

4. Model implementation in Stella

4.1 Program sectors

The model has the following components: an initialisation part, the dynamic sectors (dealing with patch-level water balance, stream and river flow, and the operational rules of reservoirs in the river network) and two sectors for keeping track of all output parameters required.

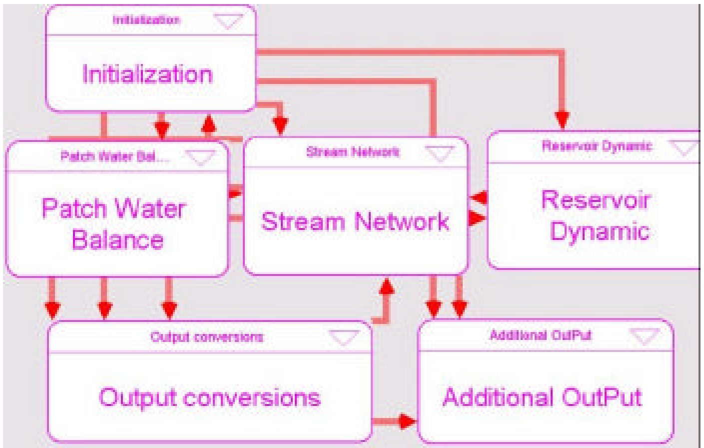


Figure 8. Sectors of the model implemented in Stella.

4.2 Patch-level water balance

The 'patch water balance' section in the model (Figure 9) closely follows the diagram in figure 5. The amount of rainfall for each land use category within each subcatchment is calculated from the rate per unit area and the respective area fractions. In order to implement the flows of this rainfall at the same day to either of the pools 'soil water', 'groundwater', 'cumulative evapotranspiration' or the surface runoff flow, the equations for the respective flows are linked and prioritized (with interception having priority and surface runoff being the residue of potential infiltration and rainfall *minus* interception.

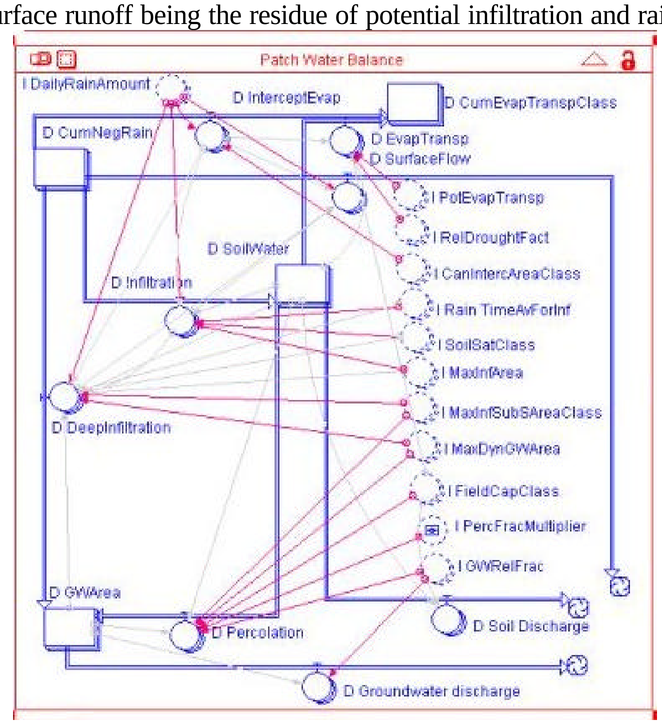


Figure 9. Diagram of the patch-level water balance with the stocks represented by rectangles, parameters by circles, water flows by double-lined arrows and information follows by single lines

4.3 River flow

The implementation of the stream and river flow sector distinguishes between the part of the surface flow that can reach the observation points on the day of rainfall and the part that has one or more days of delay before reaching that point. As each calculation step (from stock to stack) takes one time step in STELLA, special care is needed to obtain same-day flows. A similar issue holds for the soil quick flow that can reach the observation point at the earliest one day after the rainfall.

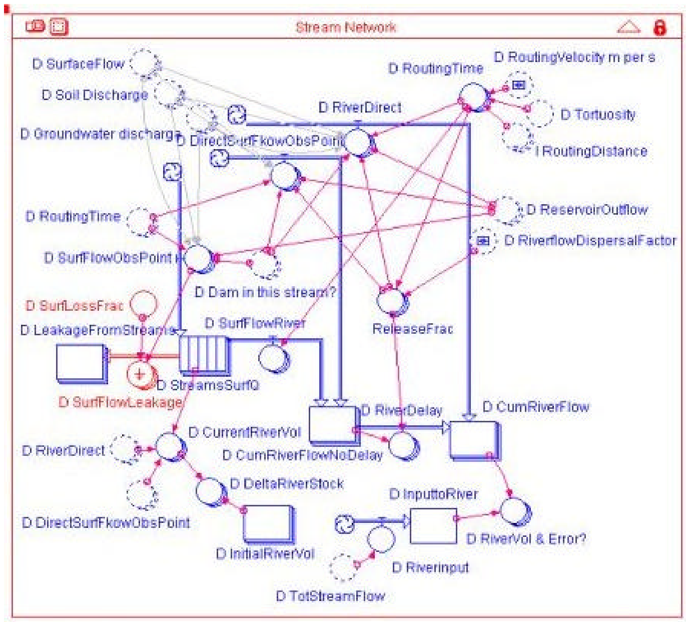


Figure 10. Diagram of the river network with the stocks represented by rectangles, parameters by circles, water flows by double-lined arrows and information follows by single lines.

4.4 Input parameters

User can define the length of simulation and the measurement period start and end. Internally, many of the input parameters are processed in the initial section to convert units and apply the area fractions of the various subcatchments and land cover fractions (in fact – these land cover fractions can be dynamic).

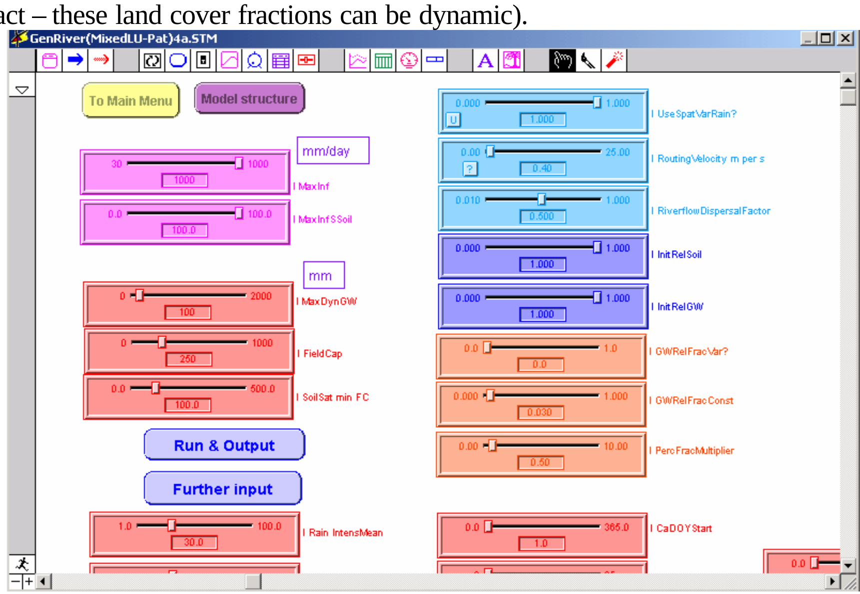


Figure 11. Input screen in Stella, the user can modify input values by using the 'sliders'.

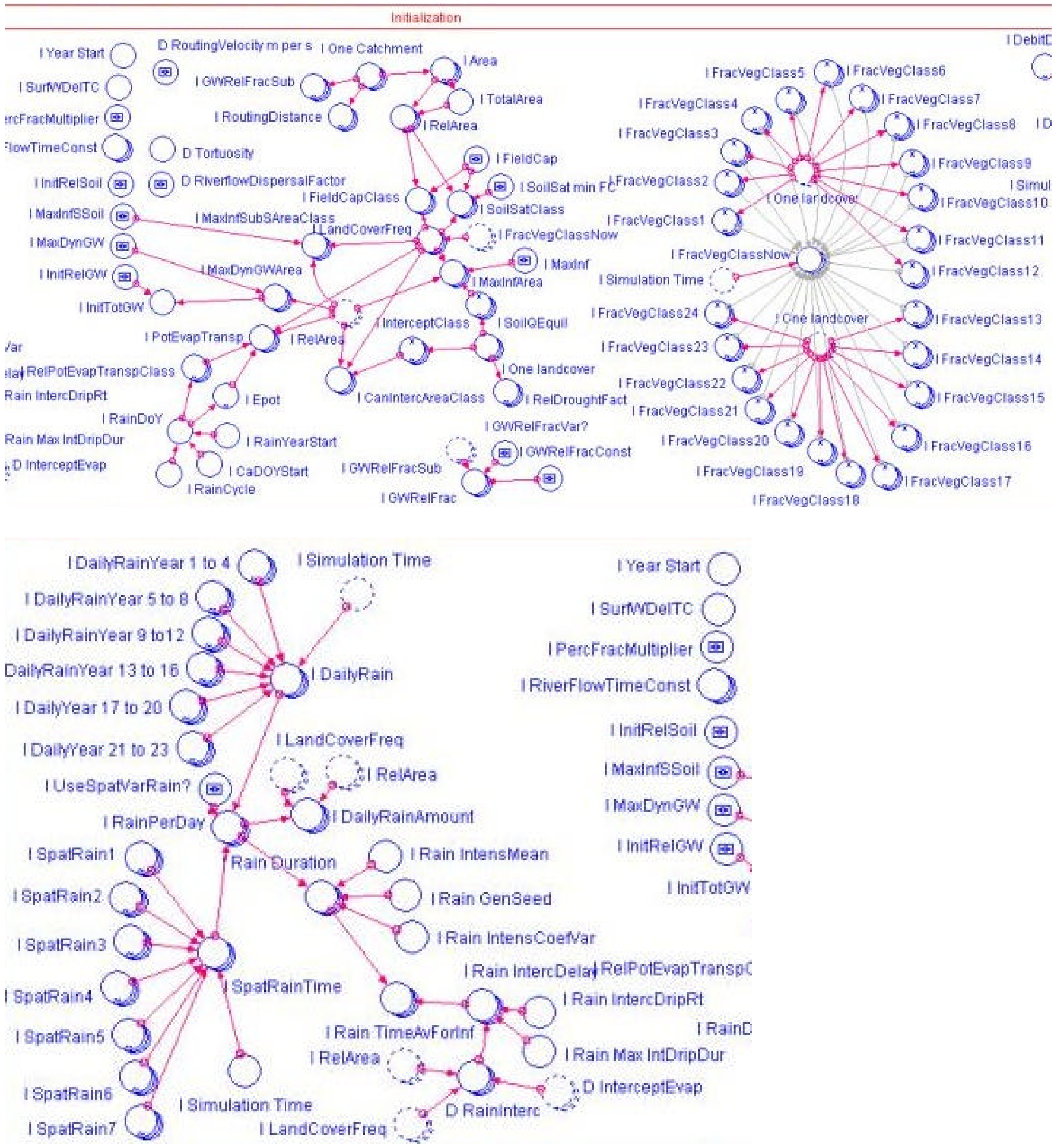


Figure 12. Sectors in the STELLA implementation that process the input parameters

4.5 Example of Output screen in Stella

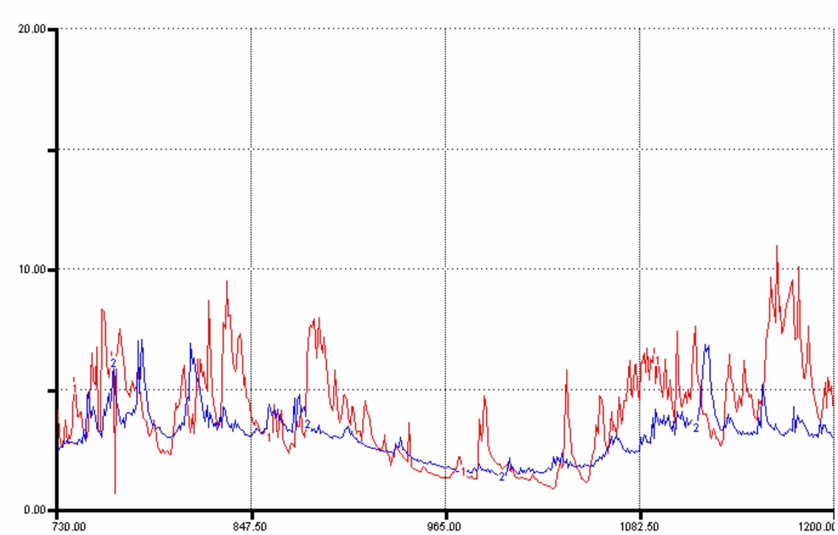


Figure 13. Example of debit prediction (blue) and debit data (red) as one of output parameter
We can compare measured and simulated river flow over time – in this example no specific effort to 'fit' the model was made.

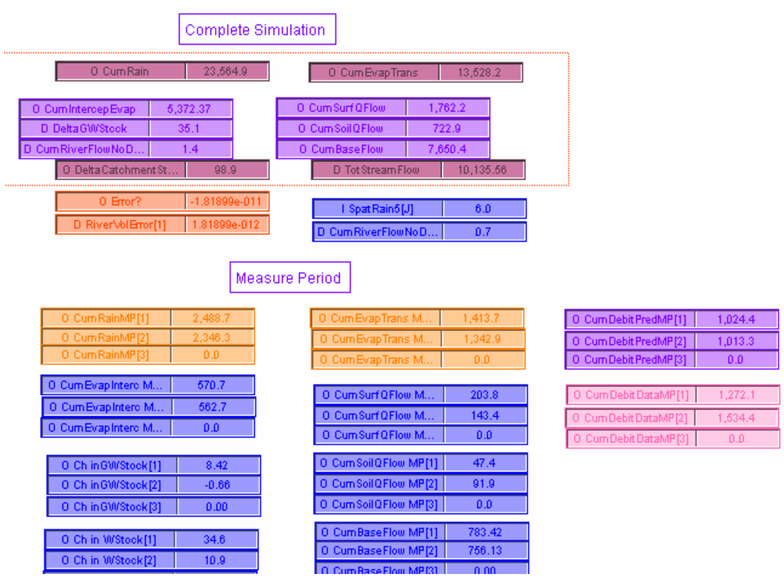


Figure 14. An example of cumulative table as output. The balance per user defined measuring period allow us to check sensitivity of the water balance paramaters.

4.6. Initialization

The model has a number of parameters that require initialization, including the soil and groundwater stocks at the start of the model. As parameter values for such can not be empirically derived, we let the model run for a number of years before we derive the output -- this way the output values become independent of the choice of initialization values (demonstrate with example for water balance in year 1, 3 and 5)

Table 5. Initialization for Soil and Groundwater content

Water Balance (mm)	Initial Soil Water Content (at initial Groundwater =1)					Initial Groundwater (at Initial Soilwater content = 1)				
	0	**0.25**	**0.5**	**0.75**	**1.0**	**0**	**0.25**	**0.5**	**0.75**	**1.0**
CumEvapTrans(1)	740	789	875	968	1018	1018	1018	1018	1018	1018
CumEvapTrans(3)	926	926	926	926	926	926	926	926	926	926
CumEvapTrans(5)	926	926	926	926	926	926	926	926	926	926
CumSurfQFlow(1)	785	785	785	785	785	785	785	785	785	785
CumSurfQFlow(3)	785	785	785	785	785	785	785	785	785	785
CumSurfQFlow(5)	781	781	781	781	781	781	781	781	781	781
CumSoilQFlow(1)	0.4	1.8	4.7	7.7	68.5	66	66	66	66	66
CumSoilQFlow(3)	10.5	10.5	10.5	10.5	10.5	10.5	10.5	10.5	10.5	10.5
CumSoilQFlow(5)	10.5	10.5	10.5	10.5	10.5	10.5	10.5	10.5	10.5	10.5
CumBaseFlow(1)	905	1003	1064	1116	1154	807	895	982	1070	1154
CumBaseFlow(3)	833	833	833	833	833	833	833	833	833	833
CumBaseFlow(5)	833	833	833	833	833	833	833	833	833	833

Large values for initial soil and groundwater stock lead to higher base flow and evapotranspiration in year 1 but in year 3 the base flow already is independent of the initial value. To be on the safe side, we take year 5 for all subsequence output shown.

4.7 New GenRiver Application

The model was initially developed to analyze river flow in Way Besai watershed in Sumberjaya, Lampung (Indonesia), so current default input parameter are based on Sumberjaya condition. In order to make a new GenRiver application for different watershed, we need to prepare following data :

1. Climate
   - ? Rainfall
     A number of formats are possible, as long as they allow a reconstruction of monthly exceedance curves of daily rainfall intensity:
     - 30 (or at least 20) years of daily rainfall records for a station that can represent the area (or multiple stations if these are supposed to be similar) ***or***
     - any 'rainfall simulator' equation with the appropriate parameters that can be used to generate a 30 year dataset for the site (e.g. MarkSim?).
   - ? Rainfall intensity
     Data on rain duration and amount for a sampling period that is deemed representative to estimate the mean and coefficient of variation of rainfall depth per hour.
   - ? Rainfall spatial correlation
     An indication of the degree of spatial correlation in rainfall (correlation coefficient of daily rainfall as function of distance between stations), or of the generic nature of rainfall (frontal rains with high spatial correlation or convective storms that are 'patchy' and show low correlation).
   - ? Potential evaporation
     Average values per month, derived from open pan evaporation measurements or from equation such as Penman's that is calibrated on such data.

2. Landform
   Coarse DEM that allows for derivation of overall difference in elevation within the subcatchment, and a delineation of subsubcatchments. If there is a generic 'language' for the shape of the subcatchments relative to the main channel, we may use this.

3. Soils
   - Mean soil depth (till major restriction for root development)
   - Average texture (or soil type in a way that allows texture to be estimated) as input to 'pedotransfer' functions to estimate soil water retention curve (saturation, field capacity, wilting point)
   - Estimated bulk density relative to the reference value for soils under agricultural use, to estimate saturated hydraulic conductivity and potential infiltration

4. Geology
   We need to estimate the 'differential storage' in 'active groundwater' as well as a 'groundwater release' fraction. So far these parameters were 'tuned' to the recession phase of actual riverflow during periods without rainfall. In the absence of such data we will need to 'guesstimate'. If data on the seasonal variation in depth of groundwater table are available, we can use those.

5. Vegetation and Land cover
   Fractions of total land cover that are
   - Deciduous (reducing LAI in dry season to near 0)
   - Semi deciduous (reducing LAI in dry season to less than 0.5 (??) of value in wet season)
   - Evergreen maintaining LAI at over 0.5 of the maximum value
   - Bare soil or build-up areas
   - Open surface water

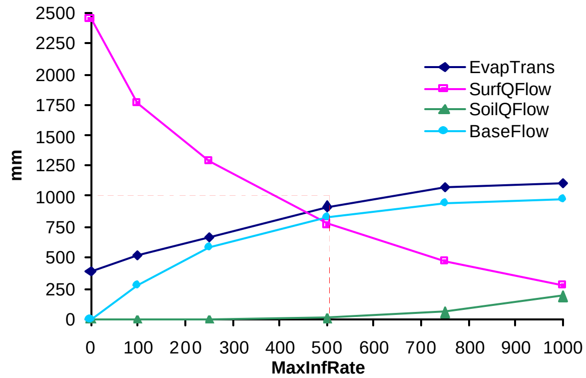
   For more detailed assessments in the Sumberjaya and Mae Chaem areas we will use the actual time course of change. On that basis we might do with an estimate whether the actual change in the 'virtual' subcatchments has been 'rapid' (like 60 - > 10% forest cover in 25 years), 'extremely rapid (faster than that), or slower.

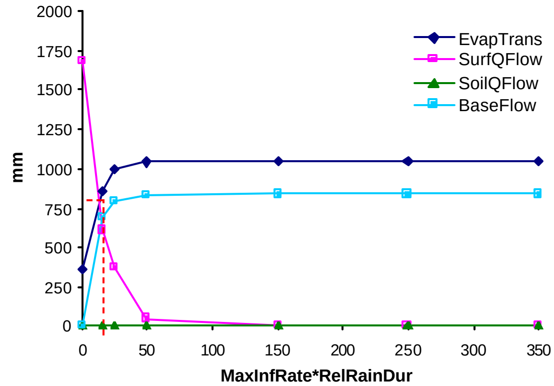
6. Actual river discharge
   If available, river debit data for any period of time (expressed in $m^3 s^{-1}$ in the river or mm day^{-1} over the whole contributing catchment) will be valuable in 'constraining' the simulations. If not available, we will simply have to 'believe' the model predictions as such.

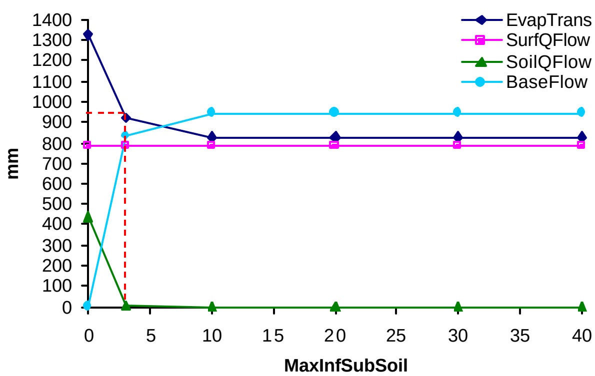
5. Sensitivity analysis around default parameter values

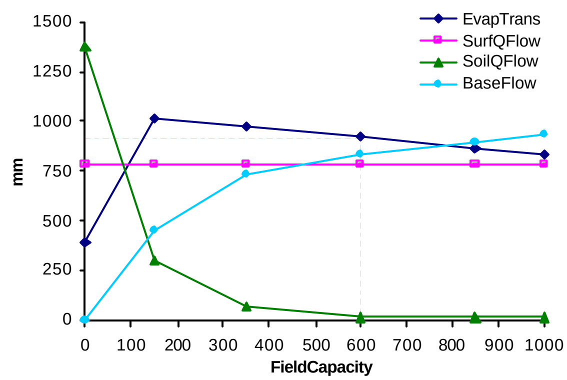
5.1 Patch-level water balance

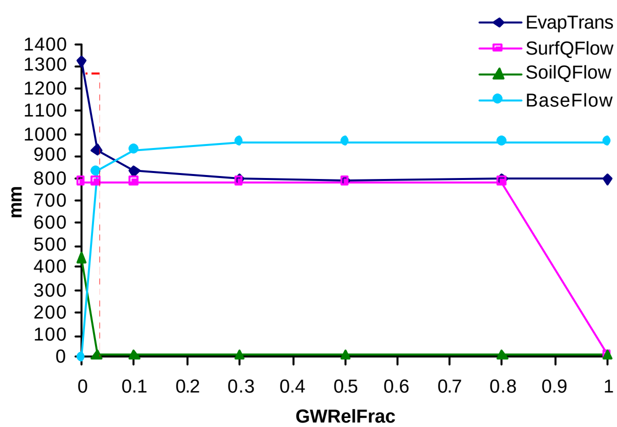
This section will discuss the sensitivity of model outputs to variations in the various parameters for the patch-level water balance. Results refer to year 5 of a simulation (to become independent of initialisation); a red line in all figures shows the default value of each parameter. We first focus on the terms of the cumulative water balance.

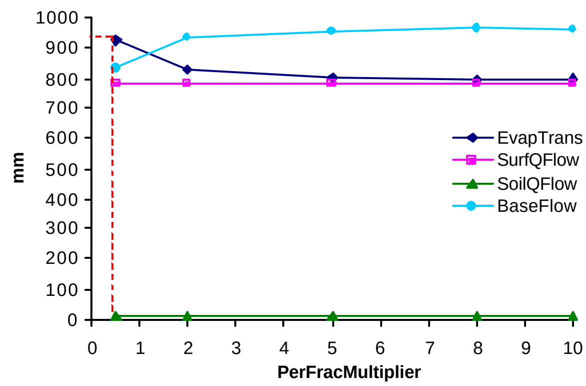












All water balance parameter respond continuously when we the change Maximum Infiltration Rate (MaxInfRate), while for other parameter's response has thresholds. Changing of MaxInfRate value will change all water balance parameters reflecting the importance of infiltration - limited runoff process. Cumulative base flow and evapotranspiration increase with larger MaxInfRate, value while decreases of Surface quick flow (SurfQFlow). The results for changes in MaxInfRate directly related to those for modifying the Relative Rain Duration (RelRainDur) parameter as in fact the only product of these two plays a role in the equations. For the product we see that a value of 50 mm approximately end the response range as only a few rain even in out series exceed this amount.

Modifying the Maximum Subsoil Infiltration Rate (MaxInfSubSoil) has a big impact on the amount of cumulative baseflow in the 0 – 5 mm day $^{-1}$range, but not beyond it . In this model MaxInfSubSoil determines to deep infiltration and percolation that lead to groundwater storage and base flow , but its impact is constrained by the maximum surface infiltration rate in combination with field capacity. At the threshold these other processes become 'bottle necks'.

The amount of cumulative baseflow also responds to changing the parameter value of Field Capacity. Soil Quick flow (SoilQFlow) changes relatively strong when the Field Capacity increases from 0 – 300 mm, but beyond 600 mm value there is no significant change. SoilQFlow arises when water can enter the soil, as 'saturated soil capacity' is not yet fully met and surface infiltration allows flow but can not be held for more than a day by the Field Capacity setting, nor can it drain fast enough to the ground water.

The Groundwater Release Fraction (GWRelFrac) has a big impact for cumulative base flow and Soil Quick Flow at range 0 – 0.1 day $^{-1}$. The Percolation Fraction Multiplier (PerFracMultuplier) at range 0.5 – 1 day $^{-1}$ will change the amount of base flow, but not beyond this value, as other in the processes become limiting.

6. Example of model application

6.1 GenRiver application for Way Besai - Sumberjaya (Lampung, Indonesia)

We used Sumberjaya as default condition for the model. We define the area into 15 subcatchments with total area 404 km^2. Fraction of land cover type for the catchment in year 3 and 20 in time series is 58 and 14% for forest, respectively, 22 and 11% for cropland and pioneer stages of fallow vegetation and 12 and 70% for coffee gardens, respectively.

We would to see how is the respond of river flow if we have three different types of rainfall (patchy, intermediate and patchy). For the same parameters setting that influence the shape of the rising stage parts, the predicted hydrograph becomes smoother when we shift from homogenous to patchy rainfall (Figure 16). While mean and range in the simulations are closed to measured ones, there is a tendency to over predict the lower flow rates, so our estimate of storage capacity may be too high yet.

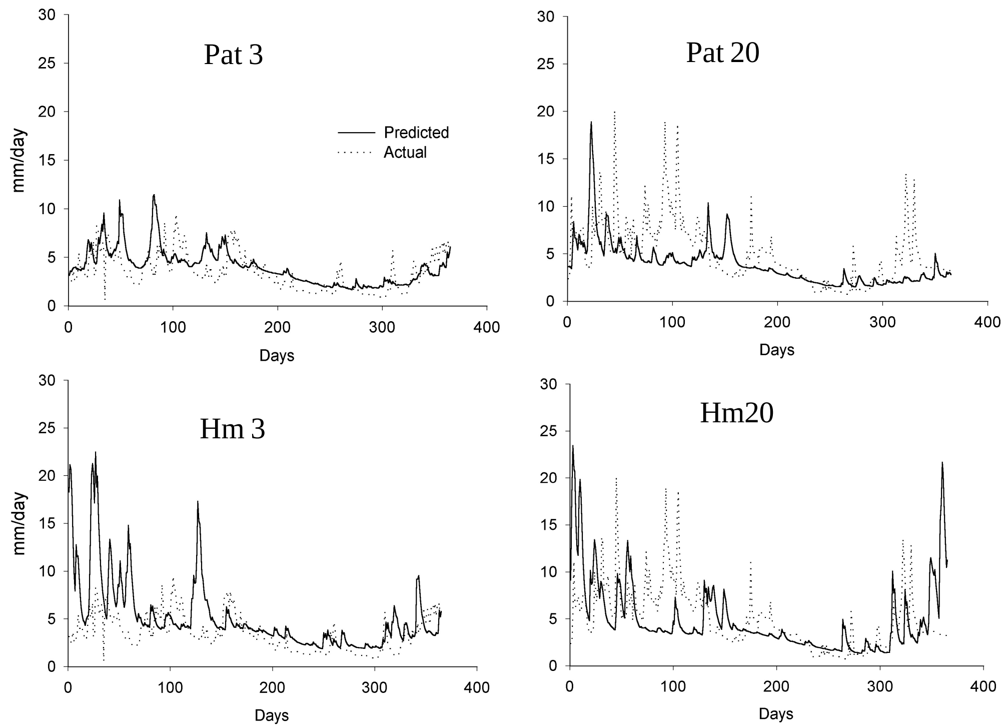


Figure 17. Predicted (lines) and measured (dots) river flow of Way Besai (expressed in mm day $^{-1}$ in year3 and 20 for the homogenous, intermediate and patchy rainfall pattern.

River flow of Way besai river in Lampung is consistent with GenRiver simulations for highly dissaggregated rainfall pattern, with a relatively small impact of the drastic land use change of the last two decades.

6.2 An Application of GenRiver for Mae Chaem (North Thailand)

We tried to set up this model using Mae Chaem data for further test. We use recorded rainfall data for Mae Chaem but we still using others default input parameter (based on Sumberjaya condition) because we have not yet have specific result for some data such as routing distance, subcatchment area and other properties related to soil sondition.

Mae Chaem watershed cover about 40,000 km^2 and 140 subcatchments are delineated from the digital elevation model (DEM). Most of the area cover by decidous forest (43%) followed by degradated forest (17%) and evergreen forest (11%).

Existing land cover in Mae Chaem area :

Existing land cover	Area (Sq km)	Area (%)
Urban	19.559	0.509
Field - Crop	24.453	0.624
Scrub	3.036	0.077
Scrub - Crop	1.998	0.051
Paddy	61.763	1.576
Paddy-shifting	18.05	0.461
Fallow	319.68	8.156
Orchard	142.042	3.624
Evergreen forest	142.042	3.624
Decidous forest	512.226	13.068
Degradated forest	1724.084	43.986
Hill pine forest	671.314	17.127

Total	3919.59	100.000

For Mae Chaem condition we have tried three different conditions:

a. Current land cover distribution
b. All evergreen forest
c. All deciduous forest

a. Current land cover distribution

Using the Mae Chaem rainfall data and others parameters as default we have a result of predicted river flow :

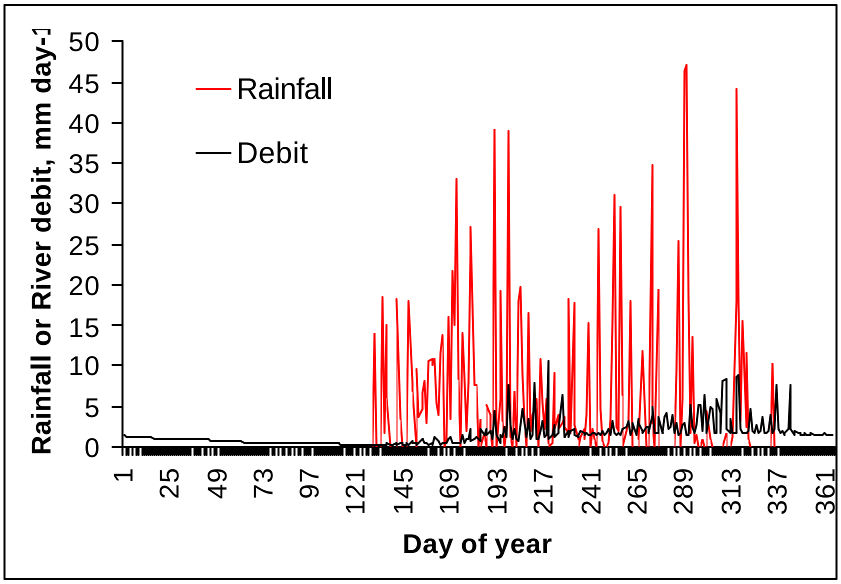


Figure 18. Predicted of the hydrograph indicating river debit.

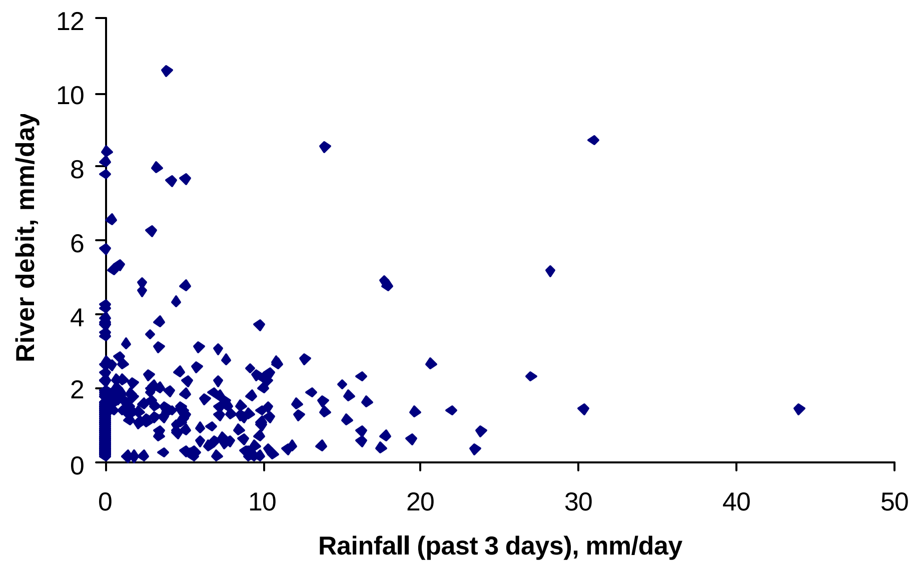


Figure 19. The model does not predict a simple relationship between rainfall (here represented as the average over the past 3 days) and river flow, as the differential storage capacity of the catchment between early and late rainy season plays a major role.

The overall water balance estimates :

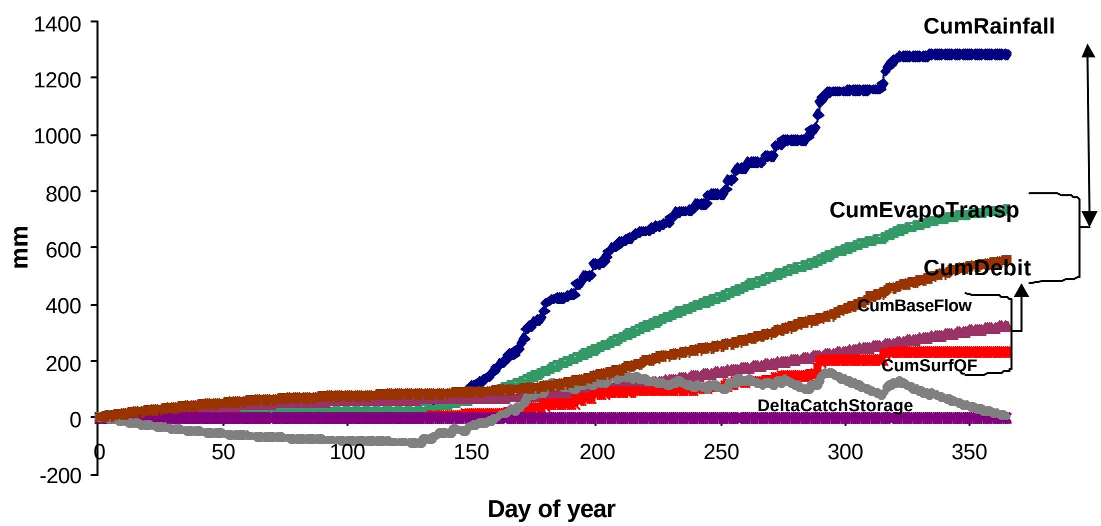


Figure 20. First estimated cumulative water balance for Mae Chaem subcathment as output of GenRiver model with site rainfall data, but otherwise using deafult parameters setting. The cumulative rainfall at the end of the year equals to the sum of evapotranspiratiom and river debit, but during the year a change in catchment storage can allow for river flow to be out of the phase of rainfall. River debit is sum of the baseflow and surface quick flow, as the third potential term, soil quickflow was predicted to be zero.

The overall water balance estimates as follow :

Annual rainfall (1285 mm) ~ evapotranspiration (730mm) + river debit (553 mm)

River debit (553 mm) ~ baseflow (324 mm) + quickflow (230 mm)

Small deviation from equality based on chages in water storage in the rivers and soil compartments

b). All Evergreen forest cover

Assuming the whole subcatchment to be under evergreen forest cover (with a relative potential evapotranspiration of 1 through the year) and using the default parameter set, the GenRiver model predicts a drastically different water balance.

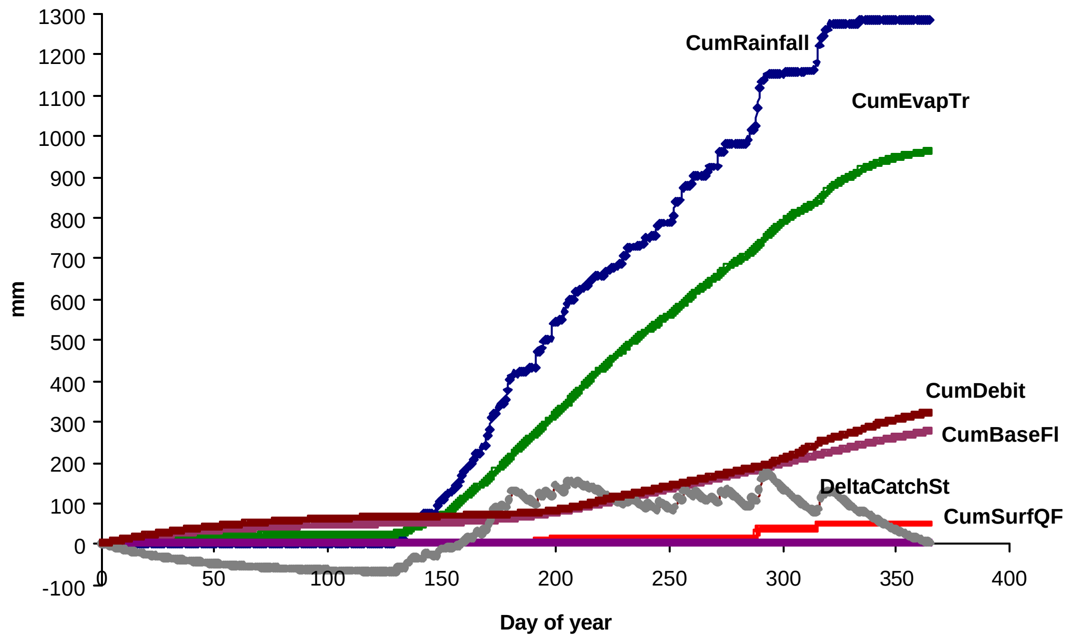


Figure 21. As figure 16?, but under assumption of complete forest cover for the whole subcatchment.

The annual water balance :

Annual rainfall (1285 mm) ~ evapotranspiration (961 mm) + river debit (322 mm)

River debit (322 mm) ~ baseflow (274 mm) + surface quick flow (48 mm)

If we take such a full forest cover as the basis for comparison, the model predicts that land use change up to the current land cover frequencies has decreases water use by vegetation by 231 mm year –1 and caused parallel increase in river debit, with an increase in baseflw 0f 50 mm year-1. Baseflow as a proportion of total riverflow has decreased from 85% (274 out of 322) to 59% (324 out of 553).

c). All deciduous forest

By imposing deciduousness forest on tree: i.e they flush new leaves on day 150 and shed leave on day 350 and have a relative potential evapotranspiration of 1 in between, we get the following result :

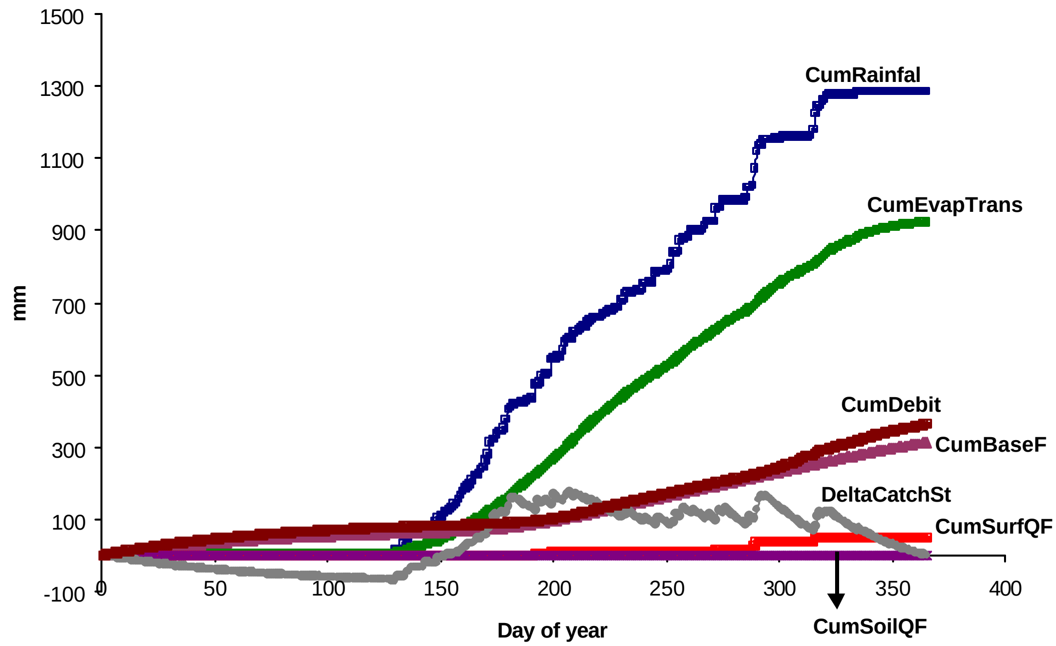


Figure 22. As figure 17, but by assuming the whole subcathment to be under deciduous forest cover.

The overall annual water balance :

Annual rainfall (1285 mm) ~ evapotranspiration (922 mm) + river debit (361 mm)

River debit (361 mm) ~ baseflow (314 mm) + surface quick flow (48 mm)

Of course, these initial model predictions critically depend on a number of parameters that need further collaboration. The water used by different types of land cover obviously are importance for Mae Chaem, more so than for Sumberjaya, as the same absolute change in water use represents a much larger share of the annual water balance.

7. SpatRain (Spatial Rainfall Distribution)

Variations in river discharge tend to decrease with increasing area of consideration, partly due to a decrease in temporal correlation of rainfall events across space. Patchiness of rainfall can contribute to an increase of yield stability over space. Existing rainfall simulators tend to focus on station-level time series, not on space/time autocorrelation. The SpatRain model described here was constructed to generate time series of rainfall that are fully compatible with existing station-level records of daily rainfall, but yet can represent substantially different degrees of spatial autocorrelation. Calculations start from the assumed spatial characteristics of a single rainstorm pathway, with a trajectory for the core area of the highest intensity and a decrease of rainfall intensity with increasing distance from this core. The model can derive daily amounts of rainfall for a grid of observation points by considering the possibility of multiple storm events per day, but not exceeding the long-term maximum of observed station-level rainfall. Options exist for including elevational effects on rainfall amount. SpatRain is implemented as an Excel workbook with macros that analyze semi-variance as a function of increasing distance between observation points, as a way to characterize the resulting rainfall patterns accumulated over specified lengths of time (day, week, month, year).

The SpatRain model starts from the spatial characteristics of a single rainstorm pathway (with a trajectory for the core area of the highest intensity and a decrease of rainfall intensity with increasing distance from this core) and can derive daily amounts of rainfall for a grid of observation points by considering the possibility of multiple storm events per day. Design features include:

- the simulated rainfall for any point in the landscape must be consistent with existing data on the frequency distribution of daily rainfall;
- the program must allow for spatial trends in mean rainfall, *e.g.* due to elevational effects;
- the program should analyze semivariance as a function of increasing distance between observation points, as a way to characterize the resulting rainfall patterns accumulated over specified lengths of time (day, week, month, year) and identify the storm-level parameters that lead to specified degrees of spatial correlation; and
- for use in combination with a hydrological model, SpatRain should allow for the identification of subcatchments in a watershed area and allow averaging the point grid pattern to derive the daily average rainfall per subcatchment.

7.1 Approach to the problem

Station-level daily records are often the only information available on the distribution of rainfall. Such data can be represented as a series of monthly 'exceedance' graphs, derived from say 30-year data. Between months of the year and locations we may expect differences in the intercept with the X-axis or 'frequency of wet days' (or days with a measurable amount of precipitation, usually defined as > 0.5 mm day^{-1}), the intercept with the Y-axis or maximum amount of rainfall in a single day recorded in that particular month of the year, and in the curvature of the (monotonously rising) line between these two points (Fig. 23).

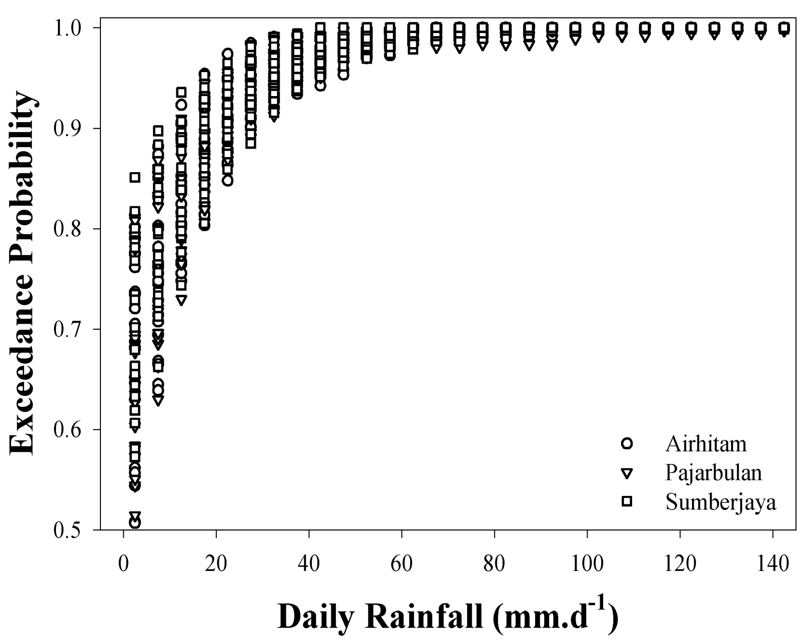


Figure 23. Input distribution of station level rainfall depth in three measurement stations that were combined, as they don't show essential differences.

Conceptually one can imagine a procedure that reshuffles measured daily sequences of rainfall while maintaining the monthly total, like rearranging a jackpot, where the variability of

3 values exposed on the window should follow our expectation: homogenous (apple-apple-apple) or heterogeneous (apple-banana-orange). The total set of permutations of 30 sets of jackpot with 30 pictures (for 30 days and only 30 cells) is enormous, and among these we can expect to find a substantial variation in degrees of spatial autocorrelation. By generating a sample of these reshuffling results, calculating autocorrelation and then selecting specific configurations, we would meet the key design criteria specified above. The program will, however, be rather cumbersome and time consuming if large areas are to be considered, and the selection of results that meet a specific change in spatial autocorrelation with increasing distance may require a large subset of reshuffling results. More efficient algorithms are desirable, but the 'jackpot' analogue shows that the design rules are not mutually incompatible. A more direct approach can be taken if we assign specific spatial properties to single storm events and then adjust the frequency of storms and the intensity of rainfall in the core area of these storms to match the existing station records.

7.2 Assumed storm properties

Three parameters are used here for describing rainfall in the core area: the length of the core trajectory, the radius of the core area and the rainfall depth in the core area (Fig. 24). Two further parameters describe the relative decrease of rainfall depth with increasing distance from the core. The combination of these can produce the full scala of 'homogenous' to 'heterogeneous' types of rain. These parameters can be related to frictional forces forming thunderstorms or convective bands causing frontal circulation (Pielke, 2002):

$$I_d \; ? \; I_0 * (1 ? e^{?(f^{?}/d)^{f^{?}}})$$

where:
d is distance of a cell from the storm core (grid unit);
I_d is rain intensity of a cell at distance d from the core ($mm.d^{-1}$);
I_0 is rain intensity at the core ($mm.d^{-1}$);
$f^?$ is spreading factor; and
$f^?$ is agglomerating factor.

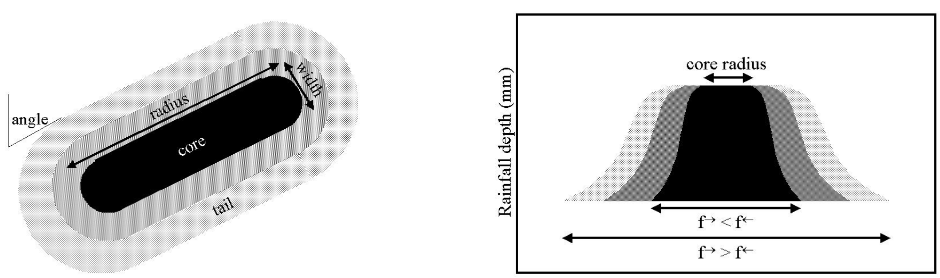


Figure 24. Assumed shape of individual storm events. Patchy rains are possibly formed when $f^? < f^?$ (frictional forces), while homogenous rains are formed when $f^? > f^?$ (frontal dynamics).

7.3 Matching spatial pattern with temporal pattern

A single storm event will 'wet' (above the measurement threshold of 0.5 mm day^{-1} used in most empirical data sets) a number of cells, some at the core intensity and some at a lower intensity. Given a set of parameters for the storm trajectory, we can derive the frequency distribution of rain depth in wetted cells, relative to the core rain intensity (p), in n classes. Once this is known, the frequency distribution of core intensities (F) can be derived from the observed station level rain intensities (f). Frequency distributions of f, p and F should have the same class number and interval order. We use the following order to define the class boundary: [max..max*q^1],[max*q^1.. max*q^2], [max*q^2.. max*q^3],…,[max*q^n.. min], where max is the maximum data, min is the minimum data and n is class intervals number. The value of q is ranging from 0 to 1 and calculated as follows:

$$q = e^{\ln(min/max)/n}, q = 0 - 1$$

We first need to recognize the combinations of classes p_j and F_k that are compatible with class f_i:

$$f_i = \sum_{j,k \to i} \{ p_j F_k \mid j \& k \sim i \}, 0 \le f_i \le 1, 0 \le p_j \le 1, 0 \le F_k \le 1$$

For the highest rainfall class only one combination, involving the highest class of both p and F will yield the desired result, but for the other classes there can be several combinations of p and F that yield the same result (the tail end of a big rainfall event, a medium fraction of a medium storm or the core area of a small storm). We can approach it working our way from the top down, but a simpler derivation starts from the observation that for all distributions f, p and F the sum equals 1. By assuming that the resultant (f) comes from the multiplication between p and F, we then get this basic equation:

$$\sum_{i=1}^{n} f_i = \sum_{j=1}^{n} p_j * \sum_{k=1}^{n} F_k$$

So that F of frequency class n can be defined as:

$$F_n = \frac{\sum_{i=1}^{n} f_i}{\sum_{j=1}^{n} p_j} - \sum_{k=1}^{n-1} F_k$$

From the equation, we can derive a criterion for the shape of the p distribution (that depends on assumed storm properties) that is compatible with the targeted f distribution. If at any point $\frac{\sum_{i=1}^{n} f_i}{\sum_{j=1}^{n} p_j}$ is less than $\sum_{k=1}^{n-1} F_k$, F_n would violate the assumption of non-negative subsequent F terms. So, a cross-over of p and f indicates incompatibility of the storm-level assumptions — that generate the p curve — with the station-level rainfall records — that generate the f curve. Figure 2.44 illustrates the compatibility of intensity distribution from two contrasting spatial patterns of 30-grid maps with temporal distribution from 30-day station record. Pattern B of exactly similar distribution to the station record rainfall produces compatible F as shown in Fig. 2.44.D, whereas pattern C is incompatible with the station record distribution as indicated by negative values of F in Fig. 2.44.E. This means, it is impossible to arrange rainfall maps of pattern C using the existing temporal distribution.

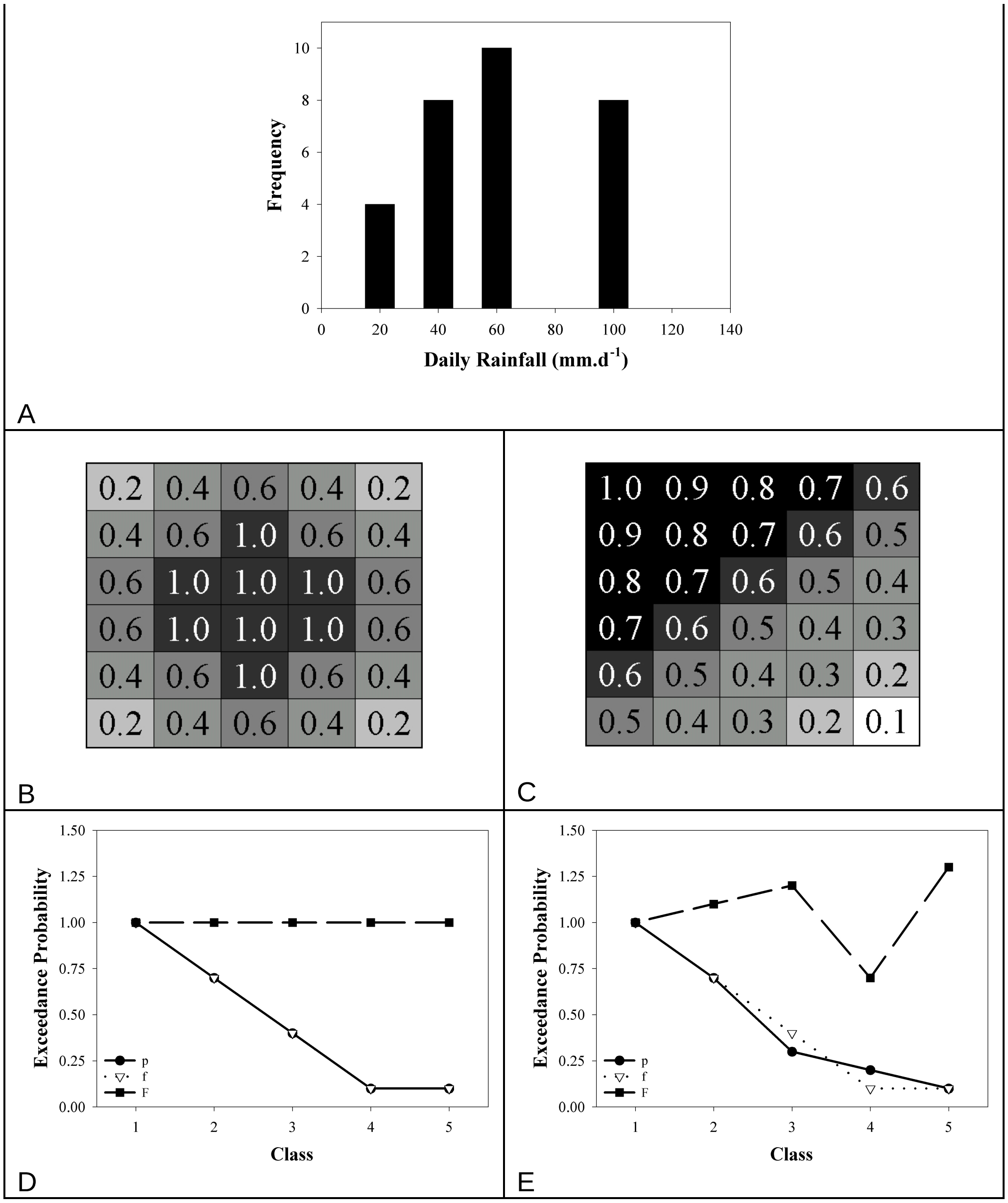


Figure 25. Compatibility of intensity distribution from two spatial patterns (B and C) with temporal distribution from station records (A). Pattern B is compatible (D) while pattern C is incompatible (E).

7.4 Considering multiple storm events

Equation $I_d ? I_0 * (1 ? e^{?(f^{?}/d)^{f^{?}}})$ may produce a narrow-spread area of single storm events, on which its wet cells ratio relative to total area (c_1) does not match the wet days fraction of that specific month (d). Hence, we need to allow for multiple storm events, depending on the area fraction wetted by a single event and the time fraction of rainy days at the measurement station level. For spatially independent multiple events on a single day we can derive that probability of dry days on a given month, P|?|, should meet the probability of dry cells during single event, P|? |, to the power of events number (N):

$$P|d| ? P|?|^{N}$$

Where P|?|=1-d and P|?|=1-$c_{1.}$ Thus, the number of events is:

$$N ? \frac{\ln(1 - d)}{\ln(1 ? c_1)}$$

7.5 Cross-scale probability of storm events

Patchy rains have less wet fraction than homogeneous rains in space. In order to conserve each cell to having uniform chance of being hit by storms in time, patchy rains should have higher probability to occur than homogeneous rains. Consequently, the cross-scale probability of storm with N number of events ($P(E_N)$) is defined from wet days fraction (d) by taking wet cells fraction of N storm events (c_N) into account:

$$P ?E_N ? ? \frac{d}{c_N}$$

7.6 Considering elevational effect

Rainfall patchiness can also be affected by elevational effects of the area. Thus, rainfalls at particular degree of patchiness generated by the above procedures should be corrected if applied on an area with elevational effects. The elevation modifier of rainfall at elevation z (X_z) is assumed as rainfall average at that elevation ($?_z$) relative to overall average (?):

$$X_z ? \frac{\mu_z}{\mu}$$

In fact we are modifying the amount of rain that any storm brings to any cell, not the preferred pathway of storm trajectories. Though similar multipliers we can introduce 'rain shadow' effects that depend on a preferential direction of storms and gradients in elevation.

7.7 Patchiness indicator

Semivariogram is used as quantitative spatial pattern indicator of simulated rainfall (Fig. 2.45). Spatial distribution of rain intensity from the storm cores can be distinguished by semiv ariance increase (dS) within the distance range of increasing semivariance (dh) or the slope (s=dS/dh). From Fig. 2.45, it is expected that patchy rains will have higher dS within shorter dh (steeper slope) than homogeneous rains. Moreover, based on beha vior of the slope, patchiness can be quantified using fractal dimension (D) (Bian,1997):

$$D ? 3 ? (s/2)$$

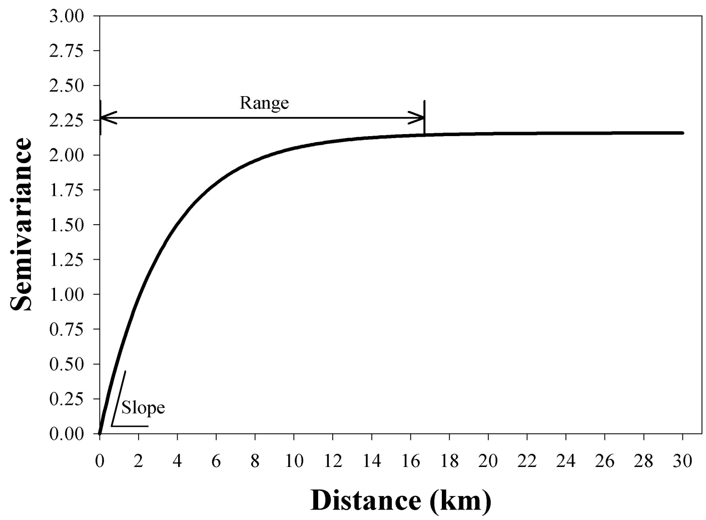


Figure 26. Fractal dimension as defined by slope of semivariance range is used as patchiness indicator of rainfall.

The value of D ranges from 0 to 3. Lower fractal dimension of a spatial pattern may be interpreted as more fragmented pattern. Thus, patchy rain is expected to have lower D than the homogeneous.

7.8 Implementation in SpatRain

A flowchart of the program that implements the above conceptualization is shown in Fig. 2.46. The SpatRain simulator is freely available on our website (http://www.worldagroforestrycentre.org/sea/products/AFmodels/spatrain.htm). The current version of the program is developed using VB macro in an Excel workbook. Application to the Mae Chaem area at a 3 km^2 grid cell resolution proved to be at the edge of the program's capability. To overcome the memory limitations, a standalone version of SpatRain has been developed using Java programming language. Fig. 2.47 shows one of the SpatRain-Java environment features in displaying the dynamic maps (daily rainfall maps as the simulation outputs), static scalar maps (*e.g.* DEM) and static discrete maps (*e.g.* sub-catchments boundary) at better resolution of 1 km^2 grid cell.

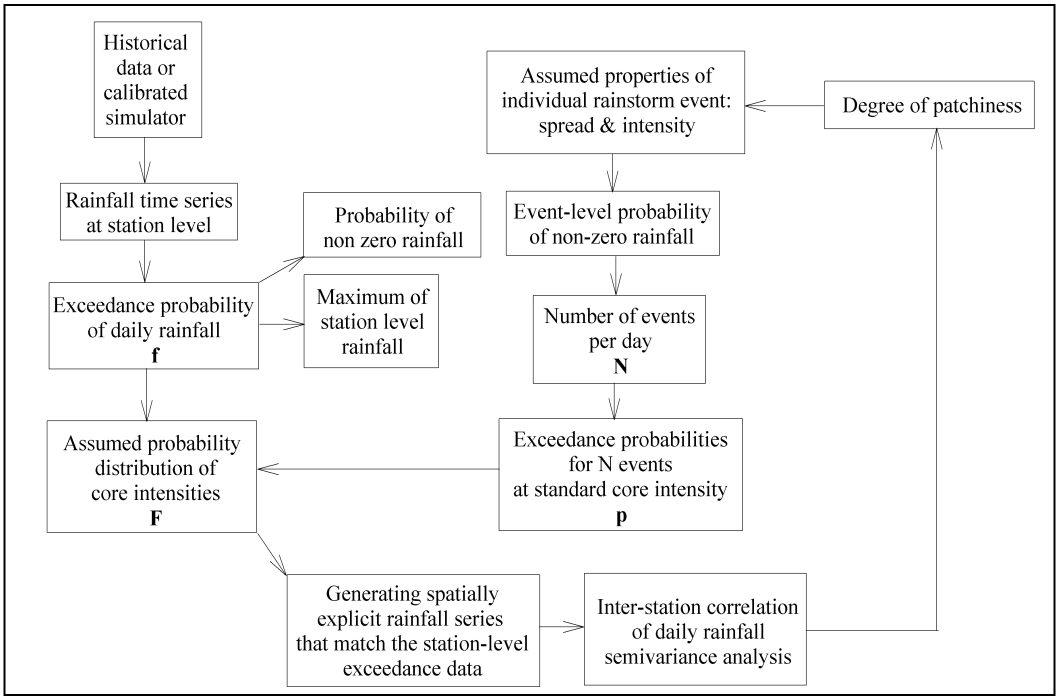
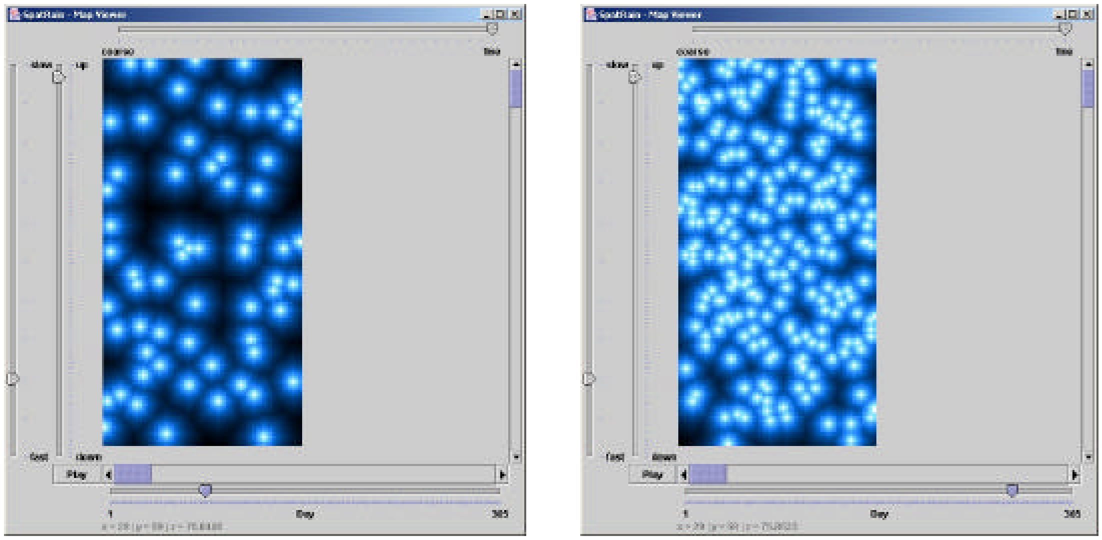
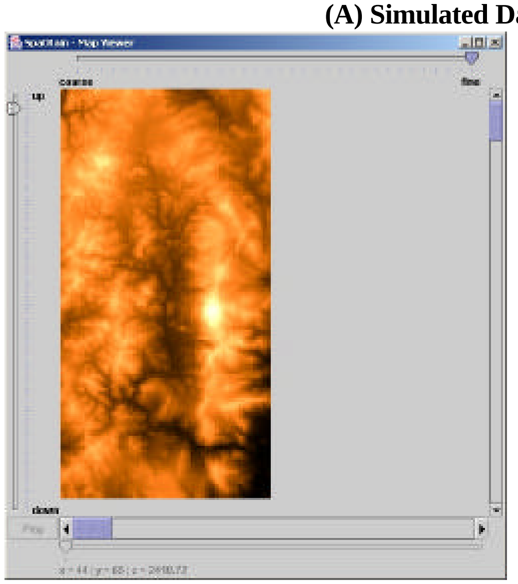
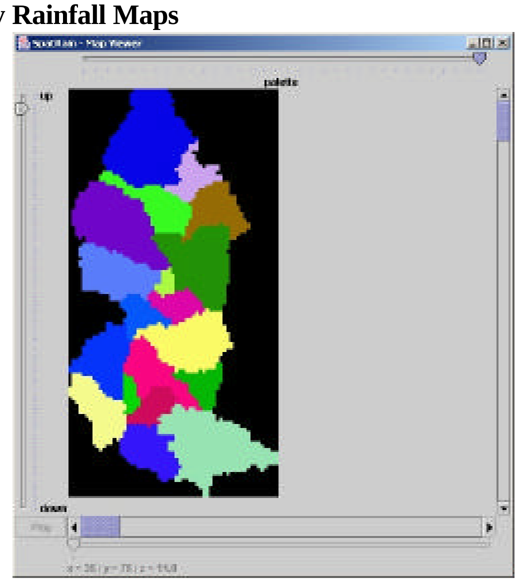


Figure 27. Flow diagram of model calculations in SpatRain



(A) Simulated Daily Rainfall Maps





(B) DEM **(C) Sub-cathment Boundary Map**

Figure 28. SpatRain-Java capability in handling maps with large number of grids: (A) dynamic maps; (B) static scalar map (C) static discrete map.

Acknowledgements

ACIAR and Barry Croke (Australian National University) for valuable input and advice for the manual.

Rerefences

Bian, L. 1997. Multiscale Nature of Spatial Data in Scaling Up Environmental Models, In: Quattrochi, D.A. and Goodchild, M.F. (Eds.): Scale in Remote Sensing and GIS, Lewis Publishers: pp: 20

Calder, I.R., 2002. Forests and hydrological services: reconciling public and science perceptions. Land Use and Water Resources Research 2, 2.1-2.12 (www.luwrr.com)

Kiersch, B. and Tognetti, S., 2002. Land-water linkages in rural watersheds. Land Use and Water Resources Research 2, 1.1-1.6(www.luwrr.com)

Pielke, Sr. R. A. 2002. Mesoscale Meteorological Modeling. 2nd edition. Academic Press, pp: 676

van Noordwijk, M, Farida, A., Suyamto, D., Lusiana, B. and Khasanah, N. (2003) Spatial variability of rainfall governs river flow and reduces effects of land use change at landscape scale: GenRiver and SpatRain simulations. In: Post, D. (ed.) Proceedings of the 2003 MODSIM Conference, July 2003, Townsville, Australia, pp: 572-577

van Noordwijk, M. and Lusiana, B., 1999 WaNuLCAS, a model of water, nutrient and light capture in agroforestry systems. Agroforestry Systems 43: 217-242