What is Υετός (Hyetos)?

Hyetos is a model that performs **disaggregation** of daily rainfall into hourly rainfall. It uses the **Bartlett-Lewis rainfall model** as a background stochastic model for rainfall generation. Then it uses **repetition** to derive a synthetic rainfall series, which resembles the given series at the daily scale, and, subsequently, an appropriate **adjusting procedure**, namely the proportional adjusting procedure (for its choice see relevant topic), to make the generated hourly series fully consistent with the given daily series.

Although the model implementation is specified for the daily higher-level scale and the hourly lower-level scale, the methodology can be used for other, coarser or finer, time scales as well.

Here the **general model characteristics** and its **implementation** are summarised along with some **hints** for the user. For more detailed information, there is a list of **References**.

In addition to the brief theoretical documentation of the model, there is a general description of the software regarding its **modes of operation**, its user interface (**Main form**, **Bartlett-Lewis model parameters form**, **Repetition options form**, **Graphs form**, and **Visual output form**) and the input and output **files format**.

Why disaggregation?

Detailed hydrological models most often use the hourly time scale, and also require inputs at this scale. However, historical **hourly records** are not widely available as most rain gauge stations provide **daily records**. An appropriate disaggregation model could generate a synthetic hourly series, so as to be **fully consistent with the known daily series** and, simultaneously, **statistically consistent with the actual hourly rainfall series**. The actual rainfall series is not known, but its stochastic structure could be hypothesized according to a specific model, which in our case is the **Bartlett-Lewis model**. Obviously, a synthetic series obtained by a disaggregation model could not coincide with the actual one, but is a likely realization of rainfall.

Another situation where disaggregation is useful is in **climate change studies**. Usually, the output of General Circulation Models (forecasts for different climate change scenarios) is generally provided at a coarse time-scale (e.g. monthly) whereas hydrological applications require a finer time scale.

In addition, there exist situations, such as in **flood studies**, where an observed rainfall series is not the one that will be used in hydrological models. Events more severe than the observed ones must be synthesised and used. Often in such cases, the total characteristics (duration, depth) of the events are known and these must be disaggregated into e.g. an hourly scale.

Furthermore, in **simulation studies**, although rainfall records may be available at the desired (e.g., hourly) scale, it could be necessary to derive and study the system using multiple (rather than the single observed) sequences of rainfall series all having some common cumulative properties. Again, disaggregation is helpful.

See also:

General model characteristics

Synopsis of disaggregation by adjusting procedures

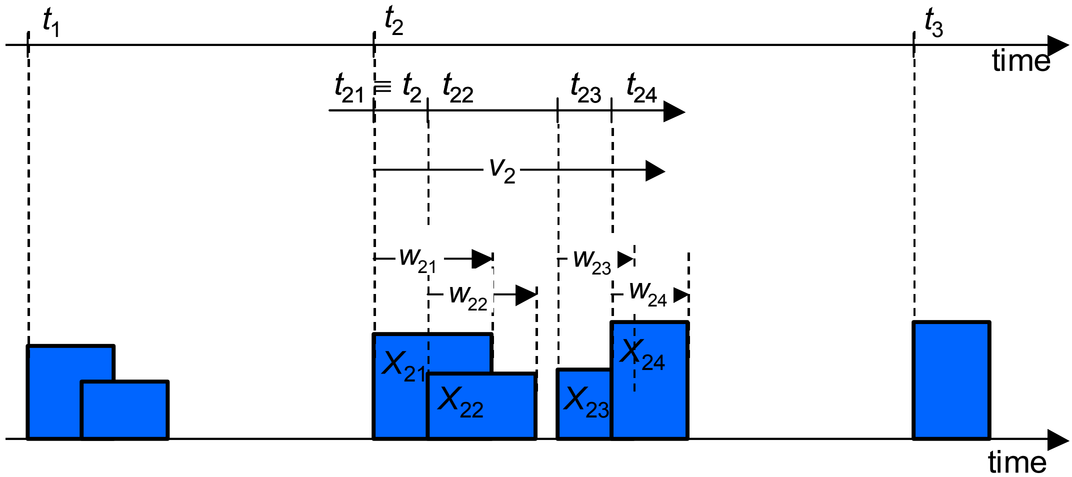
General model characteristics

Many disaggregation models of the literature are ad hoc models designed as such from the beginning. On the contrary, **Hyetos** combines an existing typical rainfall simulation model along with an appropriate technique for modifying the rainfall model output, thus performing disaggregation.

As an appropriate rainfall model, the **Bartlett-Lewis model** was chosen due to its wide applicability and experience in calibrating and applying it to several climates. (Rodriguez-Iturbe et al., 1987, 1988; Onof and Wheater, 1993, 1994)

Techniques for **disaggregation by adjusting**, i.e., for modifying a fine scale (lower-level, such as hourly) time series, generated by a specific stochastic model, so as to be consistent with a given coarser scale (higher-level, such as daily) time series, and simultaneously not affecting the stochastic structure implied by the model, have been studied by Koutsoyiannis (1994) and Koutsoyiannis and Manetas (1996).

Synopsis of the Bartlett-Lewis (BL) point process model



The general assumptions of the Bartlett-Lewis Rectangular Pulse model (Rodriguez-Iturbe Et Al., 1987, 1988; Onof and Wheater, 1993) are (see figure):

1. Storm origins t_i occur following a Poisson process (rate λ)
2. Cell origins t_{ij} arrive following a Poisson process (rate β)
3. Cell arrivals terminate after a time v_i exponentially distributed (parameter γ)
4. Each cell has a duration w_{ij} exponentially distributed (parameter η)
5. Each cell has a uniform intensity X_{ij} with a specified distribution

In the **original** version of the model, all model parameters are assumed constant. In the **modified** version, the parameter η is randomly varied from storm to storm with a gamma distribution with shape parameter α and scale parameter v. Subsequently, parameters β and γ also vary in a manner that the ratios $\kappa := \beta / \eta$ and $\varphi := \gamma / \eta$ be constant.

The distribution of the uniform intensity X_{ij} is typically assumed exponential with parameter $1 / \mu_X$. Alternatively, it can be assumed two-parameter gamma with mean μ_X and standard deviation σ_X.

Thus, in its most simplified version the model uses five parameters, namely λ, β, γ, η, and μ_X (or equivalently, λ, κ, φ, η, and μ_X) and its most enriched version seven parameters, namely λ, κ, φ, α, v, μ_X and σ_X.

Hyetos supports both the original and the modified model version with exponential or gamma intensities. To implement the original version, in the Parameters form set α greater than or equal to100 (otherwise the modified version is assumed). To implement the exponential distribution of X_{ij} in the Parameters form set $\sigma_X = \mu_X$.

Synopsis of disaggregation by adjusting procedures

Provided that a data series is known at a higher-level time scale (e.g., daily) and a lower-level (e.g. hourly) synthetic series has been generated by some stochastic model, disaggregation by adjusting procedures is a methodology to modify the lower-level series so as to make it consistent with the higher-level one. To this aim, it uses accurate adjusting procedures to allocate the error in the additive property, i.e., the departure of the sum of lower-level variables within a period from the corresponding higher-level variable. These procedures are accurate in the sense that they preserve explicitly certain statistics or even the complete distribution of lower-level variables. In addition, the methodology uses repetitive sampling in order to improve the approximations of statistics that are not explicitly preserved by the adjusting procedures.

Three such adjusted procedures have been developed and studied (Koutsoyiannis, 1994; Koutsoyiannis and Manetas, 1996). Here are some of their more important properties

1. **Proportional adjusting procedure**

   - It modifies the initially generated values $\tilde{X}_s$ to get the adjusted values X_s according to

     $$X_s = \tilde{X}_s \left(Z / \sum_{j=1}^{k} \tilde{X}_j \right) \qquad (s = 1, \ldots, k) \quad \ldots\ldots\ldots\ldots (1)$$

     where Z is the higher-level variable and k is the number of lower-level variables within one higher-level period.

   - It is the simplest in application.
   - It is exact for complete preservation of distributions if variables X_s are independent with two parameter gamma distribution and common scale parameters.
   - It also provides good approximation for dependent variables with gamma distribution.
   - It has the advantage of not resulting in negative values X_s.

2. **Linear adjusting procedure**

   - It modifies the initially generated values $\tilde{X}_s$ to get the adjusted values X_s according to

     $$X_s = \tilde{X}_s + \lambda_s \left(Z - \sum_{j=1}^{k} \tilde{X}_j \right) \qquad (s = 1, \ldots, k) \quad \ldots\ldots\ldots\ldots (2)$$

     where λ_s are appropriate coefficients depending on the covariances of X_s with Z.

   - It is exact for complete preservation of distributions for Gaussian (dependent or independent) variables.
   - It is exact in preserving second order moments of (dependent or independent) variables with any distribution.
   - It can explicitly assure preservation of correlations with lower-level variables of different higher-level periods (a more recent development).
   - Its main disadvantage is that it may result in negative values (which can then be corrected using repetitions)

3. **Power adjusting procedure**

   - It modifies the initially generated values $\tilde{X}_s$ to get the adjusted values X_s according to

$$X_s = \widetilde{X}_s \left(Z / \sum_{j=1}^{k} \widetilde{X}_j \right)^{\lambda_s / \eta_s} \qquad (s = 1,\ldots,k) \ldots\ldots\ldots\ldots (3)$$

where λ_s are appropriate coefficients depending on the covariances of X_s with Z and η_s are coefficients depending on mean values of X_s and Z.

- It is approximate apart from special cases where it is exact (e.g., when it coincides with proportional procedure).
- It needs repetitions.
- It does not result in negative values
- For stationary processes it is identical to the proportional procedure.

See also: Notes on the choice of the appropriate adjusting procedure

Notes on the choice of the appropriate adjusting procedure

These conditions hold in the problem examined:

- In the rainfall disaggregation problem, the process can be assumed stationary (within a specific period, e.g., month) and thus the power adjusting procedure is identical to the proportional one.
- The problem is characterised by a large proportion of zeros (e.g., up to 90% within rainy days). This creates difficulties if the adjusting procedure does not prohibit negative values, as all zero values can become negative after the adjusting.
- The rainfall depths in rainy intervals can be assumed approximately gamma distributed.
- The autocorrelation of rainfall is not very strong in the typical time scales.

Under these conditions, the most appropriate among the three adjusting procedures is the **proportional adjusting procedure**. This was implemented in the model.

In addition, repetition is required to better preserve the autocorrelations (especially those between different days) and skewness. The main source of bias dictating the need for repetition in the rainfall disaggregation is the variable number of nonzero intervals within any period.

See also: Some hints to the user

Notes on the implementation of the rainfall model

The main potential problem in the implementation of the methodology is that it can be computer time consuming, if applied over a long simulation period. To avoid this, the simulation period must be separated is as many subperiods as possible.

Different sequences (clusters) of wet days (separated by at least one dry day) can be assumed independent. This empirical observation is consistent with the Bartlett-Lewis model, which assumes Poisson arrivals of storms.

This allows different treatment of each cluster of wet days, which reduces computational time rapidly.

This scheme is implemented in the model. More specifically, the Bartlett-Lewis model runs separately for each cluster of wet days. Several runs are performed for each cluster, until the departure of daily sum from the given daily rainfall becomes lower than an acceptable limit.

In case of a very long sequence of wet days it is practically impossible to get a sequence of wet days with departure of daily sum from the given daily rainfall lower than the specified limit. In these cases the sequence is subdivided into sub-sequences, each treated independently of the others.

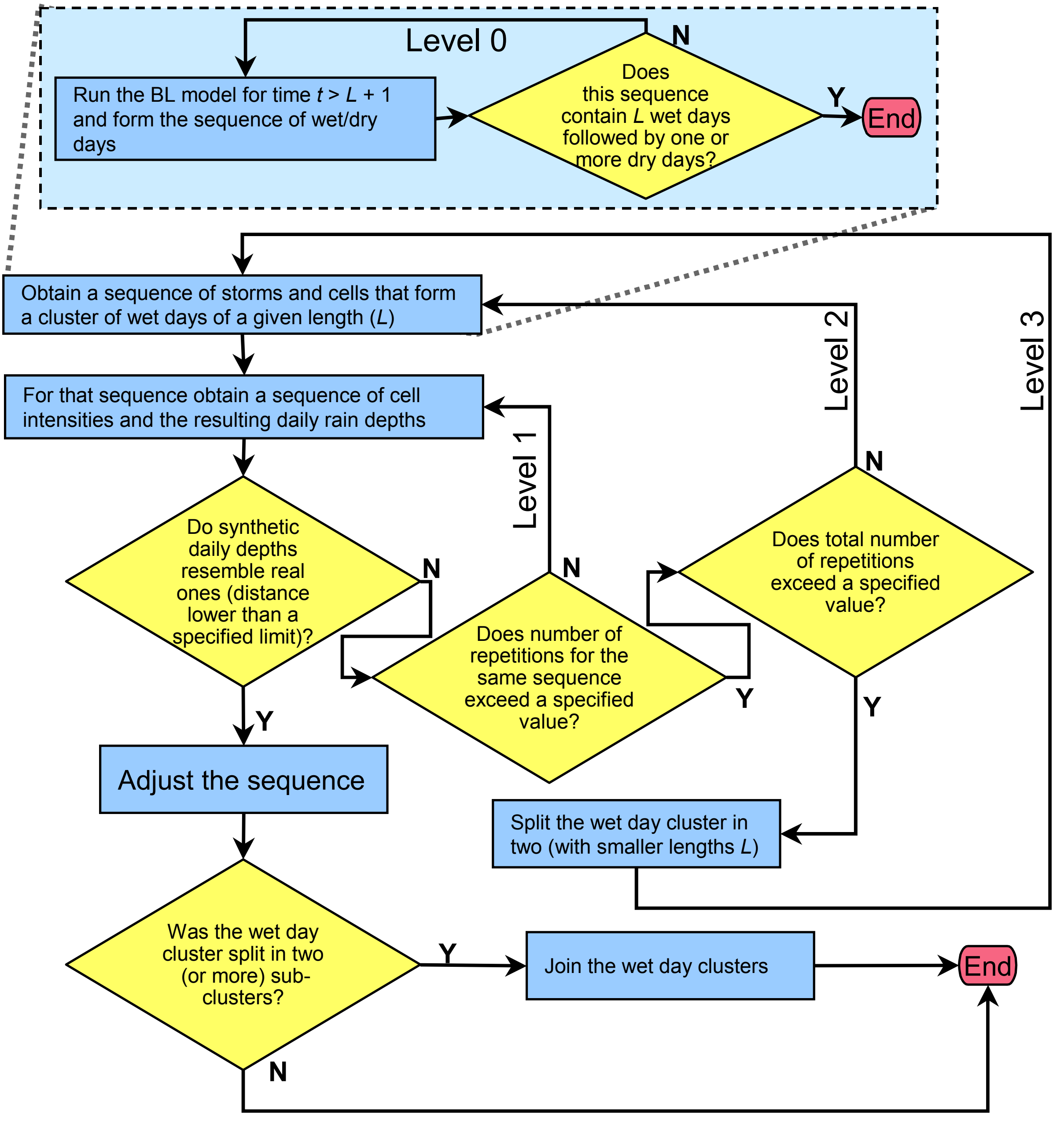
See also: **Repetition scheme**

Repetition scheme

The repetition scheme incorporates four levels of repetition, as shown in the figure below. The *distance* used to judge whether synthetic daily depths resemble the real ones is determined by

$$d = \left[\sum_{i=1}^{L} \ln \left(\frac{Z_i + c}{\widetilde{Z}_i + c} \right)^2 \right]^{1/2} \quad \text{............ (4)}$$

where d is the distance; Z_i and $\widetilde{Z}_i$ are the original and generated, respectively, daily depths of day i of the wet day sequence; L is the length of wet day sequence; and c = 0.1 mm. The adjusting of the hourly sequence is done according to equation (1) (see Synopsis of disaggregation by adjusting procedures). The accepted limit of distance and the allowed number of repetitions are defined by the user using the Repetition options form.



Some hints to the user

Parameter estimation: The current software version does not support estimation of the Bartlett-Lewis model parameters, which should be estimated using other software.

Bias remediation: If in a model application the wet spells are very rare (or the probability of dry hours in wet days very high) then the bias in second order statistics may not be remedied even by increasing the number of repetitions or decreasing the allowed distance. In such situations, it is suggested to modify the BL model parameter σx (e.g., by trial-and-error).

Preservation of skewness: The Bartlett-Lewis model does not handle the skewness of the rainfall process and therefore the disaggregation model cannot explicitly preserve the skewness. However, there are some possibilities of adapting the resulting skewness of the simulated data by modifying the BL model parameters μx and σx (e.g., by trial-and-error).

See also:

Synopsis of the Bartlett-Lewis (BL) point process model

Bartlett-Lewis model parameters form

References

The methodology of **Hyetos** and details of its application are described in

- Koutsoyiannis, D., and C. Onof, Rainfall disaggregation using adjusting procedures on a Poisson cluster model (will be available in 2000).

For the **Bartlett-Lewis rainfall model** the user is referenced to

- Rodriguez-Iturbe, D. R. Cox, and V. Isham, Some models for rainfall based on stochastic point processes, Proc. R. Soc. Lond., A 410, 269-298, 1987.
- Rodriguez-Iturbe, D. R. Cox, and V. Isham, A point process model for rainfall: Further developments, Proc. R. Soc. Lond., A 417, 283-298, 1988.
- Onof, C. and H. S. Wheater, Modelling of British rainfall using a Random Parameter Bartlett-Lewis Rectangular Pulse Model, *J. Hydrol.*, 149, 67-95, 1993.
- Onof, C. and H. S. Wheater, Improvements to the modeling of British rainfall using a Modified Random Parameter Bartlett-Lewis Rectangular Pulses Model, *J. Hydrol.*, 157, 177-195, 1994.

For **disaggregation by adjusting** the user is referenced to

- Koutsoyiannis, D., A stochastic disaggregation method for design storm and flood synthesis, *Journal of Hydrology*, 156, 193-225, 1994.
- Koutsoyiannis, D., and A. Manetas, Simple disaggregation by accurate adjusting procedures, *Water Resources Research*, 32(7) 2105-2117, 1996.

Modes of operation

The model can perform in each of the following modes depending on the user selections:

1. **Disaggregation test mode** (without input; **default mode**). An initial sequence of storms is generated using the Bartlett-Lewis model with the given parameters and then aggregated into hourly and daily scale. The daily sequence serves then as an "original" series, which is disaggregated, thus producing another synthetic hourly series. This mode is appropriate for testing the disaggregation model itself (e.g. by comparing original and disaggregated statistics; the statistics can be either printed by pressing the *Statistics* button in the Main form or displayed graphically in the Graphs form).

2. **Full test mode** (with hourly input). To enter to this mode, an input file with the appropriate format containing **hourly historical data** must be defined using the program Main form. The difference from the **Disaggregation test mode** is that the daily sequence is read from the file rather than generated. This mode is appropriate for testing (e.g. by comparing original and disaggregated statistics) the entire model performance including the appropriateness of the Bartlett-Lewis model and its parameters and the disaggregation model.

3. **Operational mode** (with daily input). This is similar to **Full test mode** the difference being that the input file contains no hourly data but only **daily**. This is the usual case for the model application. It cannot provide any means for testing.

4. **Rainfall model test mode** (with hourly input). This is similar to the **Full test mode** but with synthetic data not disaggregated but generated from the Bartlett-Lewis model with the given parameters. To enter to this mode, (a) an input file with the appropriate format containing hourly historical data must be defined using the program Main form; (b) the checkbox *Perform adjustment* must be checked off from the Bartlett-Lewis model parameters form, and (c) the *Distance allowed* must be set equal to a large number (e.g. 1000) from the Repetition options form.

5. **Simple rainfall generation mode** (without input and without disaggregation). This is similar to the **Rainfall model test mode** but with no input provided (and consequently, no input file defined). This mode is appropriate for generation of rainfall series using the Bartlett-Lewis model with the given parameters without performing any disaggregation.

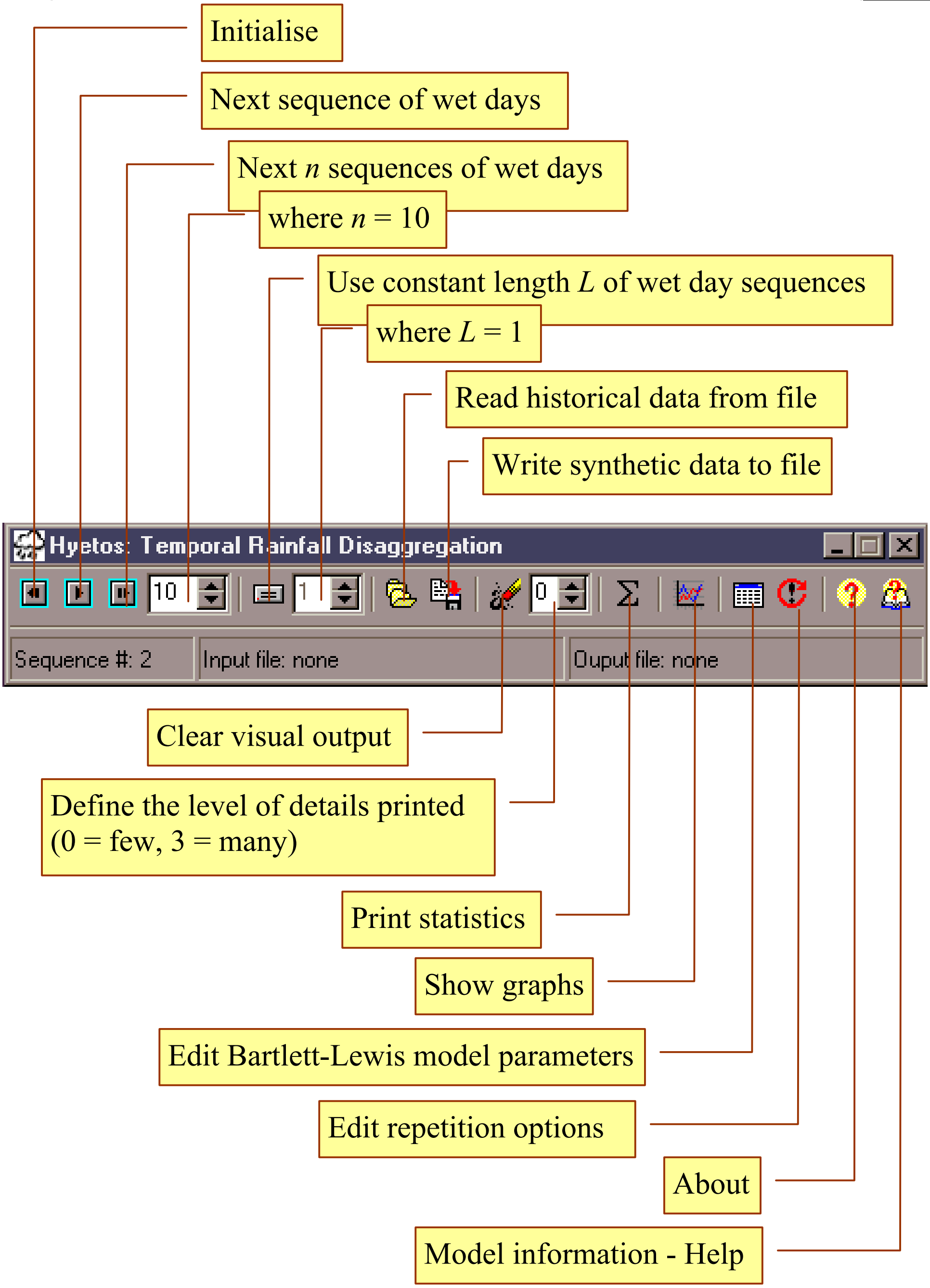
In all modes the Bartlett-Lewis model can be implemented either in its original or modified version with a number of parameters from 5 to 7.

Main form

This is the main form of **Hyetos**. Use the on-screen hints of toolbar buttons displayed by placing and pausing the mouse pointer on them. Here is a summary of toolbar buttons description.

The other forms of the software application, namely the Bartlett-Lewis model parameters form, the Repetition options form, the Graphs form and the Help About form appear by clicking the appropriate buttons of this form, whereas the Visual output form appears automatically with the program operation.

If original data must be read from a file, this file must have a specific format (see relevant topic).



Bartlett-Lewis model parameters form

Activate this form by pressing the appropriate button in the Main form.

Use this form to insert or modify the Bartlett-Lewis model parameters and some related options. Normally, one parameter sets corresponds to one month.

When the program starts, the default parameter set shown in the figure below is loaded (it corresponds to the month January for the Heathrow raingauge estimated by the spectral parameter estimation method).

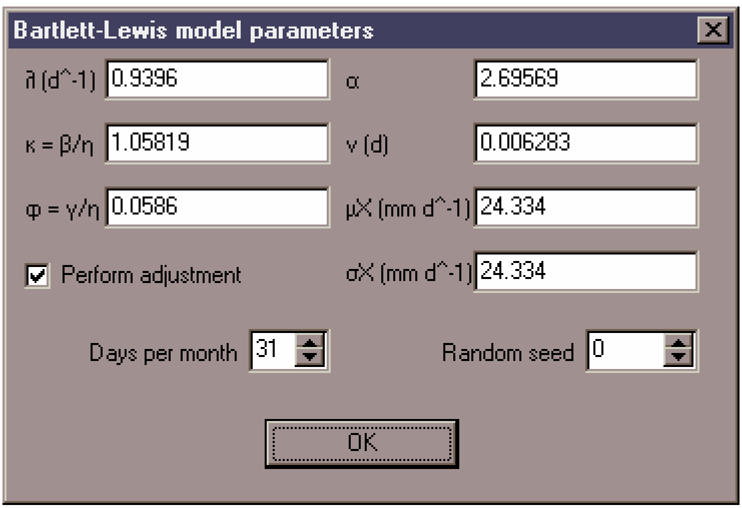
Whenever an input data file is defined, if this file contains a Parameters header (see related topic) the Bartlett-Lewis model parameters are determined from this file. However, by activating this form, the parameters can be edited as well, although all changes made using this form are not saved into the input file.

By default, the modified version of the Bartlett-Lewis model is assumed To implement the original model version with constant parameter η , set $\alpha \geq 100$ in the corresponding edit box (top right in the figure). Then the next edit box (second to the right), which normally is labeled v (d) changes to η (d^-1). Insert there the value of parameter η.

To implement the exponential distribution of the cell intensity X set the value of σx equal to that of μx. If these values are different, the gamma distribution is assumed.

To obtain a set of different synthetic series for the same input, use different values of the *Random seed*.

To operate the model in mode 4 or 5 (*Rainfall model test mode* or *Simple rainfall generation mode*, respectively) check off the *Perform adjustment* checkbox.

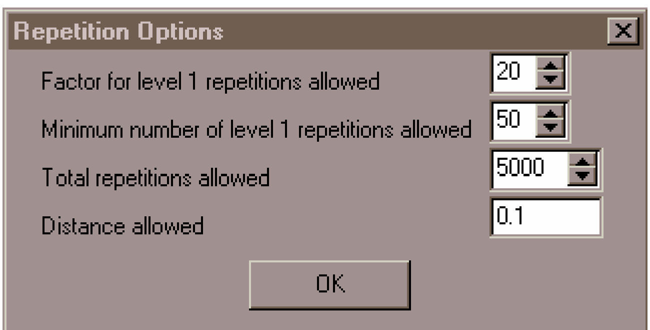


Repetition options form

Activate this form by pressing the appropriate button in the Main form.

Use this form to define the specific values used for the repetition scheme.

The number of level 1 repetitions performed by the program is determined by multiplying the entry of the edit box labeled *Factor for level 1 repetitions allowed* by the number of attempts to establish an appropriate sequence of wet days (number of level 0 repetitions, which is not given by the user but rather is determined by the program). In this manner, the more the number of required level 0 repetitions is, the more the number of level 1 repetitions, and the less the number of level 2 repetitions, will be. This practice results in a faster algorithm The number of level 1 repetitions allowed cannot be set lower than the entry of the edit box labeled *Minimum number of level 1 repetitions allowed*. The total number of level 1 and level 2 repetitions cannot exceed the entry of the edit box labeled *Total repetitions allowed*.



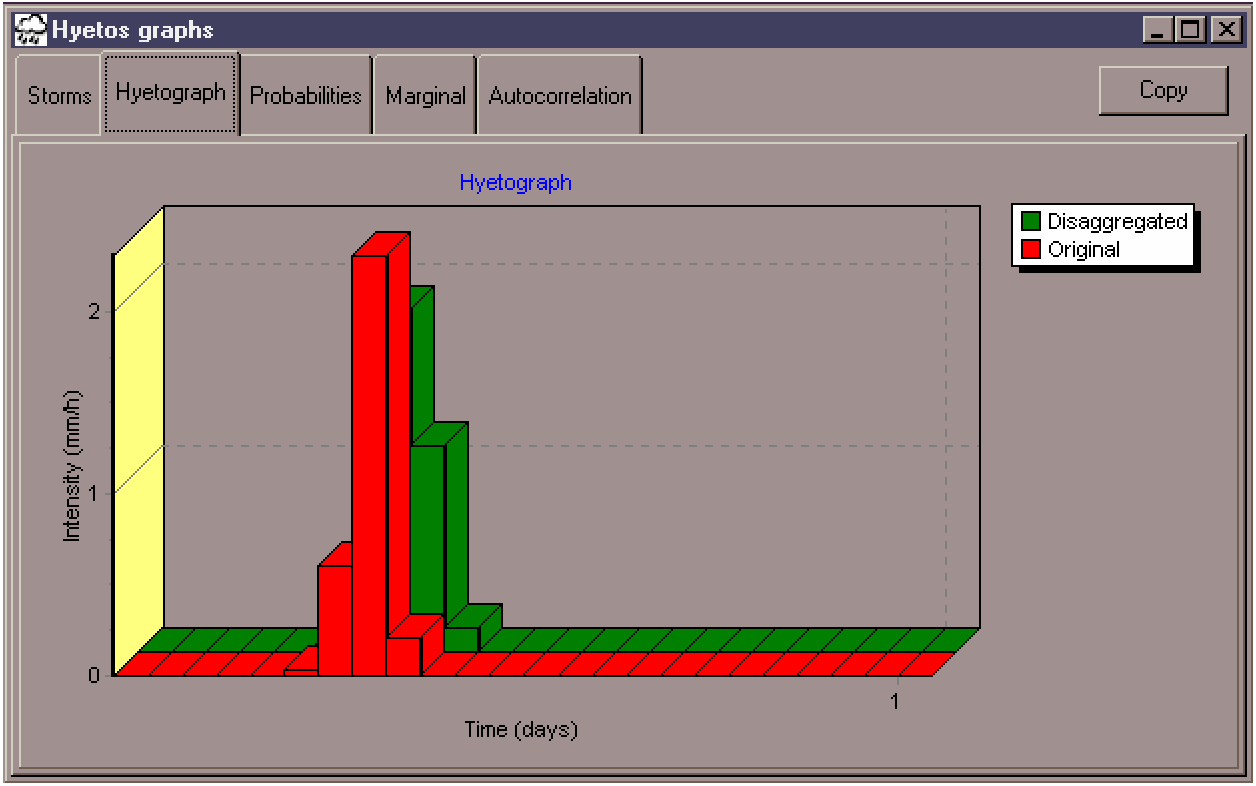
Graphs form

Activate this form by pressing the appropriate button in the Main form.

Use this form to visualise the generated and original (if applicable) rainfall series and their main statistics.

To zoom in any of the graphs, drag on the region of interest downwards. To zoom out, drag on any region within the graph upwards. To move along the graph drag to the desired direction with the right mouse button pressed.

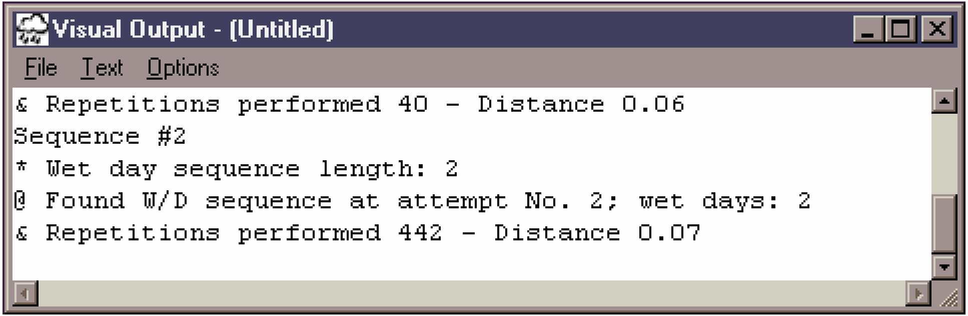
Using the Copy button, a graph is copied into the clipboard and can then be pasted to anywhere else (e.g. word processing programs etc.).



Visual output form

This form appears automatically when some information is printed by the model. The level of details printed on this form is determined by the relevant spin edit box in the Main form. The form can be cleared either by selecting all its content and pressing the *Delete* key on the keyboard, or by pressing the relevant button in the Main form.

The content of the form can be saved in a text file (use the file menu) or copied to the clipboard (press Ctrl-C).



Input and output file formats

The input file contains the daily rainfall depths to be disaggregated (when the model operates in Mode 3 – Operational mode; see Modes of operation) and optionally hourly depths if available (Mode 2 – Full test mode or Mode 4 – Rainfall model test mode). It can also contain the Bartlett-Lewis model parameters. The output file contains the daily and disaggregated hourly rainfall depths. Both input and output files are text files.

Activation of either the input file or the output file is done from the Main form by pressing the relevant buttons.

The input file may or may not contain one or two header lines followed by the data, as shown in the figure below. Dry days (with zero rainfall depths) may or may not be contained in the file.

The output file has similar structure as the input file without headers or additional numeric fields and always contain hourly values. The zero values of dry days are not printed in the output file.