# Modelling rainfall from sub-hourly to daily scale with a heavy tailed meta-Gaussian model

**Marie Boutigny**[1,2], **Pierre Ailliot**[1], **Aurore Chaubet**[2], **Philippe Naveau**[3], **Benoît Saussol**[1]

[1]Laboratoire de Mathématiques de Bretagne Atlantique, UMR 6205, Université de Brest, France
[2]SPL Eau Du Ponant, Brest, France
[3]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CNRS-CEA-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France

**Key Points:**

- A new meta-Gaussian model is developed for precipitation at a wide range of time scales
- Properties of lower and upper tail of rainfall amounts are discussed and used to create the model
- Numerical results show that the model can reproduce the full distribution from dry to heavy rainfall

Corresponding author: Marie Boutigny, `marie.boutigny1@gmail.com`

## Abstract

The time scale of usual hydrological applications can vary from a few minutes to a few days. An accurate description of the precipitation probability distributions at the appropriate time scale is needed to compute meaningful summaries like return levels and periods. However few statistical parametric models can handle the full rainfall distribution at these different temporal scales. In this context, we propose and study a new meta-Gaussian model which leads to a parametric model with four parameters for the full distribution of precipitation. The main advantage of our model is that it can be applied to a wide range of accumulation periods. In particular, it still performs well below the hourly scale. In addition, each parameter is linked to a different part of the distribution: one of them describes the probability of rainfall occurrence, two parameters are related respectively to the shape of lower and upper tails of the distribution, and the last one is a multiplicative scale parameter. The building block of our model is the use of a latent Gaussian process that offers flexibility and simple inference algorithms.

The model is fitted to rain gauge data recorded in Guipavas (France). It is shown that the proposed distribution handles accumulation periods from 6 minutes to several days. The model outperforms other meta-Gaussian models which have been proposed in the literature.


## 1 Introduction

Precipitation intensities and frequencies represent key variables for many environmental studies, not only in hydrology but also agronomy, meteorology and impact studies (see e.g. Bauer et al., 2015; Caseri et al., 2016). There is an abundant literature on the modeling of daily and monthly rainfall intensity distributions. The most popular distribution for positive daily precipitation is probably the gamma distribution (see e.g. Katz, 1999), which generally provides an adequate fit for precipitation as the monthly scale. Other distributions were used and studied, for example, the mixed exponential (see Wilks, 1999), Weibull (Castellvı et al., 2004) and the log-normal (Shoji & Kitaura, 2006). Comparisons were made for specific data sets. Woolhiser and Roldan (1982) put forward the mixed exponential first and the gamma second. Liu et al. (2011) ranked the log normal first, then mixed exponential, gamma and finally exponential. Selker and Haith (1990) compared single-parameter distributions. The performance strongly depends on location of interest as local climate strongly impacts rainfall distribution. In general, the occurrence of precipitation (dry/wet measurement) is modelled separately from the intensity.

Concerning rainfall accumulated on short time periods, most of the statistical literature and probabilistic models are developed for rainfall accumulated over an hour. Still, a particular impetus for describing rainfall distributions and simulating random rainfall draws at any finer time scaled comes from the need to test the sensitivity and robustness of urban hydrological models (Schilling, 1991). At sub-hourly scales, most measurements are frequently null and discrete due rain gauge precision. Still, very few but with high intensities events strongly skew the density to the right. These issues, in particular over-representation of zeros and the heavy tail behavior, cannot be ignored and need to be treated with care. These features create a large probability mass at zero and a heavily tail on the upper part of distribution.

To treat the aforementioned issues with one model, one root of our approach will be the versatile Gaussian random variable. The idea of building from Gaussian variables has been used in the past. For example, taking its square of root was proposed by Panofsky et al. (1958). The Box-Cox transformation (Box & Cox, 1964) has also been widely applied (see e.g. Cecinati et al., 2017; Hussain et al., 2010), like the $g$-and-$h$ transformation introduced by Tukey (1977). Tukey's transforms have the advantage of producing heavy tailed features (see e.g. Goerg, 2015; Xu & Genton, 2017). A common justification behind all these Gaussian based transformation techniques is that the normal distribution represents a solid, simple and flexible building block to handle space-time process (see e.g. Hussain et al., 2010). The main drawback is that the Gaussian layer is always hidden. Still, nowadays, different inference techniques can handle the estimation based on latent Gaussian variables, see the literature on rainfall disaggregation (see e.g. Guillot & Lebel, 1999; Allcroft & Glasbey, 2003; Allard & Bourotte, 2015), downscaling and model correction (see e.g. Maraun et al., 2010; Rebora et al., 2006; Zhao et al., 2017), short term or spatial prediction (see e.g. Sigrist et al., 2012; Benoit et al., 2018)), building stochastic weather generators (see e.g. Bardossy & Plate, 1992; Ailliot et al., 2009; Kleiber et al., 2012), data assimilation Lien et al. (2013) or merging different data sources (see e.g. Cecinati et al., 2017). This body of work leads us to investigate how the strategy of building of Gaussian variables can be adapted to sub-hourly rainfall time scales. In particular, the modeling of dry measurements within this framework can be simplified if one see Bernouili draws as occurrences of a Gaussian variable under a threshold (see e.g. Allard & Bourotte, 2015) and Figure 4.

A remaining difficulty is the modeling of heavy tailed events at sub-hourly scales. Theoretically, the Generalized Pareto (GP) distribution is justified by the field of extreme value theory (see e.g. Papalexiou et al., 2013; Katz, 1999). The main drawback of extreme value theory is that only exceedances above a high threshold are modelled, but not the full distribution. In addition, finding the optimal high threshold that defines an extreme remains a delicate task for the practitioner. Different methods have been proposed to model the full rainfall distribution range and bypass the threshold selection step. For example, Naveau et al. (2016) developed a model for which extreme value theory was both applied to its lower tail and its upper tail, and mi-range rainfall were also taken into account. But, this approach was based on the uniform distribution, not the Gaussian one, and more importantly, it did not handle dry events, but only positive values. In terms of applications, it was only tested on hourly and daily time scales. In addition, the discrete nature of rain gauge measurements ware not fully taken into account.

In this context, the paper focuses on developing a new meta-Gaussian model based on the desired properties for the upper and lower tails of the distribution. In particular the model should be able to produce heavy tails in order to fit precipitation data at small time steps. It should also be flexible enough to fit rainfall accumulated on a wide range of time scales.

The data set that will be used throughout the paper are introduced in Section 2. The proposed model is justified and presented in Section 3, along side with other models that will be used for comparison. Finally results are shown in Section 4, including a comparisons between the different models and a test of the flexibility of the models by varying the time step.

All the models and data that will be used in the paper can be fitted with the R package available on Github at `https://github.com/mbtgy/tcG`.

## 2 Data

In this section the data set used throughout the paper is be presented and the properties of the precipitation distribution raised in the introduction are illustrated. The data used is a 12 years series of precipitation measured at Guipavas, France (geographical coordinates 48.45°N, 4.38°W) provided by Météo France. Guipavas is located in the North-West part of France, close to the city of Brest (cf. Figure 1). Its climate is influenced by oceanic conditions, characterised by a low temperature amplitude and alternation of rainy frontal systems coming from the Atlantic and high pressure systems which bring dryer conditions, with a mean annual precipitation of 1200 mm and a wet day ($\geq 1$ mm) frequency of about 2 days out of 5.

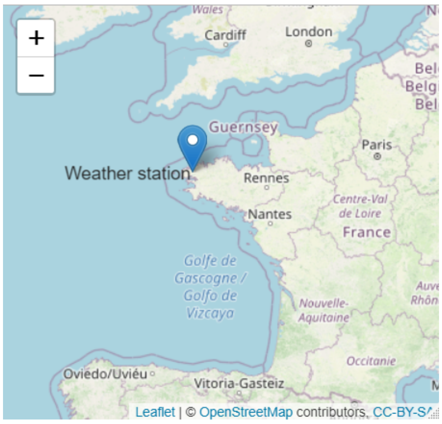



**Figure 1.** Location of the weather station (Guipavas)

Precipitation is available at a 6 minutes time step[1] from 2006 to 2017 and in order to remove the seasonal components, a focus is made on the three months of Autumn, i.e. October, November and December. Figure 2 shows as an example the time series for 2006. Figure 3 shows the histogram of precipitation for the whole series, with the entire distribution on the left, the wet measurements only in the middle and a focus on low and moderate intensities (between 0.2 and 2 mm) on the right.

The measurement device is a tipping bucket gauge with a 0.2 mm precision. Hence the data present a strong discretization, visible in Figure 3 on the right. In Figure 2 many 0.2 mm measurements can be observed in what seem to be dry periods (especially in October). They can be due to the dew that is sometimes sufficient for the gauge to toggle. A drizzle with an intensity lower than 0.2mm/6min can also make 0.2 measurement appear more or less regularly in a "dry" period.

The histogram on the left panel of Figure 3 shows a strong peak in zero, the rest of the histogram being almost invisible. It is expected at a small time step and a fully continuous distribution obviously can not handle the observed frequency of dry measurement (0.94). As for the positive rainfall (Fig. 3, left and middle), the distribution is strongly skewed as most measurements correspond to low intensity rainfall. The positive distribution has a 99.9% quantile of 3.2 mm. The observations above this quantile are highlighted on the left histograms using stars and maximum of the series is 9.4 mm. Some on these heavy rainfall events are also visible in Figure 2.

To sum up the goal is to model precipitation at a sub-hourly scale, hence to have a strongly skewed distribution with a discrete component in zero and the ability to produce heavy tails. The next section discusses the choice of such model.

## 3 Models

### 3.1 Meta-Gaussian Models

A classical approach for modelling rainfall, sometimes called meta-Gaussian model, is to assume that rainfall amounts $Y$ can be linked to a Gaussian variable according to

$$Y = 0 \times \mathbb{1}_{X<0} + \psi(X) \times \mathbb{1}_{X \geq 0}, \quad \text{with } X \sim \mathcal{N}(\mu, 1), \tag{1}$$

[1] Data can be found for free online at `https://donneespubliques.meteofrance.fr/`, but only at 1 hour time step.

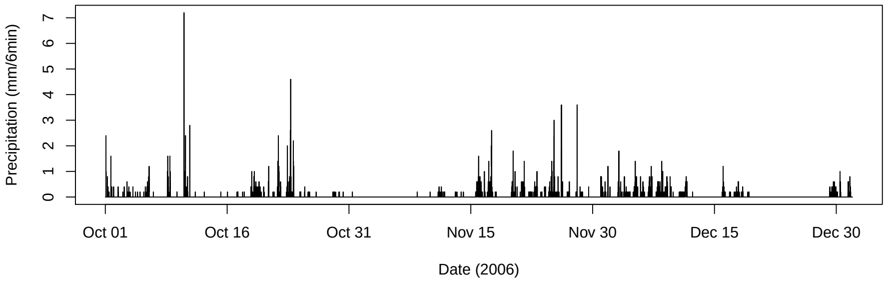



**Figure 2.** Time series of precipitation for Autumn 2006 in Guipavas.

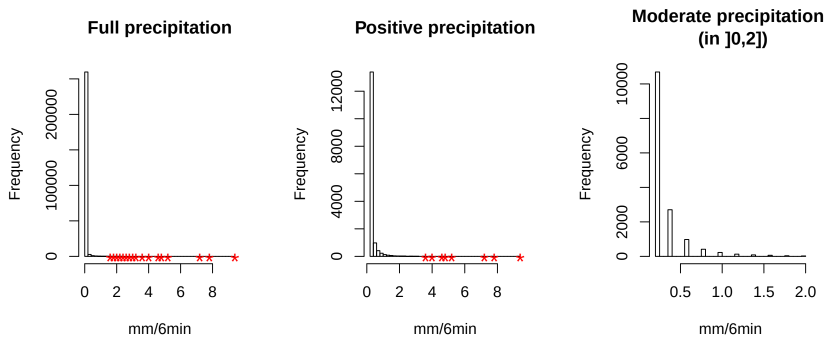



**Figure 3.** Histograms of precipitation at a 6 minutes time step (2006-2017, Guipavas). Whole distribution on the left, positive part in the middle, low intensities (between 0.2 and 2 mm) on the right. The stars represent the observations above the 99.9% quantile of the represented distribution (i.e. 1.4 mm on the left, and 3.2 mm in the middle).

where $\mathbb{1}_A$ is the indicator function equal to 1 if condition $A$ is true, and equal to 0 otherwise. $Y$ denotes the rainfall, $X$ is a Gaussian random variable with mean $\mu$ and variance 1 and $\psi : [0, +\infty[ \to ]0, +\infty[$ is an increasing function which is generally refereed to as the anamorphosis in the literature. The operation of such model is schematised in Figure 4. The censorship in 0 produces dry conditions with a proportion linked to the mean of the Gaussian (step 1 in Figure 4), whereas the transformation $\psi$ acts on the positive part of the distribution which corresponds to wet conditions (step 2 in Figure 4).

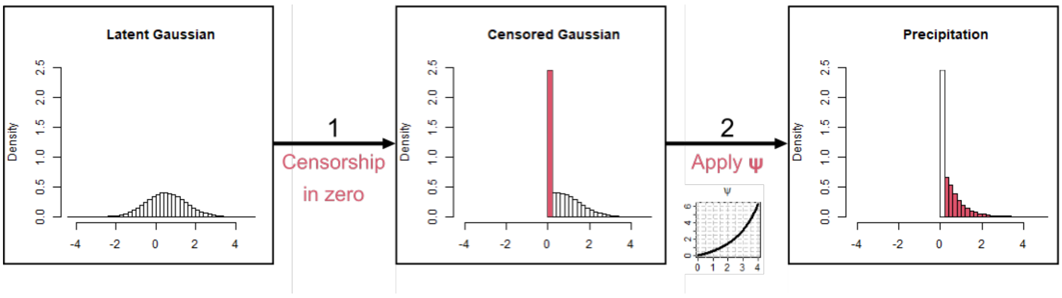



**Figure 4.** Schematic functioning of a meta-Gaussian model. The coloured areas in the histograms represent the part of the distribution modified at each step.

The cumulative distribution function (cdf) of such model can be written as

$$F_Y(y) = \begin{cases} \Phi(\psi^{-1}(y) - \mu) & \text{if } y > 0 \\ \Phi(-\mu) & \text{if } y = 0 \end{cases} \tag{2}$$

where $\Phi$ is the cdf of the standard normal distribution.

Remark that this meta-Gaussian model is general since any positive random variable $Y$ which has a discrete component at the origin like precipitation can be written as (1) using

$$\psi(x) = F_Y^{-}(\Phi(x - \mu)) \tag{3}$$

where $\mu = -\Phi^{-1}(P(Y = 0))$ and $F_Y^{-}$ denotes the quantile function of $Y$ (generalized inverse function of the cdf $F_Y$ of $Y$). Plugging a non-parametric estimate of the quantile function of $Y$ in (3) allows building non-parametric estimates of $\psi$, see e.g. Lien et al. (2013); Cecinati et al. (2017). The dots in Figure 5 show the estimate obtained on the particular data set introduced in Section 2. The shape of $\psi$ near zero is linked to the small precipitations: a horizontal tangent at the origin means that they are more low rainfall than expected low values in the truncated Gaussian. The probability of low rainfall increases if $\psi$ is flatter at the origin. The growth speed is linked to the upper tail: a convex-exponential shape indicates that the tail is heavier than a Gaussian one.

However, parametric approaches are generally favoured in the applications and many models have been proposed in the literature. The most classical is the one of Bardossy and Plate (1992); Sigrist et al. (2012)

$$\psi(x) = \sigma x^{1/\alpha}, \tag{4}$$

but other transformation have been proposed such as Allcroft and Glasbey (2003) which uses a quadratic power function, Rebora et al. (2006) which uses a simple exponential or finally in Allard and Bourotte (2015)

$$\psi(x) = \sigma_2(\exp(\sigma_1 x^{1/\alpha}) - 1) \tag{5}$$

is used. To force the resulting distribution to match a specific distribution the inverse of a cdf can also be used, as it is the case with the Gamma distribution in Kleiber et al. (2012).

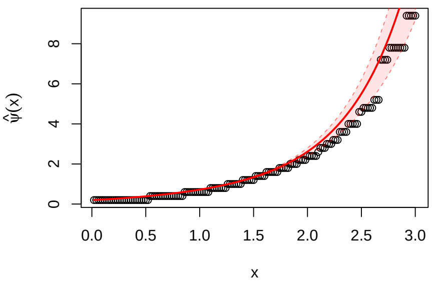



**Figure 5.** Non-parametric estimate of the anamorphosis function based on (3) (dots). The plain curve corresponds to the proposed parametric model fitted to the data (see Section 4.1), and the dotted line is its 95% confidence interval computed on 500 non parametric bootstrap replicates.

Transformation (4) being the most commonly used, it will be a point of comparison and will be referred to as the classical meta-Gaussian model.

### 3.2 Low and Heavy Rainfall Modelling with Meta-Gaussian Models

The choice of an appropriate anamorphosis function for a particular application is typically a trade-off between model complexity, versatility, tractability and interpretability. In this section, it is advocated that the properties of lower and upper tails of the positive part of the rainfall distribution may also provide interesting insights.

Different studies have shown that rainfall at daily or sub-daily scales are generally heavy tailed (see e.g. Papalexiou et al., 2013, and references therein). In this situation, $\psi$ should be chosen such that the transformed Gaussian variable defined by (1) is tail equivalent with a Pareto distribution with positive shape parameter $\xi$. According to Appendix Appendix B, this holds true if and only

$$\lim_{x\to\infty} \frac{x\psi(x)}{\psi'(x)} = \frac{1}{\xi}. \tag{6}$$

The first function that comes to mind which satisfies (6) is $x \mapsto \exp \frac{\xi x^2}{2}$. By rewritting $\psi$ as

$$\psi(x) = \exp \frac{\xi x^2}{2} \exp u(x)$$

- which is always possible - condition (6) becomes

$$\lim_{x\to\infty} \frac{u'(x)}{x} = 0.$$

This condition seems easier to work with as it allows understanding that loosely speaking, the anamorphosis should increase "like" the function $x \mapsto \exp \frac{\xi x^2}{2}$ when $x \to +\infty$ to get heavy tail distributions. In particular, one can verify that most of the anamorphosis that can be found in the literature - including the classical meta-Gaussian model (4) introduced previously - do not satisfy condition (6) and hence are not suitable for modelling rainfall with heavy tail. An exception is the model (5) of Allard and Bourotte (2015) which is tail equivalent with a Pareto distribution if and only if $\alpha = \frac{1}{2}$.

Naveau et al. (2016) advocated, using arguments of the extreme value theory applied to low rainfalls, that the lower part of the distribution of the positive amount should approximately follow a power-law, i.e. satisfy

$$\lim_{y \downarrow 0} \frac{F_Y(y) - F_Y(0)}{y^\alpha} = C$$

for some positive constant $C$ and shape parameter $\alpha > 0$ . Remark that this condition holds true in particular for the Gamma distribution (with shape parameter $\alpha$) which is often used to model daily rainfalls. Using (2) to derive a first order expansion of $F_Y$ close to 0, it can be shown that this holds true if and only if

$$\psi(x) = x^{\frac{1}{\alpha}} K(x) \tag{7}$$

with $K$ such that $\lim_{x \downarrow 0} K(x)$ exists and is strictly positive. This condition is verified by most of the anamorphosis that can be found in the literature, including the classical meta-Gaussian model (4) obviously. Remark however that for model (5) the same parameter $\alpha$ controls the shape of the distribution for low and heavy rainfall, and that is not possible to create an heavy tailed distribution with a power shape parameter different from $\alpha = \frac{1}{2}$ for low rainfalls.

To conclude, for the distribution to have a Pareto upper tail and a lower tail that follows a power law, the anamorphosis $\psi$ should be chosen such that conditions (6) and (7) are satisfied.

### 3.3 Generalized Pareto Meta-Gaussian Model

Based on conditions (6) and (7), this paper advocates the use of the simplest anamorphosis function that satisfies both conditions, i.e.

$$\psi(x) = \sigma x^{\frac{1}{\alpha}} \exp \frac{\xi x^2}{2} \tag{8}$$

with $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^{+*}$, $\alpha \in \mathbb{R}^{+*}$ and $\xi \in \mathbb{R}$. The distribution of the random variable $Y$ defined through (1) with $X \sim \mathcal{N}(\mu, 1)$ and $\psi$ given by (8) will be referred to as the GP meta-Gaussian distribution with parameter $(\mu, \sigma, \alpha, \xi)$.

The expected roles of those four parameters result from the analytical study of a meta-Gaussian model (1) with anamorphosis (8). $\mu$ is directly related to the dry probability through (2), $\sigma$ is a scale parameter, $\alpha$ controls the shape of the lower tail and $\xi$ the shape of the upper tail.

If $\xi > 0$, the distribution is tail equivalent with a Pareto distribution with shape parameter $\xi$. It implies in particular that $E[X^p] = +\infty$ if $p > \frac{1}{\xi}$. The case $\xi = 0$ corresponds to the classical meta-Gaussian model (4). A negative $\xi$ can seem counter intuitive for modelling rainfall as it creates an upper bound to the distribution, but when considering rainfall accumulated on a long period (several days) the fitted model naturally goes for negative $\xi$, as it will be shown in Section 4.3 (Fig. 10). When $\xi < 0$, $\psi$ is strictly monotonic increasing only on the interval $(0, x_{sup})$ with

$$x_{sup} = \sqrt{\left| \frac{-1}{\min(\alpha\xi, 0)} \right|}. \tag{9}$$

For negative $\xi$, the GP meta-Gaussian distribution is thus defined by applying (1) with $\psi$ given by (8) to the Gaussian variable $X \sim \mathcal{N}(\mu, 1)$ truncated at $x_{sup}$. Remind that truncation means that values above $x_{sup}$ are not observed - unlike a censorship, which is used to create the dry component, where the observations above the bound take the value of the bound. The support of the distribution is $[0, y_{sup}]$ with

$$y_{sup} = \sigma \left( \frac{e^{-1}}{\max(-\alpha\xi, 0)} \right)^{\frac{1}{2\alpha}}$$

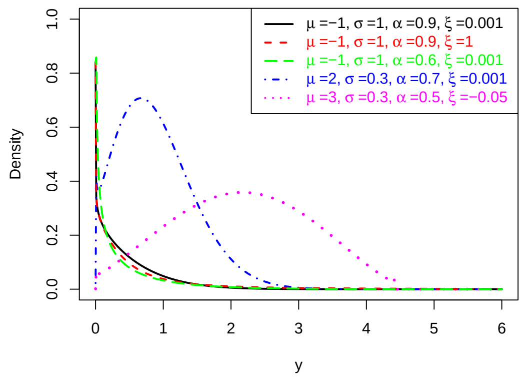



**Figure 6.** Examples of distributions obtained with different parametrizations of the GP meta-Gaussian model.

the upper bound in the precipitation domain. Note that when $\xi \geq 0$ the bounds become $x_{sup} = y_{sup} = +\infty$, so those notations can be used for $\xi \in \mathbb{R}$. When the Gaussian is truncated above $x_{sup}$, the cdf (2) must be corrected by the probability of truncation (see Appendix A).

Another advantage of the GP meta-Gaussian transformation is the possibility to derive an analytical expression for the inverse of $\psi$

$$\psi^{-1}(y) = \sqrt{\frac{1}{\alpha\xi} W\left(\alpha\xi\left(\frac{y}{\sigma}\right)^{2\alpha}\right)} \tag{10}$$

where $W$ denotes the Lambert $W$ function (see Goerg, 2015) defined as the inverse of the function $x \mapsto x \log x$, which is available in usual statistical software. Using (2) an analytical expression can be derived for the cdf and the probability density function (pdf) of the GP meta-Gaussian model. This allows in particular to compute easily the likelihood function and fit the model to data (see Section 4.1). Analytical expressions for the finite moments can also be derived, which is not the case for many meta-Gaussian models that can be found in the literature. To our knowledge the classical transform (4) is the only other meta-Gaussian model with analytical moments.

Expressions for the GP meta-Gaussian model pdf, cdf, quantile function and moments can be found in Appendix A.

### 3.4 Extended Generalized Pareto Model

As mentioned in the introduction, the extended GP model of Naveau et al. (2016) will be used as a benchmark in this paper. The specificity of the extended GP is that extreme value theory is applied to the upper tail but also the lower tail, allowing to avoid threshold selection. Note that here the lower tail does not include the dry component of precipitation, and that the model is developed only for strictly positive rainfall amounts $Y_+$. The general model proposed in Naveau et al. (2016) is defined as

$$Y_+ = \sigma H_\xi^{-1}(U^{1/\alpha}) \tag{11}$$

where $U$ follows a standard uniform distribution and $H_\xi$ is the cdf of a GP distribution with shape parameter $\xi$. We use the same notation for the parameters than in the GP meta-Gaussian model since they have the same interpretation with $\sigma$ a scale parameter, $\alpha$ controlling the power shape of the lower tail, and $\xi$ upper tail.

More sophisticated transformation of the uniform distribution are proposed in Naveau et al. (2016) but they do not provide a better fit for the particular data set considered in the next section.

The cdf of $Y_+$ can be related to the cdf of $U$ in the same way that the cdf of the GP meta-Gaussian model can be related to the Gaussian cdf, hence one can write

$$F_{Y_+}(y) = \{H_\xi(y/\sigma)\}^\alpha. \tag{12}$$

# 4 Detailed Example of Inference

The previous section introduced three models: first a classical meta-Gaussian model, i.e. (1) with anamorphosis (4), then the proposed GP meta-Gaussian model, i.e. (1) with anamorphosis (8) and finally the extended GP model, i.e. (11) for strictly positive rainfall.

Those three models will be tested and compared in this section using the data presented in Section 2.

## 4.1 Inference Method: Dealing with Discretization

With the meta-Gaussian model (1) the dry measurements are supposed to be created by the censorship in 0 and hence controlled by $\mu$ - the mean of the latent variable. Remark that the anamorphosis (8) can produce values that are lower than 0.2, the minimal value that can be measured - that will be noted $y_m$ - and when discretizing those values will become zeros. In other words zeros are produced by the censorship (controlled by $\mu$) and also by the anamorphosis (controlled by $\{\sigma, \alpha, \xi\}$). Therefore in order to have separable parameters $y_m$ needs to be introduced in $\psi$ as follows.

$$\psi(x) = y_m + \sigma x^{\frac{1}{\alpha}} \exp \frac{\xi x^2}{2} \tag{13}$$

Note the $\psi^{-1}$ and $y_{sup}$ are consequently modified.

A similar reasoning can be applied to the extended GP model: in (11) when $U \to 0$ we have $Y_+ \to 0$, while the minimal value that can be observed is 0.2. Hence $y_m$ is also introduced in $Y_+ = y_m + \sigma H_\xi^{-1}(U^{1/\alpha})$.

The introduction of $y_m$ in the models greatly improves the results obtained when fitting the three models to the particular data set considered in this study: a completely different solution is chosen for the parameters and the fit is better for the whole distribution.

For the meta-Gaussian models, the likelihood is usually computed directly from the continuous density (see Appendix A), however it has been noticed that taking into account discretization significantly improves the results. It is valid for the data considered in this study, i.e. rain gauge measurements with a bucket that tips every 0.2 mm (see Figure 11 and discussion in Conclusion).

The discrete likelihood is based on the functioning of a tipping bucket rain gauge, which means that $P(G = g) = P(g \leq Y < g + \mathit{step})$, $G$ being the discrete measurement, $Y$ being the continuous rainfall and *step* being the precision of $G$, i.e. 0.2 mm in our case. Therefore the discrete log likelihood for both meta-Gaussian models is based on the cdf (2):

$$\log \mathcal{L}_\theta^{MG} = n_0 \log(\Phi(-\mu)) + \sum_{i:\, y_i > 0} \log \left\{ \Phi(\psi^{-1}(y_i + \mathit{step}) - \mu) - \Phi(\psi^{-1}(y_i) - \mu) \right\},$$

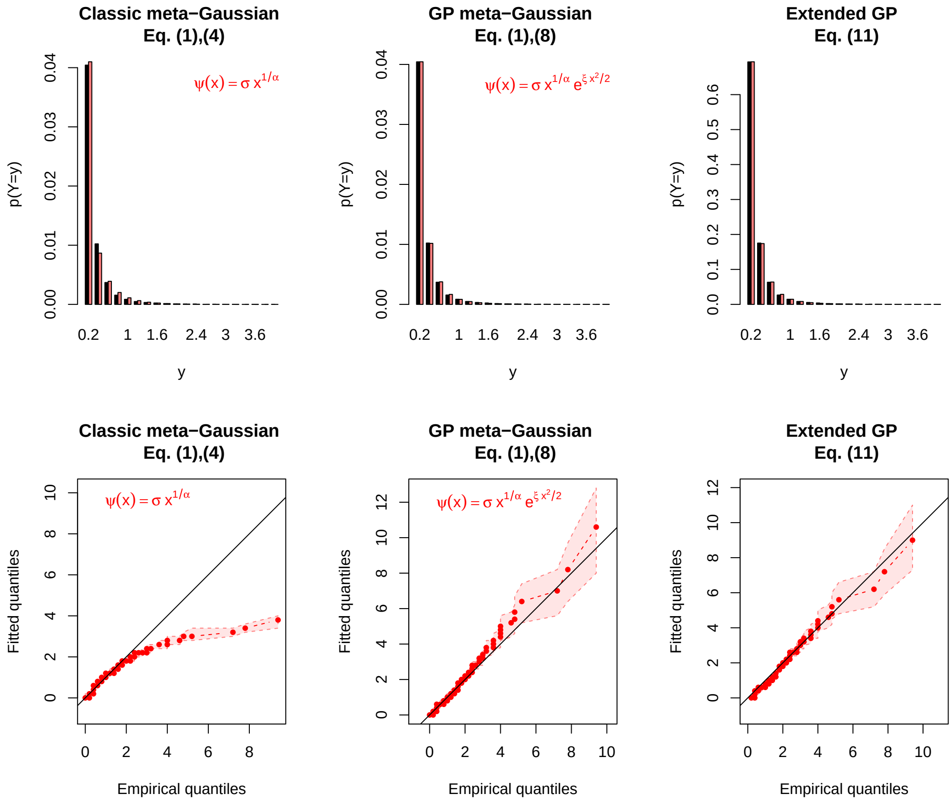



**Figure 7.** Fitting the three models to the whole series. 1st row: density of low intensities ($\leq 4$ mm), empirical in black, model in red. 2nd row: quantile-quantile plot of the full distribution, empirical versus fitted quantile. The light area gives the 95% intervals computed with 500 non parametric bootstrap replicates.

where $n_0$ is the number of dry measurements. For the GP meta-Gaussian model, $\psi$ is given in (13) and $\theta = \{\mu, \sigma, \alpha, \xi\}$ and the the classical meta-Gaussian corresponds to the particular case where $\xi = 0$.

The discrete log likelihood of the extended GP model is based on its cdf (12) in the same way, with only three parameters $\theta = \{\sigma, \alpha, \xi\}$ as there is no parameter for the dry measurement.

The moments could also be used to infer the parameters of the GP meta-Gaussian (see Appendix A). Remark however that the usual method of moment is tricky to implement when working with heavy tail distributions since some moments are infinite. Furthermore it may also be sensitive to the discretization of the data.

### 4.2 Results

Figure 7 shows the results obtained with of the three models adjusted to 6 minutes rain gauge data with discrete likelihood. First of all a focus is made on the proposed GP meta-Gaussian model, hence the second column. The global fit of the model, observable in the quantile-quantile (QQ) plot on the bottom, is very satisfying. The density (top) shows a quasi perfect fit for low intensities. For medium intensities (4-6 mm) a slight deviation can be observed on the QQ plot, the model producing more of these values that

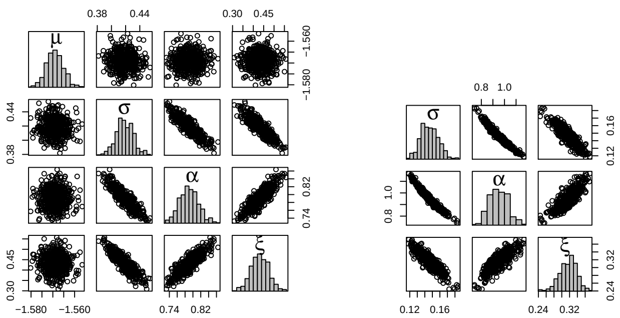


**Figure 8.** Correlation between parameters, computed with 500 non parametric bootstrap replicates. GP meta-Gaussian on the left, extended GP on the right. Histograms of the parameters are shown in the diagonal.

what is present in the data. Remark that the GP meta-Gaussian model was tested on many data sets and it was observed that it is not a recurrent issue. Finally the tail of the distribution is very well reproduced by the model, and is quite heavy as the tail parameter is 0.45. It means that moments of order greater than 2.2 are infinite. Hence even when simply computing the variance of some precipitation series at a fine scale (minutes), one must remember that the variance of the estimator is likely to be infinite.

Those results can also be found in Figure 5, where the plain curve represents the estimated transformation function $\psi$, and the dots are the empirical one. $\psi$ is almost perfectly estimated for low intensities and the good reproduction of extremes is satisfying.

Figure 8 (left) shows the correlation between the parameters of the GP meta-Gaussian model. $\mu$, that controls the dry component of the distribution, is completely uncorrelated with the parameters of the transformation, which is satisfying. However there is a quite strong correlation between the other parameters. It will be an important point to keep that in mind when trying to interpret the parameters.

As expected the classical meta-Gaussian model is unable to reproduce the tail of the distribution (Fig. 7, first column). The lower tail is also affected, and even though it is not too bad the other models do better even for low intensities. The parameters are uncorrelated for this model, but the poor fit explains why the classical meta-Gaussian model will not be further discussed in the following section.

Finally the extended GP and the GP meta-Gaussian models give very similar results in terms of goodness of fit (Fig. 7, second and third column). The parameters of the extended GP model seem to be slightly more correlated than the ones obtained for the proposed model (Fig. 8, right). It was also noted that the parameters were quite close in terms of values, which will be investigated in the next section.

Note that meta-Gaussian model with transform (5) was also tested: it gave a satisfying fit, as it only unhooked in the very end of the tail, but the quality of fit was not the main problem with this model. In Section 3.1 it was already mentioned that the low and strong intensities are controlled by the same parameter, and the correlation between parameters showed a quasi deterministic link between the three parameters of the transform.

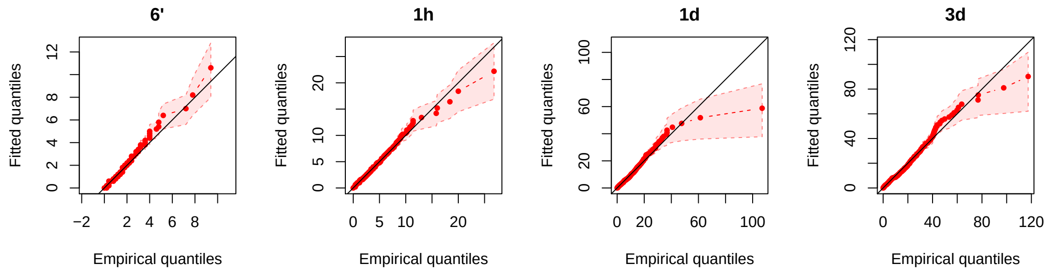



**Figure 9.** Fitting the GP meta-Gaussian distribution at various time lags (6 minutes, 1 hour, 1 day and 3 days). The light area gives the 95% intervals computed with 500 non parametric bootstrap replicates.

### 4.3 Time Lag

Accumulating precipitation data over various periods of time allows exploring the interpretability of the parameters and checking the flexibility of the model. As it was said through the paper, the GP meta-Gaussian distribution aims at modelling precipitation at a wide range of time scales, from sub-hourly to daily rainfall. The extended GP model will also be used in this section, and as it was already mentioned the classical meta-Gaussian model will not be further discussed.

The GP meta-Gaussian model is fitted to data ranging from 6 minutes up to daily scale gave very satisfying results (see the QQ plots in Figure 9), which demonstrates that the model is flexible enough to reproduce the distribution of rainfall at a wide range of time lags. Remark that the extended GP model is not shown in Figure 9 but the QQ plots obtained were so similar that they were almost indistinguishable.

Figure 10 shows the evolution of the model parameters with time aggregation (note that the time axis is non linear). The first thing to notice is the fact that the evolution of the parameters is smooth and monotone. Note that with other meta-Gaussian models such smoothness and monotony was not observed.

$\mu$ is increasing with aggregation, which is expected as there are less and less dry measurements. $\sigma$ is the global scale parameter, hence it is expected to increase as well. The tail parameter $\xi$ is decreasing which is coherent with the intuition that averaging random variable will tend to "gaussianize" them and produce distributions with lighter tails. Plus practical knowledge of precipitation in the considered region tells us that with a larger time steps there are less extreme values.

As $\alpha$ decreases the distribution becomes more "peaky" at the origin. The reason why $\alpha$ decreases is not straightforward, but the rainfall accumulated over a given time period is the sum of a random number (because of the dry measurements) of correlated (because of the temporal dependence) random variables and hence it may have a complicated behaviour. Furthermore $\alpha$ controls not only the lower tail but has also a strong impact on the medium intensities. Another lead is that it might be due to the strong correlation between the parameters (Fig. 8). This reminds us that caution is needed and one should not over interpret the parameters.

As the GP meta-Gaussian model and the extended GP one are closely related, it is also interesting to check if the parameters have the same role in both models, and hence the same evolution with time aggregation. The parameters estimation of the models are represented in Figure 10, showing that from 6 minutes to 1 day the parameters of both models exhibit similar variations. After 1 day they diverge to different solutions, and it

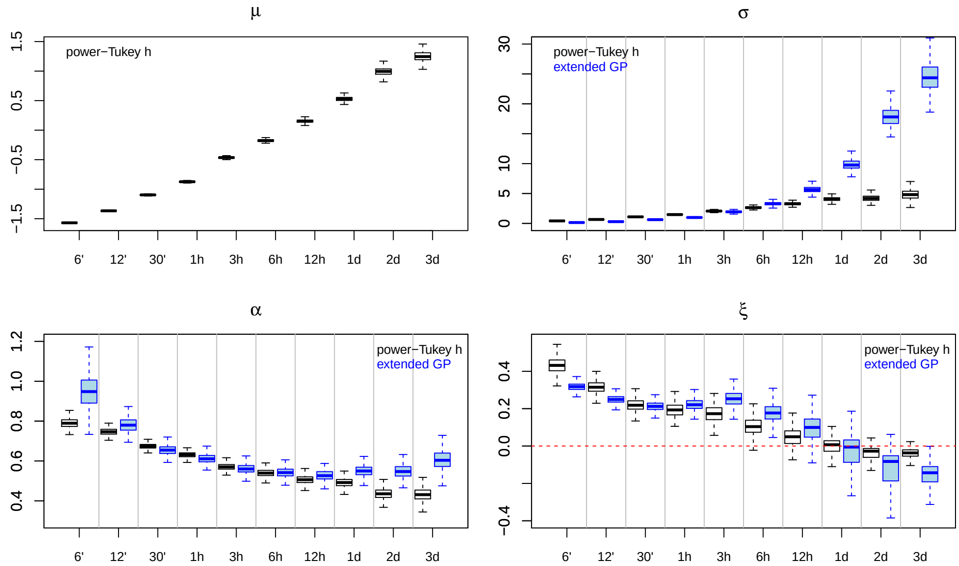



**Figure 10.** Parameter estimation at various time lags, boxplots computed on 500 non parametric bootstrap replicates. The empty boxes are parameters obtained with the GP meta-Gaussian distribution and the fulled ones with the extended GP.

could be explained by the fact that after one day a non-zero mode starts to emerge on the histogram of the data. However both models still have the same quality of fit.

## 5 Conclusion

The goal of the GP meta-Gaussian distribution is to extend the class of meta-Gaussian models to small time steps - several minutes. Properties of the lower and upper tails motivate the choice of the transformation, using extreme value theory to derive two conditions for the anamorphosis. The proposed GP meta-Gaussian model is tractable and analytical expressions exist for the pdf, cdf, quantile function and for the moments.

Results are very satisfying for a wide range of time steps - from 6 minutes to several days, demonstrating the flexibility of the model. Even though the data presented in this article did not allow to further aggregate (only 12 years of data), the model was also tested up to monthly scale using another data set of daily data and gave similarly satisfying results.

Comparison with a classical meta-Gaussian model shows what the proposed transform brings to this class of models: a better fit at small time scale due to its capacity to produce heavy tails. The GP meta-Gaussian model is quite similar to the extended GP model Naveau et al. (2016) in terms of construction but also in terms of performance. The advantage of the meta-Gaussian model is its link with the latent Gaussian that allows using methods developed for Gaussian data (multivariate, spatiotemporal models, Kalman-like algorithm, etc.).

The evolution of parameters with aggregation is very interesting for several reasons. First it demonstrates the similarities between the GP meta-Gaussian and the extended GP model. Second the smooth and monotonic evolution makes it possible to interpret the roles of the parameters, even though $\alpha$ is a bit harder to interpret than the other parameters. Finally when seeing such smoothness in the variation one can won-

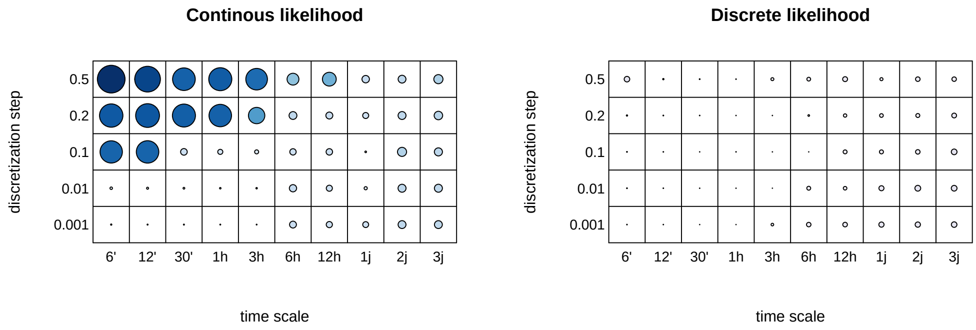



**Figure 11.** Quantile 0.95 of parameters error between the truth and both likelihood estimates (the size and colour of the dots are on a log scale).

der if a unique model with parameters varying with time aggregation could be developed. Having such model would be very interesting as it would mean that hourly or daily rainfall could say something about the parameters values at a few minutes time scale.

A final point that was not central in the paper but that is very important for practical applications is the role of discretization. To demonstrate the performance of the discrete likelihood against the continuous one (see Section 4.1), Figure 11 shows the results of the optimisation on simulated data with various discretization precisions and various "time scales", i.e. various parameter sets that were the ones found in Section 4.3. What is shown is the 0.95 quantile of the absolute error between the true parameters and the ones estimated by the two likelihoods. Both likelihoods could be used for data with 0.01 precision but for higher discretization steps the discrete likelihood performs way better than the continuous one, even with hourly data. Note that most of the parameter error is due to $\alpha$ and $\xi$. For the models that were tested in this paper, the discrete likelihood was a efficient and easy way to deal with the discretization issues.

# Appendix A Some Theoretical Properties of the GP Meta-Gaussian Distribution

The density, cdf and quantile function of a meta-Gaussian model as defined in (1) are

$$f_Y(y) = c \times \begin{cases} \phi_\mu\left(\psi^{-1}(y)\right)/\psi'\left(\psi^{-1}(y)\right) & \text{if } y > 0 \\ \Phi_\mu(0) & \text{if } y = 0 \end{cases},$$

$$F_Y(y) = c \times \begin{cases} \Phi_\mu(\psi^{-1}(y)) & \text{if } y > 0 \\ \Phi_\mu(0) & \text{if } y = 0 \end{cases},$$

$$F_Y^{-}(u) = \begin{cases} \psi(\Phi_\mu^{-1}(u/c)) & \text{if } u > \Phi_\mu(0) \\ 0 & \text{if } u = \Phi_\mu(0) \end{cases},$$

with $\phi_\mu$ and $\Phi_\mu$ denoting respectively the pdf and cdf of a normal distribution with mean $\mu$. $c$ is the normalisation constant that deals with the probability of truncation when $\xi < 0$ with the GP meta-Gaussian transform. Hence $c = 1$ for the classical transform (4), and for the GP meta-Gaussian transform (8) $c = 1/\Phi_\mu(x_{sup})$, with $x_{sup}$ the upper bound in the Gaussian domain as defined in (9).

An explicit writing of the moments was found for the GP meta-Gaussian distribution when $\xi \geq 0$. Let us write $Y_+$ the wet measurements.

$$\begin{aligned} E(Y_+^p) &= \frac{1}{\sqrt{2\pi}(1-\Phi(-\mu))} \int_0^{+\infty} \psi(x)^p \exp\left\{-\frac{1}{2}(x-\mu)^2\right\} dx \\ &= \frac{\sigma^p}{\sqrt{2\pi}(1-\Phi(-\mu))} \exp\left(-\frac{\mu^2}{2}\right) \int_0^{+\infty} x^{p/\alpha} \exp\left\{-\frac{1-\xi p}{2}x^2 + \mu x\right\} dx \end{aligned}$$

By identification in Gradshteyn and Ryzhik (2007) (eq. 3.462.1, page 365), with $\gamma = -\mu$, $\nu - 1 = p/\alpha$ and $\beta = (1-\xi p)/2$,

$$E(Y_+^p) = \frac{\sigma^p (1-\xi p)^{-\frac{1}{2}(\frac{p}{\alpha}+1)}}{\sqrt{2\pi}(1-\Phi(-\mu))} \exp\left\{\frac{\mu^2}{2}\left(\frac{1}{2(1-\xi p)} - 1\right)\right\} \Gamma\left(\frac{p}{\alpha}+1\right) D_{-(\frac{p}{\alpha}+1)}\left(-\frac{\mu}{\sqrt{1-\xi p}}\right)$$

$\Gamma$ is the Gamma function and $D_\nu$ can be expressed with Kummer's confluent hypergeometric function of first kind (see Gradshteyn & Ryzhik, 2007, eq. 9.240, page 1028).

The expression of $E(Y_+^p)$ is valid if $-\alpha < p < 1/\xi$. Note that for $\xi > 0$ first estimation of $\xi$ is needed to determine to moments that can be computed, which is due to the fact that some moments are infinite.

## Appendix B Pareto Tail for Meta-Gaussian Models

**Proposition 1** *Let $Z$ be any positive absolutely continuous random variable with pdf $f_Z$ and with a Pareto survival function $\overline{F}_Z$. Let $X$ be any standardized normal distributed random variable, and let us define the positive random variable*

$$Y \overset{d}{=} \psi(X),$$

*where $\overset{d}{=}$ means equality in distribution and $\psi(.)$ represents a continuous and increasing function from the real line to $[0, \infty)$. The two random variables $Z$ and $Y$ are tail-equivalent if and only if*

$$\lim_{x\to\infty} \frac{x\psi(x)}{\psi'(x)} = \frac{1}{\xi}, \tag{B1}$$

*where $\xi$ corresponds the common positive GP shape parameter of $Z$.*

**Proof of Proposition 1:** Let $\phi$ and $\overline{\Phi}$ denote respectively the pdf and survival function of a standard normal distribution $X$.

Recall that $Z$ and $Y$ are tail-equivalent, if and only

$$\lim_{y\to\infty} \frac{\overline{F}_Z(y)}{\mathbb{P}[Y > y]} = c \in (0, \infty),$$

This condition is satisfied if they have the same tail index. Assuming a Pareto tail with positive shape parameter $\xi$ for $Z$ implies that $Z$ is regularly varying with index $1/\xi$. Proposition A.3.8(b) from Embrechts et al. (2013) recalled that this regular variation type is equivalent to

$$\lim_{z\to\infty} \frac{z \times f_Z(z)}{\overline{F}_Z(z)} = \frac{1}{\xi}.$$

Hence, to show that $Y$ and $Z$ are tail equivalent, one needs to determine under which condition it can be written that

$$\lim_{z\to\infty} \frac{z \times f_Y(z)}{\overline{F}_Y(z)} = \frac{1}{\xi}.$$

where $f_Y$ and $\overline{F}_Y$ denote the pdf and survival function of $Y$, respectively.

By construction, the survival function of $Y$ equals to

$$\overline{F}_Y(z) = \mathbb{P}[X > \psi^{-1}(z)] = \overline{\Phi}\left[\psi^{-1}(z)\right],$$

The density of $Y$ is

$$f_Y(z) = \left(\psi^{-1}(z)\right)' \phi\left[\psi^{-1}(z)\right].$$

Then one can write

$$\frac{z \times f_Y(z)}{\overline{F}_Y(z)} = \left(z \times \psi^{-1}(z) \times \left(\psi^{-1}(z)\right)'\right) \times \left(\frac{\phi\left[\psi^{-1}(z)\right]}{\psi^{-1}(z)\overline{\Phi}\left[\psi^{-1}(z)\right]}\right).$$

Mill's ratio (see Embrechts et al., 2013) tells us that the ratio in the last bracket goes to one as $\psi^{-1}(z)$ goes to $\infty$ (i.e. as $z$ grows). Hence, the condition

$$\lim_{z\to\infty}\left(z \times \psi^{-1}(z) \times \left(\psi^{-1}(z)\right)'\right) = \frac{1}{\xi}, \tag{B2}$$

is equivalent to

$$\lim_{z\to\infty}\frac{z \times f_Y(z)}{\overline{F}_Y(z)} = \frac{1}{\xi}.$$

This is equivalent to have tail equivalence between $Z$ and $Y$.

Changing variables with $z = \psi(x)$, $x = \psi^{-1}(z)$ and $\left(\psi^{-1}(z)\right)' = dx/dz$, condition (B2) is equivalent to condition (B1).

This is the necessary and sufficient condition on $\psi(.)$ to build a Pareto random variable of tail index $\xi$ from a standardized normal random variable $X$.

## Acknowledgments

The authors acknowledge Météo France for the Guipavas precipitation time series that is available at `https://donneespubliques.meteofrance.fr/` at a 1 hour step for free, and at a 6 minutes step with charges.

This work was supported by Eau du Ponant SPL, and took place in the context of the MEDISA (Méthodologie de Dimensionnement des Systèmes d'Assainissements) project.

Part of P. Naveau's work was supported by the European DAMOCLES-COST-ACTION on compound events, and he also benefited from French national programs, in particular FRAISE-LEFE/INSU, MELODY-ANR, ANR-11-IDEX-0004 -17-EURE-0006 and ANR T-REX AAP CE40.

## References

Ailliot, P., Thompson, C., & Thomson, P. (2009). Space–time modelling of precipitation by using a hidden markov model and censored gaussian distributions. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *58*(3), 405–426.

Allard, D., & Bourotte, M. (2015). Disaggregating daily precipitations into hourly values with a transformed censored latent gaussian process. *Stochastic environmental research and risk assessment*, *29*(2), 453–462.

Allcroft, D. J., & Glasbey, C. A. (2003). A latent gaussian markov random-field model for spatiotemporal rainfall disaggregation. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *52*(4), 487–498.

Bardossy, A., & Plate, E. J. (1992). Space-time model for daily rainfall using atmospheric circulation patterns. *Water resources research*, *28*(5), 1247–1259.

Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*, *525*(7567), 47–55.

Benoit, L., Allard, D., & Mariethoz, G. (2018). Stochastic rainfall modeling at sub-kilometer scale. *Water Resources Research*, *54*(6), 4108–4130.

Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, *26*(2), 211–243.

Caseri, A., Javelle, P., Ramos, M.-H., & Leblois, E. (2016). Generating precipitation ensembles for flood alert and risk management. *Journal of Flood Risk Management*, *9*(4), 402–415.

Castellvı, F., Mormeneo, I., & Perez, P. (2004). Generation of daily amounts of precipitation from standard climatic data: a case study for argentina. *Journal of hydrology*, *289*(1-4), 286–302.

Cecinati, F., Wani, O., & Rico-Ramirez, M. A. (2017). Comparing approaches to deal with non-gaussianity of rainfall data in kriging-based radar-gauge rainfall merging. *Water Resources Research*, *53*(11), 8999–9018.

Embrechts, P., Klüppelberg, C., & Mikosch, T. (2013). *Modelling extremal events: for insurance and finance* (Vol. 33). Springer Science & Business Media.

Goerg, G. M. (2015). The lambert way to gaussianize heavy-tailed data with the inverse of tukey'sh transformation as a special case. *The Scientific World Journal*, *2015*.

Gradshteyn, I. S., & Ryzhik, I. M. (2007). *Table of integrals, series, and products* (7th ed.; A. Jeffrey & D. Zwillinger, Eds.). Academic press.

Guillot, G., & Lebel, T. (1999). Disaggregation of sahelian mesoscale convective system rain fields: Further developments and validation. *Journal of Geophysical Research: Atmospheres*, *104*(D24), 31533–31551.

Hussain, I., Spöck, G., Pilz, J., & Yu, H.-L. (2010). Spatio-temporal interpolation of precipitation during monsoon periods in pakistan. *Advances in water resources*, *33*(8), 880–886.

Katz, R. W. (1999). Extreme value theory for precipitation: sensitivity analysis for climate change. *Advances in water resources*, *23*(2), 133–139.

Kleiber, W., Katz, R. W., & Rajagopalan, B. (2012). Daily spatiotemporal precipitation simulation using latent and transformed gaussian processes. *Water Resources Research*, *48*(1).

Lien, G.-Y., Kalnay, E., & Miyoshi, T. (2013). Effective assimilation of global precipitation: Simulation experiments. *Tellus A: Dynamic Meteorology and Oceanography*, *65*(1), 19915.

Liu, Y., Zhang, W., Shao, Y., & Zhang, K. (2011). A comparison of four precipitation distribution models used in daily stochastic models. *Advances in Atmospheric Sciences*, *28*(4), 809–820.

Maraun, D., Wetterhall, F., Ireson, A., Chandler, R., Kendon, E., Widmann, M., . . . others (2010). Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*, *48*(3).

Naveau, P., Huser, R., Ribereau, P., & Hannart, A. (2016). Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection. *Water Resources Research*, *52*(4), 2753–2769.

Panofsky, H. A., Brier, G. W., & Best, W. H. (1958). *Some application of statistics to meteorology.* Earth and Mineral Sciences Continuing Education, College of Earth and . . . .

Papalexiou, S., Koutsoyiannis, D., & Makropoulos, C. (2013). How extreme is extreme? an assessment of daily rainfall distribution tails. *Hydrology & Earth System Sciences*, *17*(1).

Rebora, N., Ferraris, L., von Hardenberg, J., & Provenzale, A. (2006). Rainfarm: Rainfall downscaling by a filtered autoregressive model. *Journal of Hydrometeorology*, *7*(4), 724–738.

Schilling, W. (1991). Rainfall data for urban hydrology: what do we need? *Atmospheric Research*, *27*(1-3), 5–21.

Selker, J. S., & Haith, D. A. (1990). Development and testing of single-parameter precipitation distributions. *Water resources research*, *26*(11), 2733–2740.

Shoji, T., & Kitaura, H. (2006). Statistical and geostatistical analysis of rainfall in central japan. *Computers & Geosciences*, *32*(8), 1007–1024.

Sigrist, F., Künsch, H. R., Stahel, W. A., et al. (2012). A dynamic nonstationary spatio-temporal model for short term prediction of precipitation. *The Annals of Applied Statistics*, *6*(4), 1452–1477.

Tukey, J. W. (1977). Modern techniques in data analysis. In *Proceedings of the nsf-sponsored regional research conference* (Vol. 7).

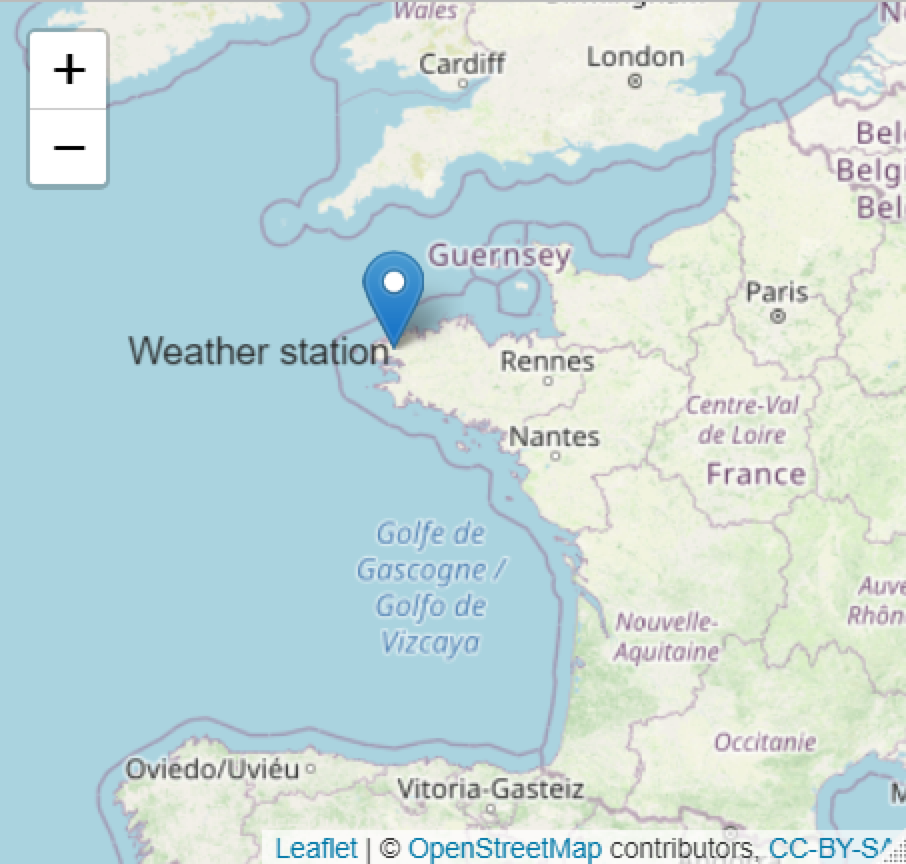
Wilks, D. S. (1999). Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agricultural and forest meteorology*, *93*(3), 153–169.

Woolhiser, D. A., & Roldan, J. (1982). Stochastic daily precipitation models: 2. a comparison of distributions of amounts. *Water resources research*, *18*(5), 1461–1468.

Xu, G., & Genton, M. G. (2017). Tukey g-and-h random fields. *Journal of the American Statistical Association*, *112*(519), 1236–1249.

Zhao, T., Bennett, J. C., Wang, Q., Schepen, A., Wood, A. W., Robertson, D. E., & Ramos, M.-H. (2017). How suitable is quantile mapping for postprocessing gcm precipitation forecasts? *Journal of Climate*, *30*(9), 3185–3196.

**Figure1.**



Wales
Cardiff
London
Guernsey
Paris
Weather station
Rennes
Nantes
Centre-Val
de Loire
France
Golfe de
Gascogne /
Golfo de
Vizcaya
Nouvelle-
Aquitaine
Occitanie
Oviedo/Uviéu
Vitoria-Gasteiz
Leaflet | © OpenStreetMap contributors

**Figure2.**

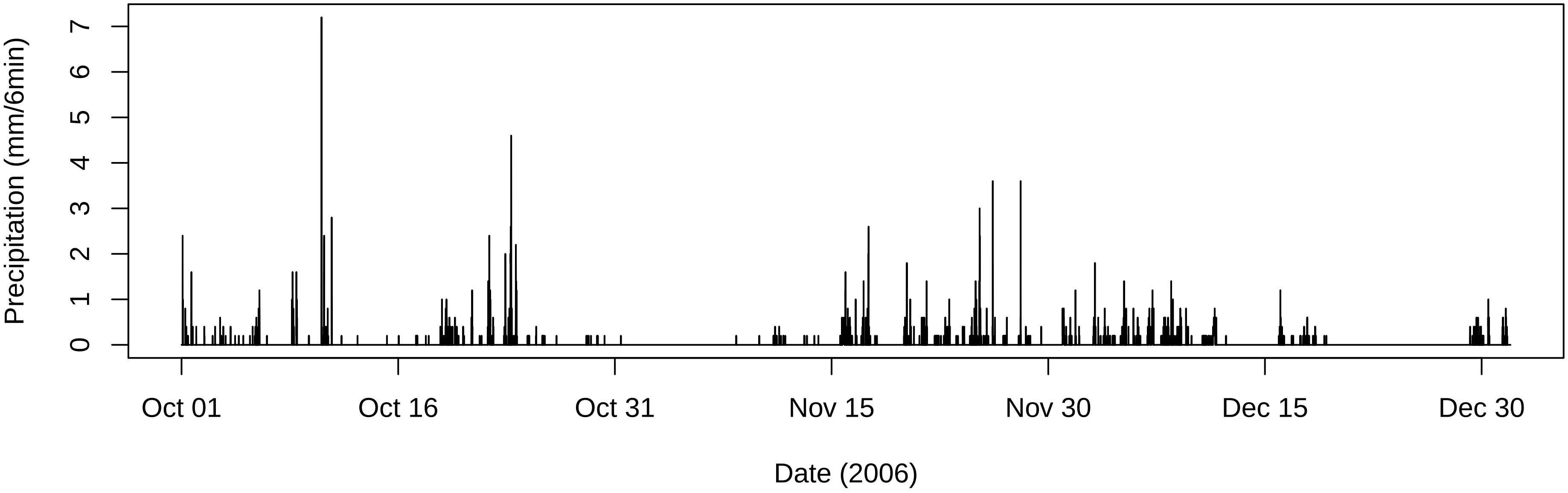
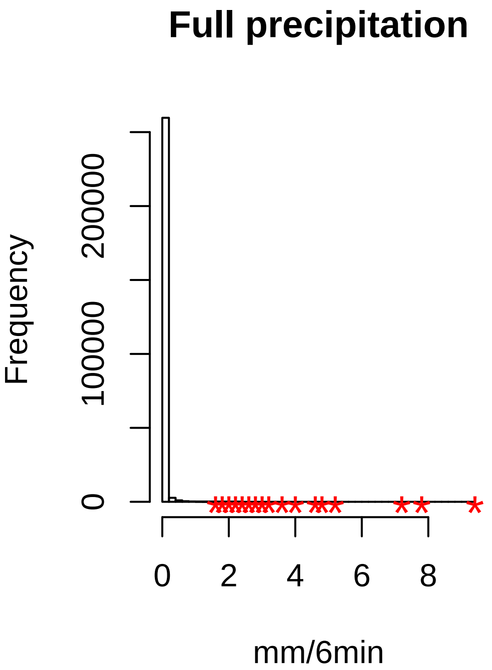
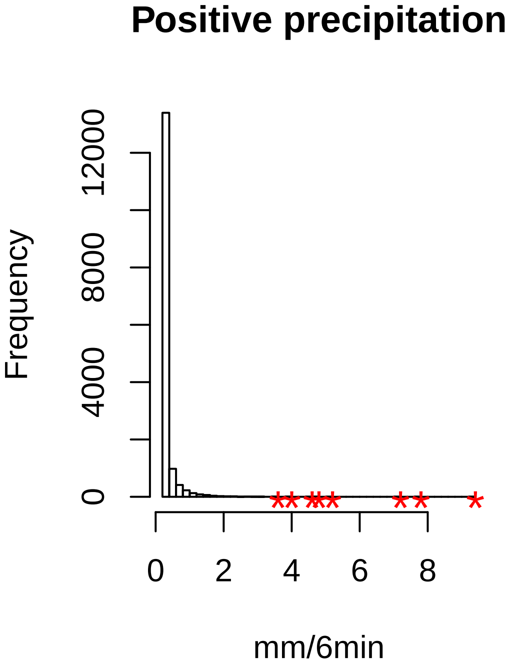

Precipitation (mm/6min)
0
1
2
3
4
5
6
7
Oct 01
Oct 16
Oct 31
Nov 15
Nov 30
Dec 15
Dec 30
Date (2006)

**Figure3.1.**



Full precipitation
Frequency
0
100000
200000
0
2
4
6
8
mm/6min

**Figure3.2.**



Positive precipitation
Frequency
0
4000
8000
12000
0
2
4
6
8
mm/6min

**Figure3.3.**

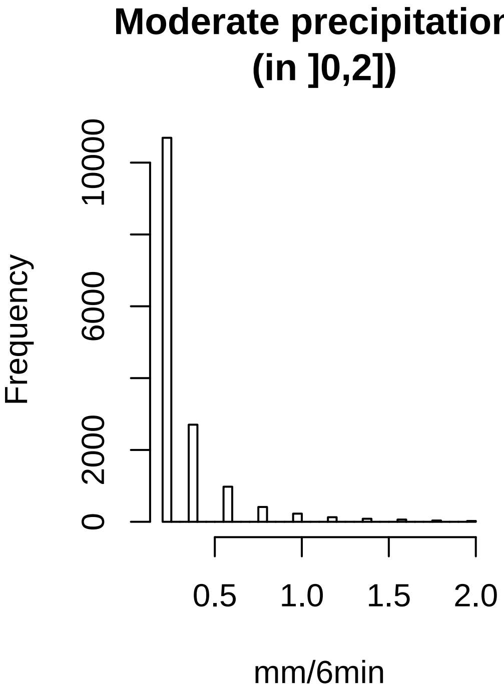

Moderate precipitation
(in ]0,2])
Frequency
0
2000
6000
10000
0.5
1.0
1.5
2.0
mm/6min

**Figure4.**

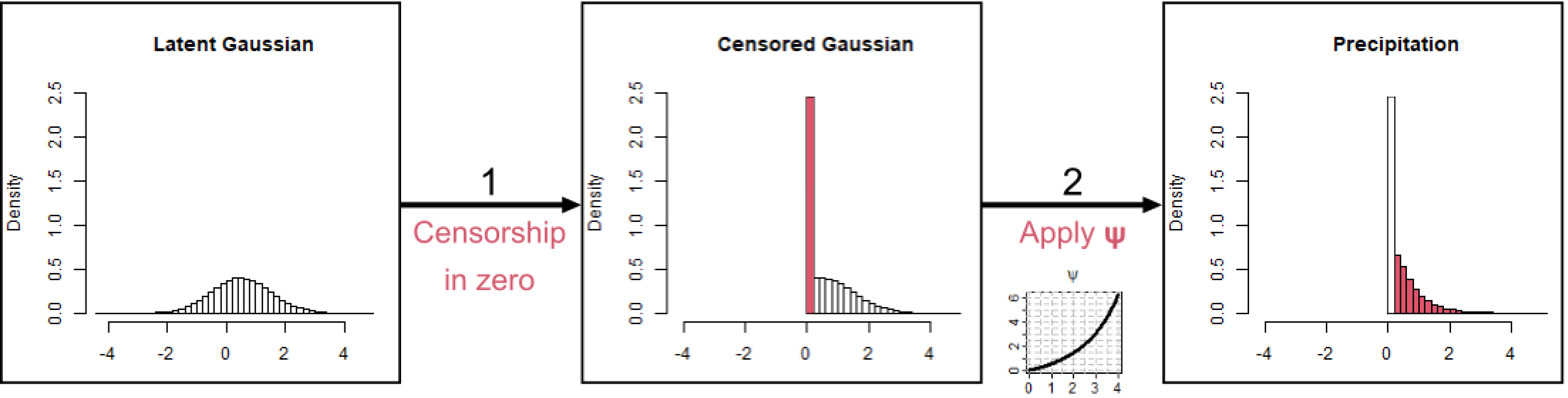

Latent Gaussian
Density
1
Censorship
in zero
Censored Gaussian
Density
2
Apply ψ
ψ
Precipitation
Density

**Figure5.**

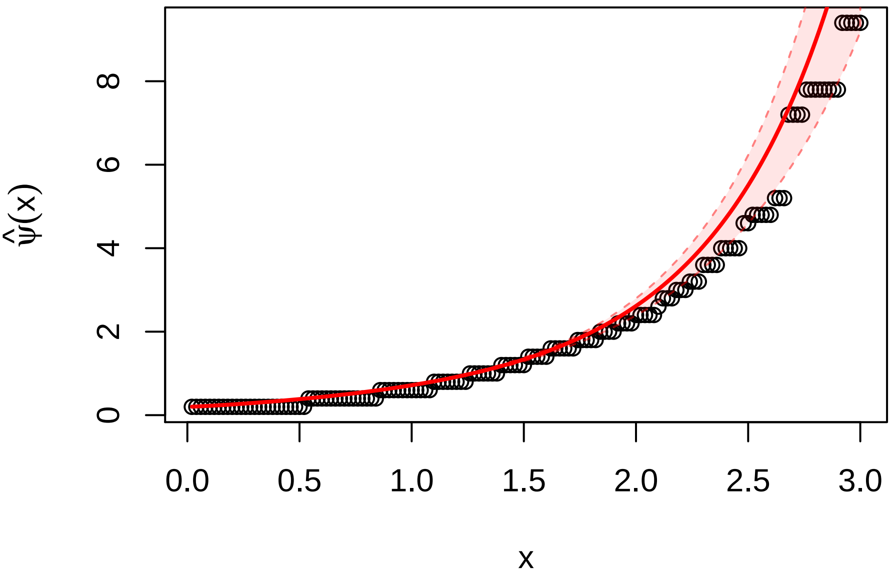
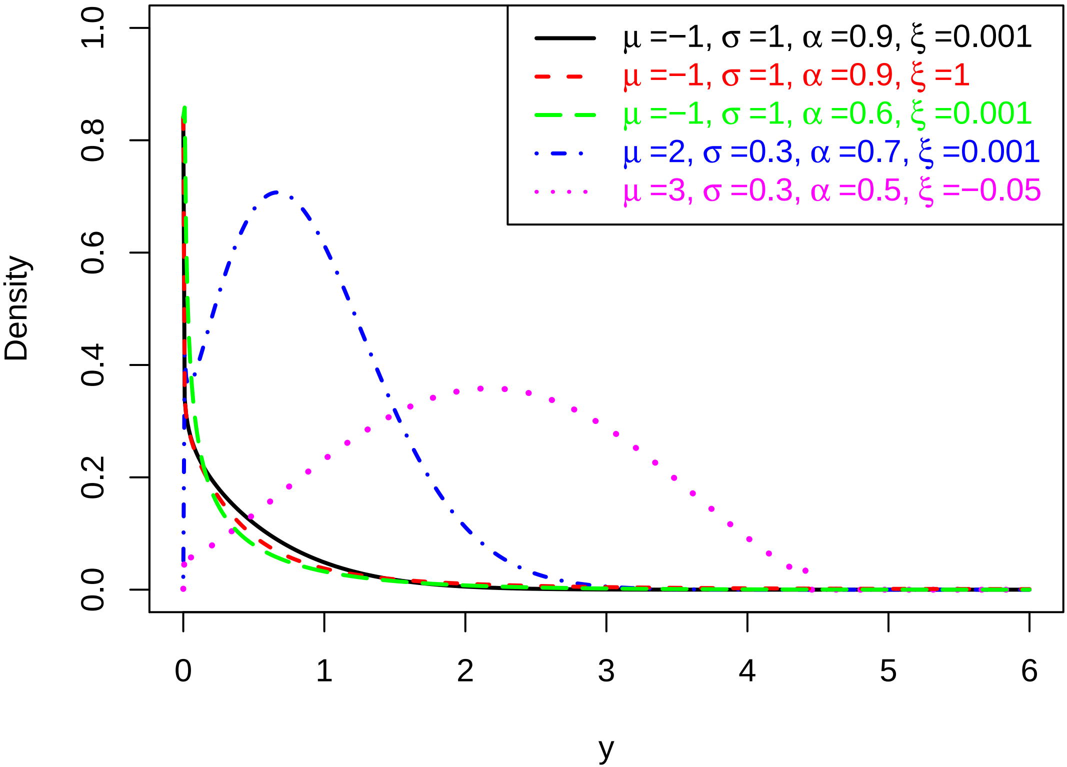

0
2
4
6
8
$\hat{\psi}(x)$
0.0
0.5
1.0
1.5
2.0
2.5
3.0
x

**Figure6.**



μ =−1, σ =1, α =0.9, ξ =0.001
μ =−1, σ =1, α =0.9, ξ =1
μ =−1, σ =1, α =0.6, ξ =0.001
μ =2, σ =0.3, α =0.7, ξ =0.001
μ =3, σ =0.3, α =0.5, ξ =−0.05
Density
y
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
6

**Figure7.1.**

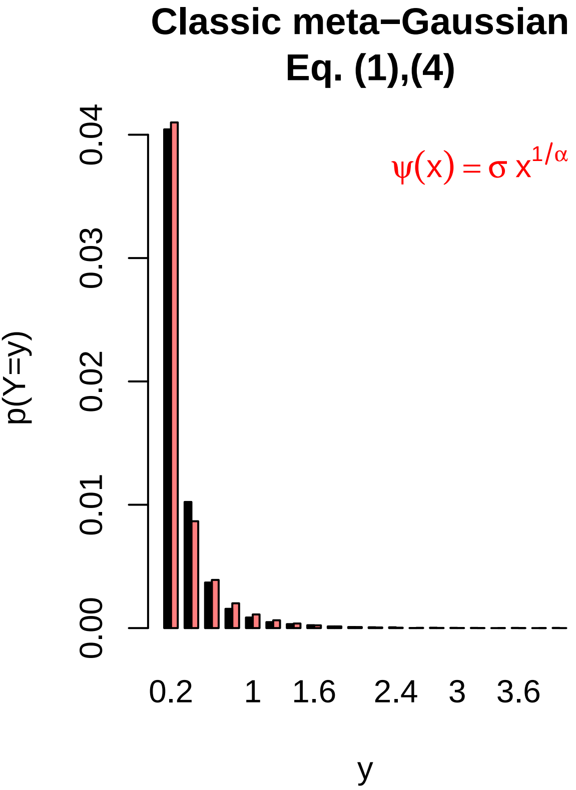

Classic meta−Gaussian
Eq. (1),(4)
$\psi(x) = \sigma\, x^{1/\alpha}$
p(Y=y)
0.00
0.01
0.02
0.03
0.04
0.2
1
1.6
2.4
3
3.6
y

**Figure7.2.**

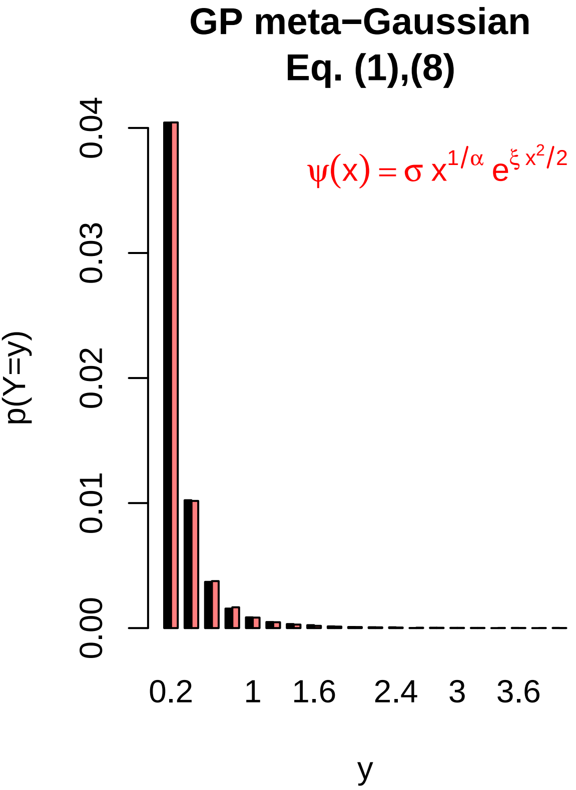

GP meta−Gaussian
Eq. (1),(8)
$\psi(x) = \sigma\, x^{1/\alpha}\, e^{\xi\, x^2/2}$
0.00
0.01
0.02
0.03
0.04
p(Y=y)
0.2
1
1.6
2.4
3
3.6
y

**Figure7.3.**

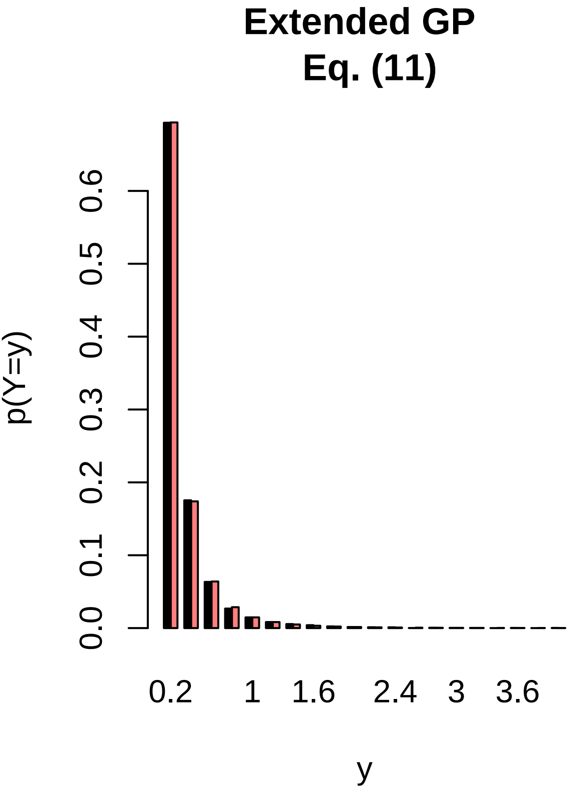
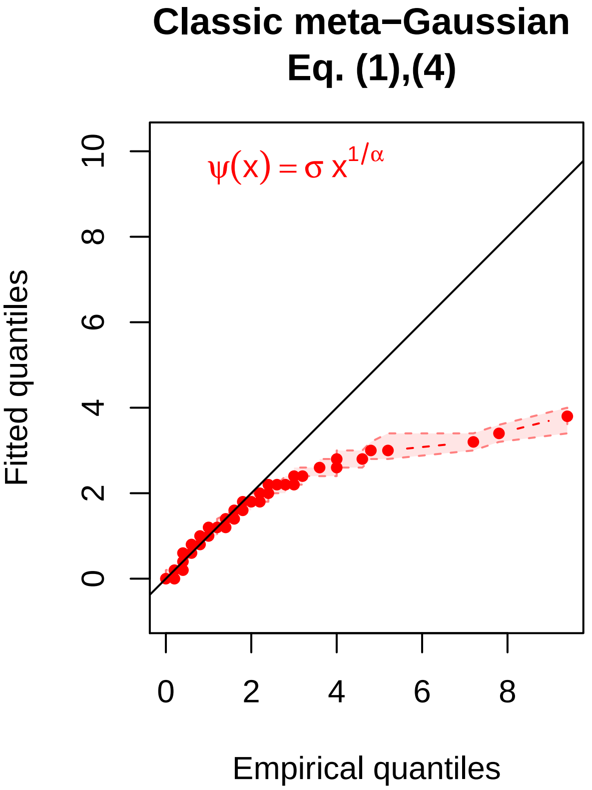

Extended GP
Eq. (11)
p(Y=y)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.2
1
1.6
2.4
3
3.6
y

**Figure7.4.**



Classic meta−Gaussian
Eq. (1),(4)
ψ(x) = σ x^{1/α}
Fitted quantiles
Empirical quantiles
0
2
4
6
8
10

**Figure7.5.**

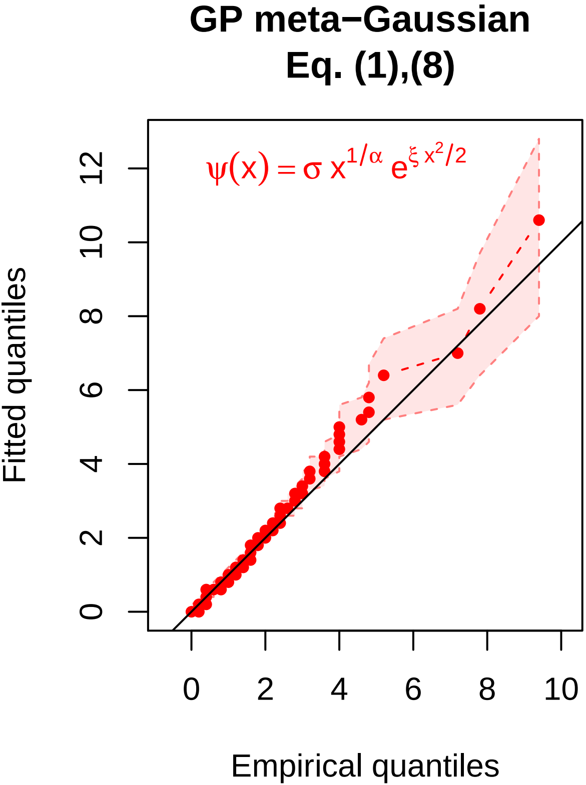

GP meta−Gaussian
Eq. (1),(8)
$\psi(x) = \sigma\, x^{1/\alpha}\, e^{\xi\, x^2/2}$
Fitted quantiles
Empirical quantiles
0
2
4
6
8
10
12

**Figure7.6.**

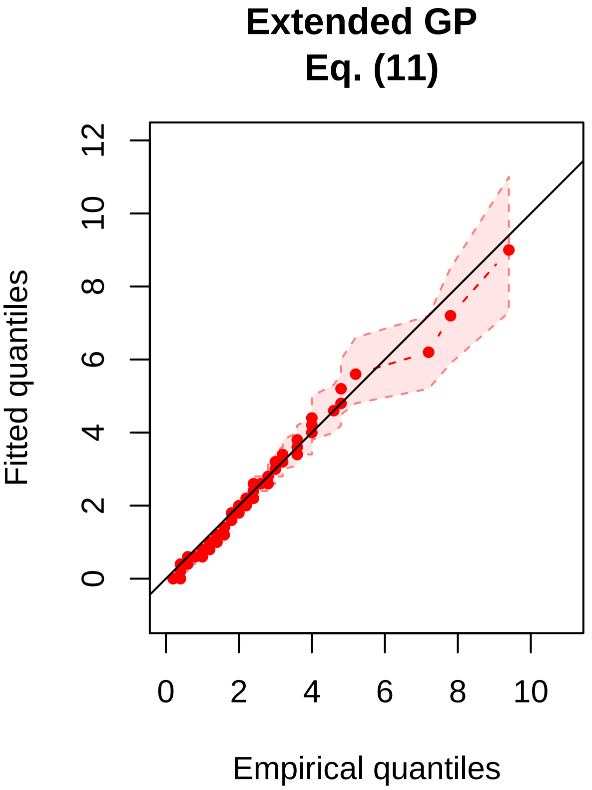

Extended GP
Eq. (11)
Fitted quantiles
Empirical quantiles
0
2
4
6
8
10
12

**Figure8.1.**

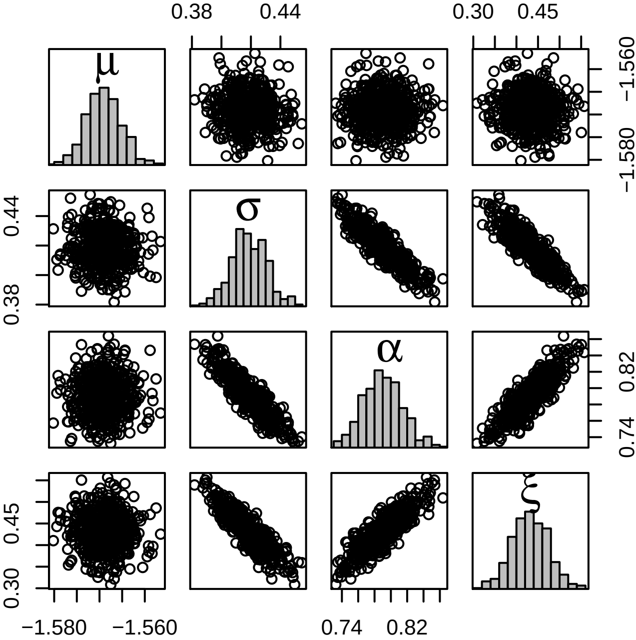

μ
σ
α
ξ
0.38
0.44
0.30
0.45
−1.560
−1.580
0.74
0.82
−1.580
−1.560
0.74
0.82
0.38
0.44
0.30
0.45

**Figure8.2.**

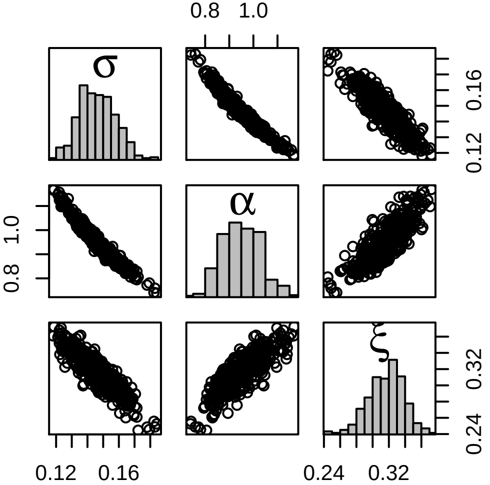
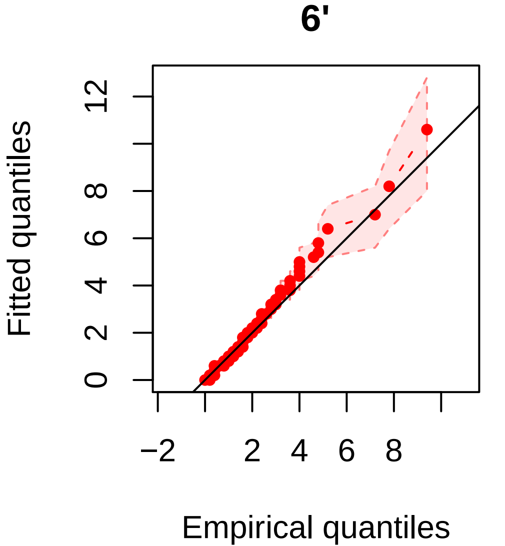

σ
α
ξ
0.8
1.0
0.16
0.12
1.0
0.8
0.32
0.24
0.12
0.16
0.24
0.32

**Figure9.1.**



6'
Fitted quantiles
Empirical quantiles
0
2
4
6
8
12
−2
2
4
6
8

**Figure9.2.**

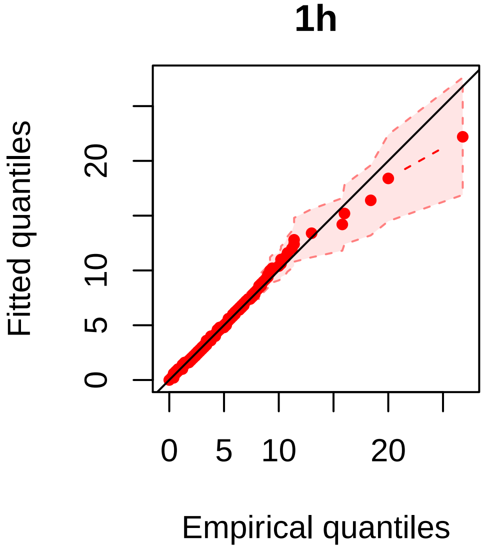

1h
Fitted quantiles
Empirical quantiles
0
5
10
20
0
5
10
20

**Figure9.3.**

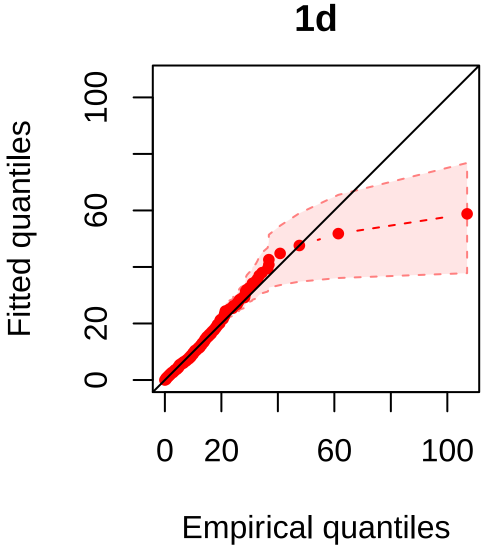
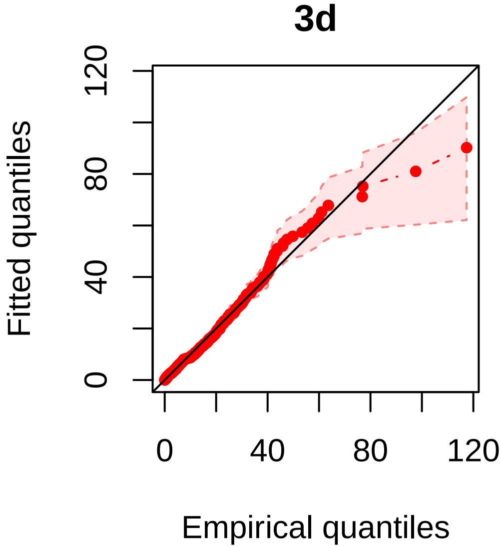

1d
Fitted quantiles
Empirical quantiles
0
20
60
100
0
20
60
100

**Figure9.4.**



3d
Fitted quantiles
Empirical quantiles
0
40
80
120
0
40
80
120

**Figure10.1.**

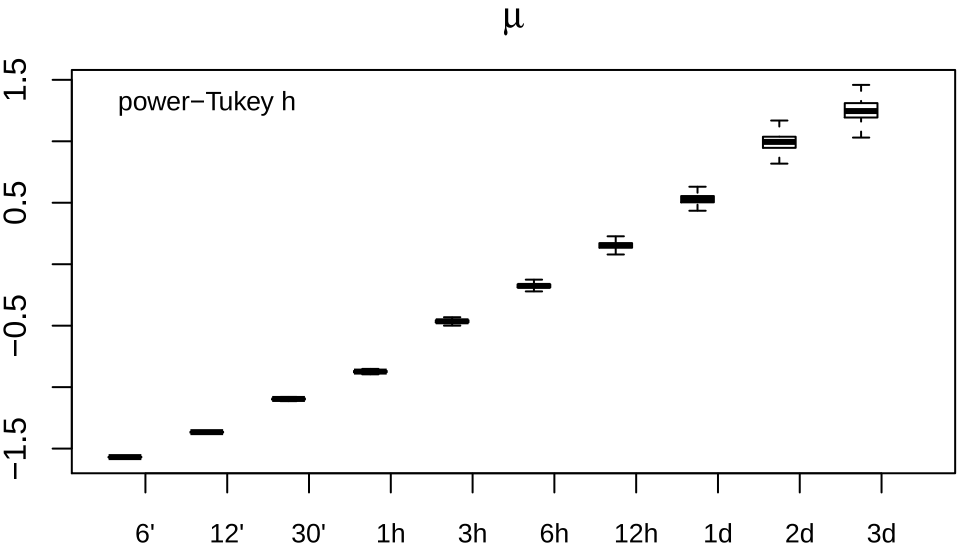

μ
power−Tukey h
1.5
0.5
−0.5
−1.5
6'
12'
30'
1h
3h
6h
12h
1d
2d
3d

**Figure10.2.**

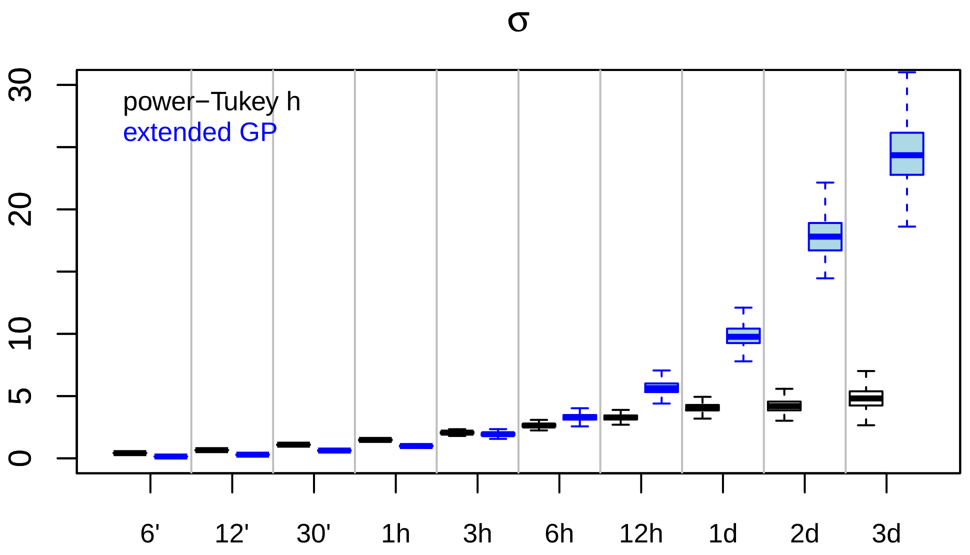

σ
power−Tukey h
extended GP
0
5
10
20
30
6'
12'
30'
1h
3h
6h
12h
1d
2d
3d

**Figure10.3.**

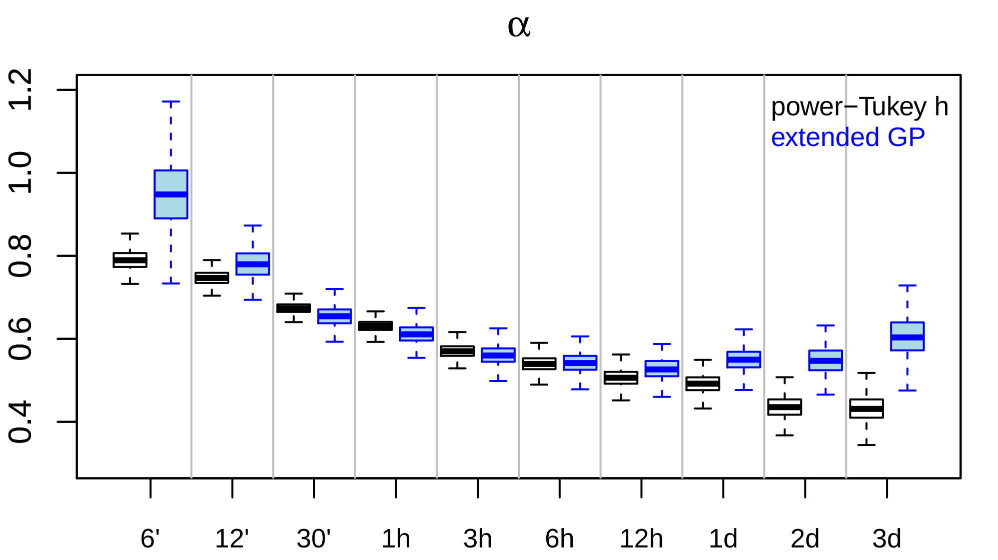

α
power−Tukey h
extended GP
0.4
0.6
0.8
1.0
1.2
6'
12'
30'
1h
3h
6h
12h
1d
2d
3d

**Figure10.4.**

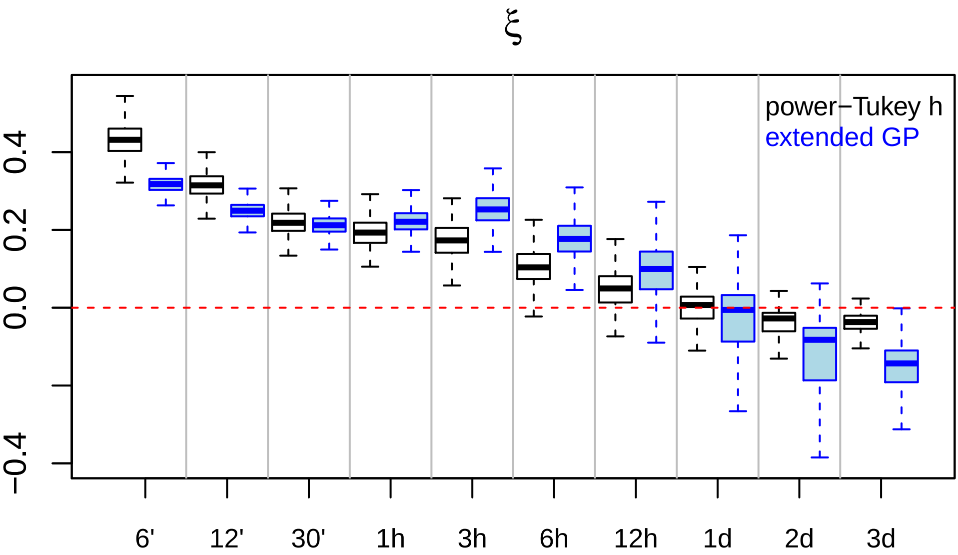
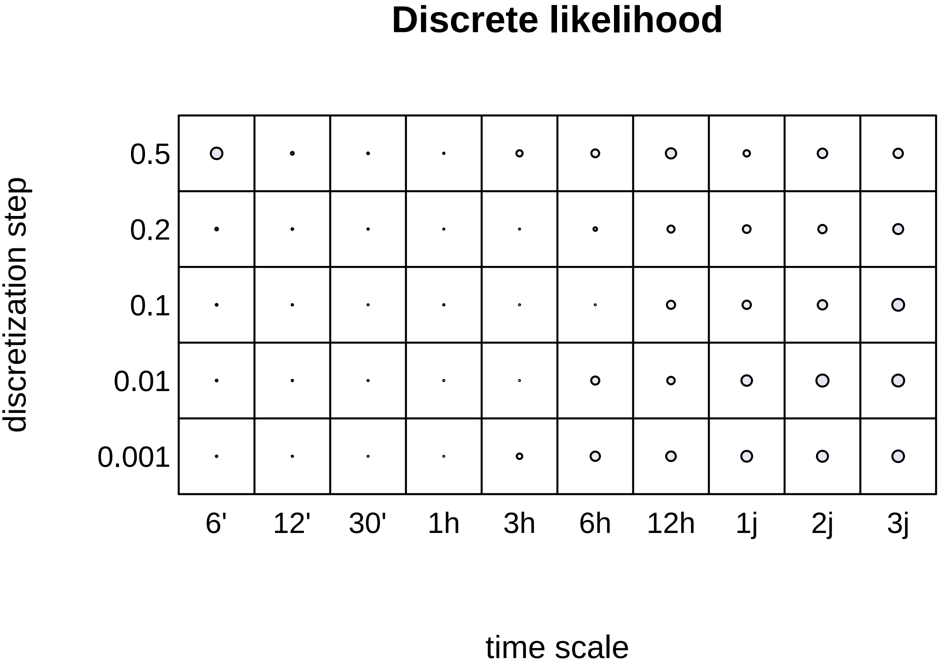
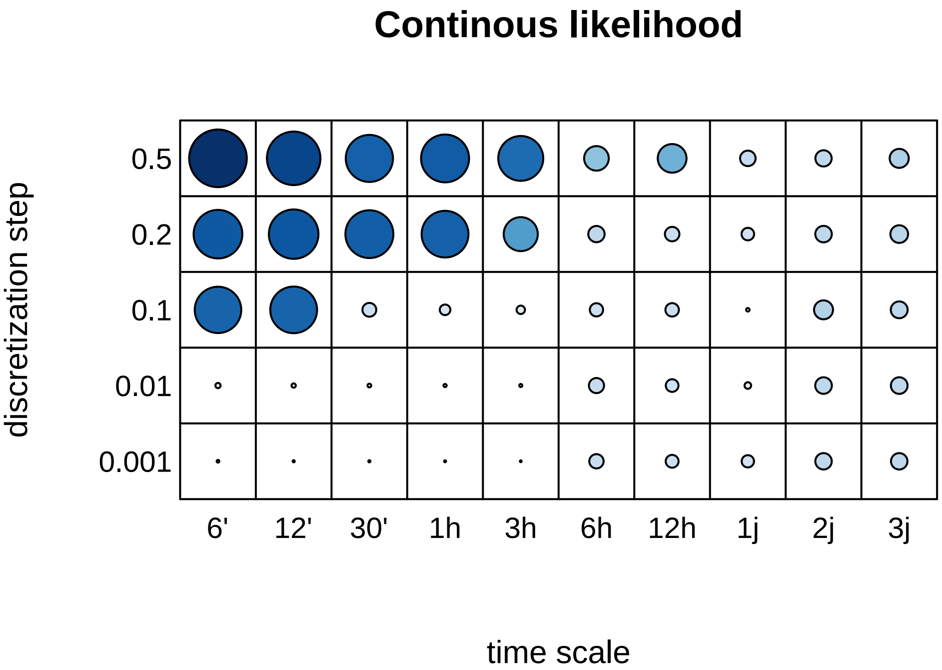

ξ
power−Tukey h
extended GP
0.4
0.2
0.0
−0.4
6'
12'
30'
1h
3h
6h
12h
1d
2d
3d

**Figure11.2.**



Discrete likelihood
discretization step
0.5
0.2
0.1
0.01
0.001
6'
12'
30'
1h
3h
6h
12h
1j
2j
3j
time scale

**Figure11.1.**


Continous likelihood
discretization step
0.5
0.2
0.1
0.01
0.001
6'
12'
30'
1h
3h
6h
12h
1j
2j
3j
time scale