

Treatise on Water Science

Chapter 27: Precipitation

Demetris Koutsoyiannis

Department of Water Resources and Environmental Engineering, Faculty of Civil Engineering, National Technical University of Athens, Heroon Polytechniou 5, GR 157 80 Zographou, Greece
Tel. +30 210 772 2831, Fax +30 210 772 2832; dk@itia.ntua.gr; <http://www.itia.ntua.gr/dk>

Andreas Langousis

Department of Civil and Environmental Engineering, Massachusetts Institute of Technology (MIT), 77 Massachusetts ave., Room 1-245, Cambridge, MA, 02139, USA
Tel. +1 (617) 253-7657, Fax +1 (440) 756-2310; andlag@alum.mit.edu; www.mit.edu/~andlag

Abstract

The study of precipitation has been closely linked to the birth of science, by the turn of the 7th century BC. Yet, it continues to be a fascinating research area, since several aspects of precipitation generation and evolution have not been understood, explained and described satisfactorily. Several problems, contradictions and even fallacies related to the perception and modelling of precipitation still exist. The huge diversity and complexity of precipitation, including its forms, extent, intermittency, intensity, and temporal and spatial distribution, do not allow easy descriptions. For example, while atmospheric thermodynamics may suffice to explain the formation of clouds, it fails to provide a solid framework for accurate deterministic predictions of the intensity and spatial extent of storms. Hence, uncertainty is prominent and its understanding and modelling unavoidably relies on probabilistic, statistical and stochastic descriptions. However, the classical statistical models and methods may not be appropriate for precipitation, which exhibits peculiar behaviours including Hurst-Kolmogorov dynamics and multifractality. This triggered the development of some of the finest stochastic methodologies to describe these behaviours. Inevitably, because deduction based on deterministic laws becomes problematic, as far as precipitation is concerned, the need for observation of precipitation becomes evident. Modern remote sensing technologies (radars and satellites) have greatly assisted the observation of precipitation over the globe, whereas modern stochastic techniques have made the utilization of traditional raingauge measurements easier and more accurate. This chapter reviews existing knowledge in the area of precipitation. Interest is in the small- and large-scale physical mechanisms that govern the process of precipitation, technologies and methods to estimate precipitation in both space and time, and stochastic approaches to model the variable character of precipitation and assess the distribution of its extremes.

Keywords

hydrology; meteorology; climatology; hydrological modelling; hydrological design; thermodynamics; Hurst-Kolmogorov dynamics; multifractals; atmosphere; precipitation; rainfall; snow; flood; raingauges; remote sensing; radars; satellites.

1 Introduction

1.1 The entrancement of precipitation

Precipitation and its related phenomena such as cloud formation and movement, thunders, rainbow, are spectacular (Fig. 1) due to their huge diversity and complexity. This complexity makes them difficult to comprehend, model and predict. Hence, it is understandable that ancient civilizations explained these phenomena in a hyperphysical manner, assuming that deities were responsible for their creation. For example, in the Greek mythology some of the phenomena were deified (e.g. Iris is the name of a goddess as well as of the rainbow) whereas the most impressive among them, thunders in particular, were attributed to the action of the King of the Gods, Zeus (Jupiter in the Roman mythology; similar deities are Indra in Hinduism, Thor in Norse, etc.). Demystification of these processes and formation of the physical concept of the hydrological cycle is closely related to the birth of science, by the turn of the 7th century BC. While the hydrological cycle was founded as a concept in the 6th century B.C. by Anaximander, Anaximenes and Xenophanes, and was advanced later by Aristotle (Koutsoyiannis *et al.*, 2007), certain aspects related to precipitation can be understood only within the frame of modern science. The fact that a solid or liquid hydrometeor resists gravity and suspends in the atmosphere in a cloud is counterintuitive and needs advanced knowledge of physics, fluid dynamics and statistical thermodynamics to be understood and modelled.

The complexity of the processes involved in precipitation, and their enormous sensitivity to the initial conditions (where tiny initial differences produce great differences in the final phenomena), retain, to this day, some of the ancient mythical and magical magnificence of the societal perception of precipitation. People still believe in hyperphysical interventions in matters concerning precipitation. As put by Henri Poincaré (1908), father of the notion of “chaos”:

Why do the rains, the tempests themselves seem to us to come by chance, so that many persons find it quite natural to pray for rain or shine, when they would think it ridiculous to pray for an eclipse?

Amazingly, however, and at the very same time, there is little disbelief in some climate modellers' prophecies (or outputs of global circulation models—GCMs) of the precipitation regimes over the globe in the next 100 years or more. This indicates an interesting conflict between perceptions of precipitation: that it is so unstable, uncertain and unpredictable that prayers are needed to invoke precipitation, and that for some scientists the future evolution of precipitation on Earth is still predictable in the long run. The latter belief concerns not only the general public, but also the scientific community. For example, a Google Scholar search with either of the keywords “precipitation” or “rainfall” plus the keywords “climate change” and “GCM” locates 21 700 publications (as of August 2009), of which about 200 have been cited 100 times or more. This long list of results appears despite the fact that climate modellers themselves admit to the performance of their models being low, as far as precipitation is concerned (Randall *et al.*, 2007). An independent study by Koutsoyiannis *et al.* (2008), which compares model results for the 20th century with historical time series, has shown that the models are not credible at local scales and do not provide any basis for assessment of future conditions. These findings demonstrate that, even today, the perception of precipitation, not only by the general public, but even by scientists specialized in the

study of precipitation, meteorologists, climatologists and hydrologists, continues to be contradictory, problematic and, in some sense, mysterious.



Fig. 1 Precipitation and related phenomena (from upper-left to lower-right): Monsoon rainfall (Pune, India, September 2009; photo by first author); snowy mountainous landscape (Mesounta, Greece, December 2008; from www.mesounta.gr/mesounta/ist_eik1/07_xion_03.htm); thunder in the Parthenon (Athens, Greece, 30th June 2010; from <http://www.dailymail.co.uk/news/worldnews/article-1290289/Greece-lightning-Ancient-Parthenon-lit-storm-breaks-Athens.html>); rainbow (Mystras, Greece, April 2008, from laspistasteria.wordpress.com/2008/04/08/rainbow-3/).

1.2 Forms of precipitation

Precipitation occurs in a number of forms, either liquid or solid or even mixed (sleet). Liquid precipitation includes *rainfall* and *drizzle*, where the former is the most common and most significant, and the latter is characterized by much smaller drop sizes and lighter intensity. *Dew* is another liquid form, formed by condensation of water vapour (mostly at night) at cold surfaces (e.g. on tree leaves).

Most important among the solid forms of precipitation are *snow* and *hail*. At high latitudes or at high altitudes snow is the predominant form of precipitation. Snowfall may occur when the temperature is low and snow accumulates at the ground until the temperature rises sufficiently for it to melt. On the other hand, hail may fall in relatively high temperature and usually melts rapidly. While hailstones are amorphous and usually large (one to several centimetres diameter), snowflakes are

symmetrical and visually appealing with a tremendous variety of shapes, so that no two snowflakes are the same.

Occult precipitation is induced when clouds or fog is formed in forested areas and it includes liquid (*fog drip*) and solid (*rime*) forms. Fog drip occurs when water droplets are deposited on vegetative surfaces, and the water drips to the ground. Rime is formed when supercooled air masses encounter exposed objects, such as trees, that provide nucleation sites (see sections 2.1 and 2.2) for formation and buildup of ice, much of which may fall to the ground in solid or liquid form. In some places (e.g. in humid forested areas) precipitation of this type may reach significant amounts; for example rime constitutes about 30% of the annual precipitation in a Douglas fir forest in Oregon (Harr, 1982; Dingman, 1994) and about 30% of total precipitation in fir forested mountainous areas of Greece (Baloutsos *et al.*, 2005), and it is the sole precipitation type in the rainless coast of Peru (Lull, 1964; Dingman, 1994).

1.3 Precipitation metrics

The principal metric of precipitation is the rainfall depth h (commonly expressed in mm) that falls at a specific point during a certain period of time t ; this can be easily perceived and measured by a bucket exposed to precipitation. A derivative quantity is the precipitation intensity,

$$i := \frac{dh}{dt} \quad (1)$$

with units of length divided by time (typically mm/h, mm/d, mm/year). Since it cannot be measured directly (at an instantaneous time basis), it is typically approximated as

$$i = \frac{\Delta h}{\Delta t} \quad (2)$$

where Δh is the change of the depth in a finite time interval Δt . The intensity derived from equation (2) is a time averaged value—but at a point basis. Spatial averaging at various scales is always very useful as can be seen in section 1.4. This averaging needs precipitation measurement at several points, followed by appropriate numerical integration methods (see section 3.1). While the traditional precipitation measurement networks are sparse, thus making the estimation of areal precipitation uncertain, in the recent decades new measurement techniques have been developed implementing radar and satellite technologies (sections 3.2 and 3.3). These provide a detailed description of the spatial distribution of precipitation, thus enabling a more accurate estimation. The latter techniques inherently involve the study of other metrics of precipitation such as the distribution of the size, velocity and kinetic energy of precipitation particles, and the so called radar reflectivity (section 3.2).

Furthermore, the quantitative description of the processes related to the fall, accumulation and melting of snow, involve a number of additional metrics, such as the snowfall depth (new snow falling), the snowcover depth or snowpack depth (the depth of snow accumulated at a certain point at a particular time), the snow density ρ_s , and the water equivalent of snowfall or snowpack, defined as $h := h' \rho_s / \rho_w$, where h' is the snowfall or the snowcover depth and $\rho_w = 1000 \text{ kg/m}^3$ is the liquid water density. Typical values of ρ_s for snowfall range between 0.07-0.15 ρ_w (e.g. Dingman, 1994) but a commonly used value is $\rho_s = 0.1 \rho_w = 100 \text{ kg/m}^3$. For this value, a snowfall depth of, say, 10 cm,

corresponds to a precipitation water equivalent of 10 mm. The density of snowpack is generally larger than $0.1\rho_w$ (because of compaction due to gravity and other mechanisms) and depends on the elapsed time and snowpack depth. After a few days it is about $0.2\rho_w$, whereas after some months it may become about $0.4\rho_w$.

1.4 The enormous variability of precipitation

The different phases and forms of precipitation, and the different shapes of precipitation particles (drops, flakes, hailstones) are just a first indicator of the great diversity of precipitation phenomena. At a macroscopic level and in quantitative terms, this diversity is expressed by the enormous variability of the precipitation process, in space and time, at all spatial and temporal scales. Intermittency is one of the aspects of variability, but even in areas or time periods in which precipitation is nonzero, the precipitation depth or average intensity are highly variable.

Fig. 2 shows the spatial variability of precipitation over the globe in mm/d at climatic scale (average for the 30-year period 1979-2008) and at an annual scale (average for the year 2006), based mostly on satellite data (see figure caption and section 3.3). While the average precipitation rate over the globe and over the specified 30-year period is 2.67 mm/d or 977 mm/year, we observe huge differences in different areas of the globe. In some areas, mostly in tropical seas and in equatorial areas of South America and Indonesia, this rate exceeds 10 mm/d or 3.65 m/year. On the other hand, in large areas in the subtropics, where climate is dominated by semi-permanent anticyclones, precipitation is lower than 1 mm/d or 365 mm/year. Significant portions of these areas in Africa, Australia, and America are deserts, where the average precipitation is much lower than 1 mm/d. In addition, in polar regions, where the available atmospheric moisture content is very low due to low temperature (see section 2.1 and Fig. 14) the amounts of precipitation are very small or even zero. For example, it is believed that certain dry valleys in the interior of Antarctica have not received any precipitation during the last two million years (Uijlenhoet, 2008).

Fig. 3 depicts the zonal precipitation profile and shows that the climatic precipitation rate at an annual basis is highest at a latitude of 5°N, reaching almost 2000 mm/year and has a second peak of about 1500 mm/year at 5°S. Around the Tropics of Cancer and Capricorn, at 23.4°N and S, respectively, the rainfall rate displays troughs of about 600 mm/year whereas at mid latitudes, between 35° and 60° both N and S, rainfall increases again and remains fairly constant, close to the global average of 977 mm/year. Then, toward the poles, it decreases to about 100 mm in Antarctica and slightly more, 150 mm/year in the Arctic. Fig. 3 also shows monthly climatic profiles for the months of January and July. It can be seen that the rainfall conditions for the two months are quite different, with the largest differences appearing at about 15°N and S and the smallest at about 30°N and S. Below 30° in the Northern Hemisphere, as well as above the Arctic Circle (66.6°), rainfall is higher during the summer (July) than during winter (January) but at mid latitudes this relationship is reversed. Similar conditions are met in the Southern Hemisphere (where July is winter and January is summer).

In both Fig. 2 and Fig. 3, apart from climatic averages, the specific values for a certain year, namely 2006, are also shown. We observe that there are differences in the climatic values, manifesting temporal variability over the different years. This variability seems to be lower in comparison to the spatial variability over the globe, as well as to the seasonal variability reflected in the profiles of different months. However, while the spatial variability over the globe and the seasonal variability

are well comprehended and roughly explainable in terms of basic physical and astronomical knowledge (i.e. solar radiation, relationship of temperature and atmospheric moisture content, motion of Earth), in other words they are “regular”, the interannual variability is “irregular”, and difficult or even impossible to predict.

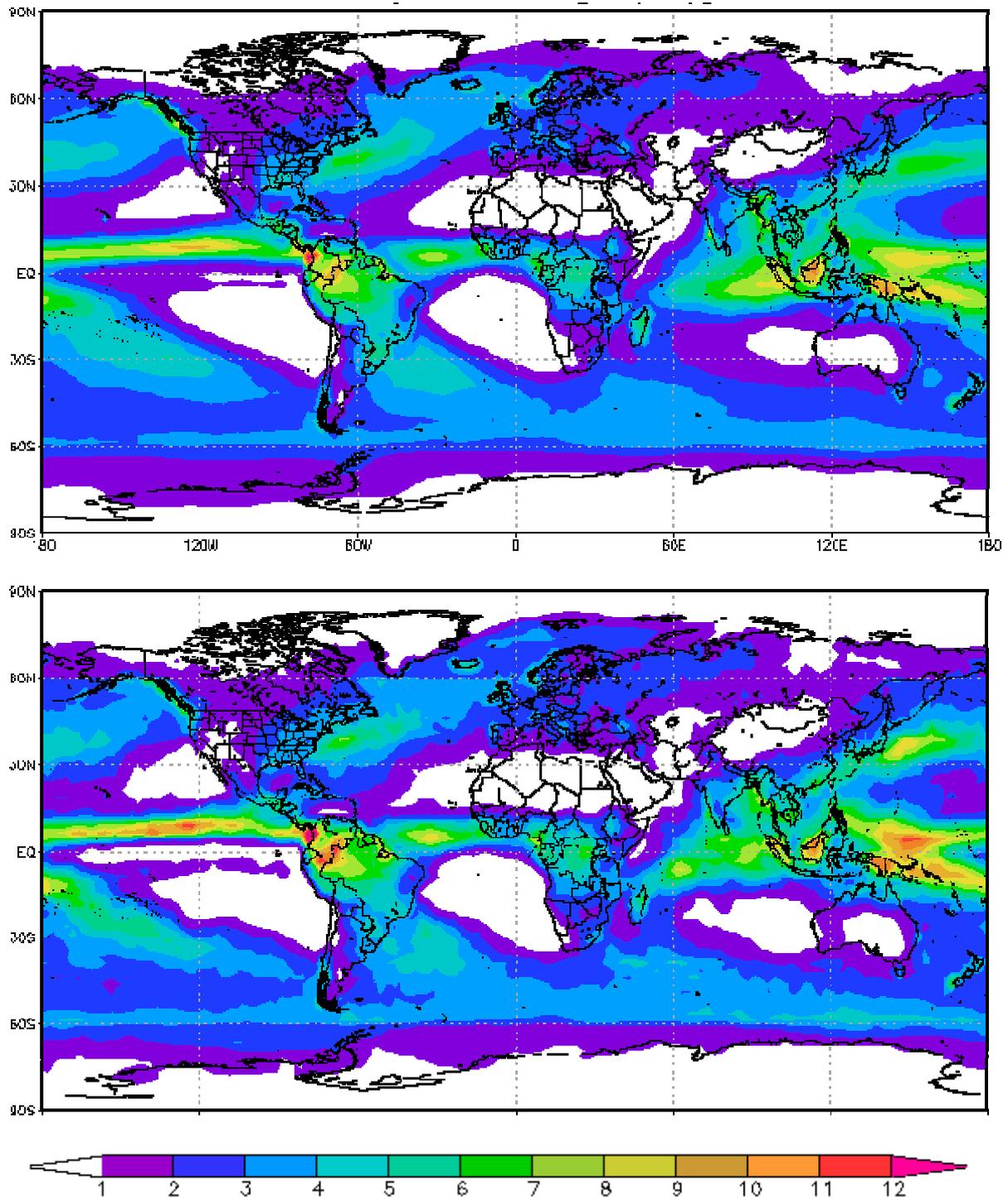


Fig. 2 Precipitation distribution over the globe in mm/d, (upper) at a climatic scale (average for the 30-year period 1979-2008) and (lower) at an annual scale (average for year 2006). Data and image generation due to the Global Precipitation Climatology Project (GPCP) made available by NASA (disc2.nascom.nasa.gov/Giovanni/tovas/rain.GPCP.2.shtml); resolution $2.5^{\circ} \times 2.5^{\circ}$.

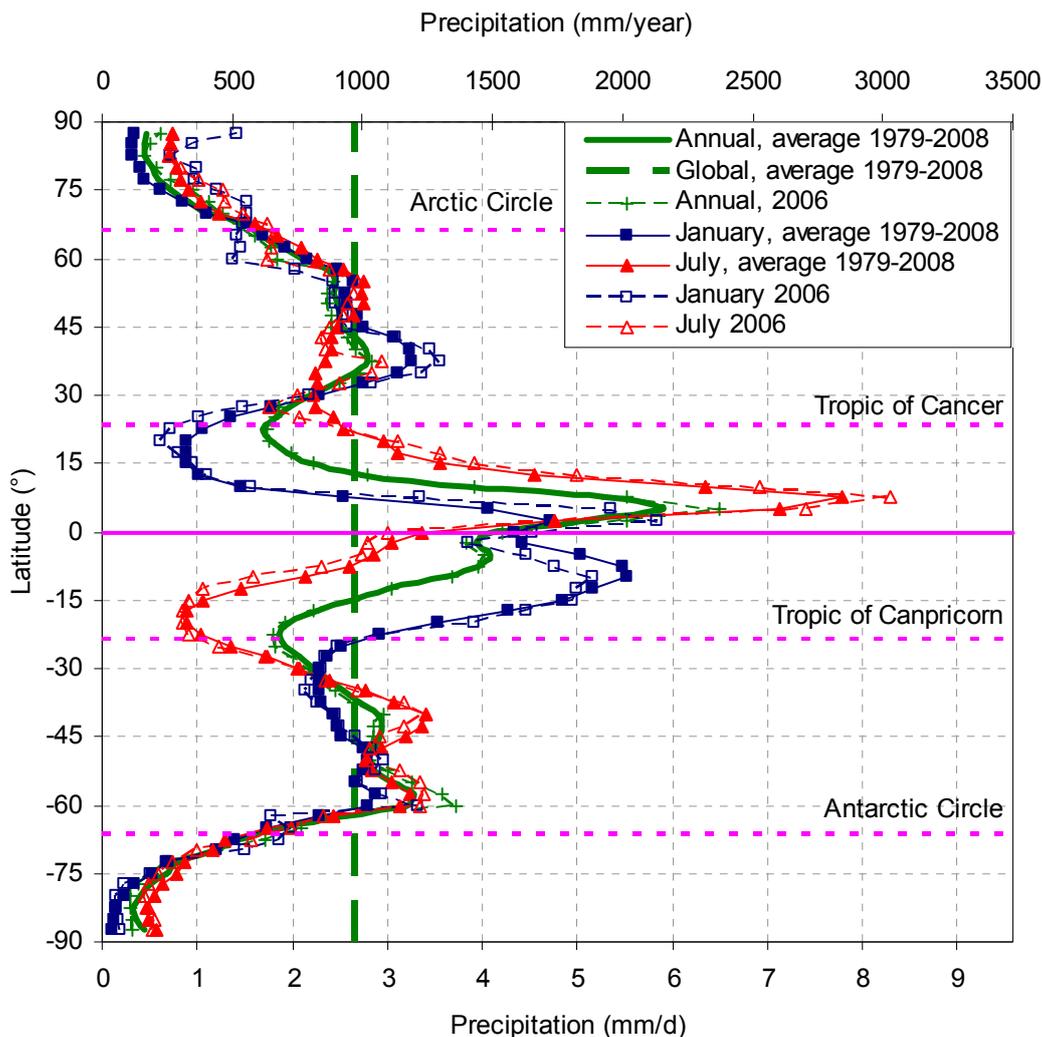


Fig. 3 Zonal precipitation profile: precipitation averaged over all longitudes for different latitudes varying from 90°N to 90°S (–90°N); data from GPCP.

Such irregular variability appears also at finer time scales as well as at finer spatial scales. In fact, as easily understood from elementary statistics, as the spatial and/or temporal scale becomes finer, the variability increases. Fig. 4 demonstrates how the variability of the spatial distribution of rainfall at a monthly temporal scale (January 2006) increases when the spatial scale decreases from 2.5°×2.5° (upper panel) to 0.25°×0.25° (lower panel). Clearly, the areas of equal rainfall amount (including areas of negligible rainfall i.e. < 1 mm/d ≈ 0.04 mm/h), which are smooth in the upper panel become rough and erratic in the lower panel. Also the maximum observed rainfall is 21 mm/d (monthly amount 651 mm) in the upper panel and 1.2 mm/h (monthly amount 893 mm) in the lower panel.

Fig. 5 demonstrates the increasing variability with decreasing time scale. Specifically, it depicts how the image of the rainfall distribution changes at a daily scale (9 January 2006) and at a sub-daily scale, at 3-hourly intervals of the same day. The differences between Fig. 5 and Fig. 4 are prominent. Especially at the 3-hourly scale a vast part of the globe receives no rainfall, and the part that receives rainfall is irregularly distributed, yet not resembling a totally random pattern. The maximum observed rate during this 3-hourly interval is 22 mm/h, about 18 times higher than the maximum rate at the monthly scale shown in Fig. 4. The lowest panels of Fig. 5 provide a zoom in over the area

lying between 9°N-5°S and 78-92°E, which is located in the Indian ocean south-east of Sri Lanka. This area received a large amount of rainfall this particular day, with a rate that is non uniform in space and time.

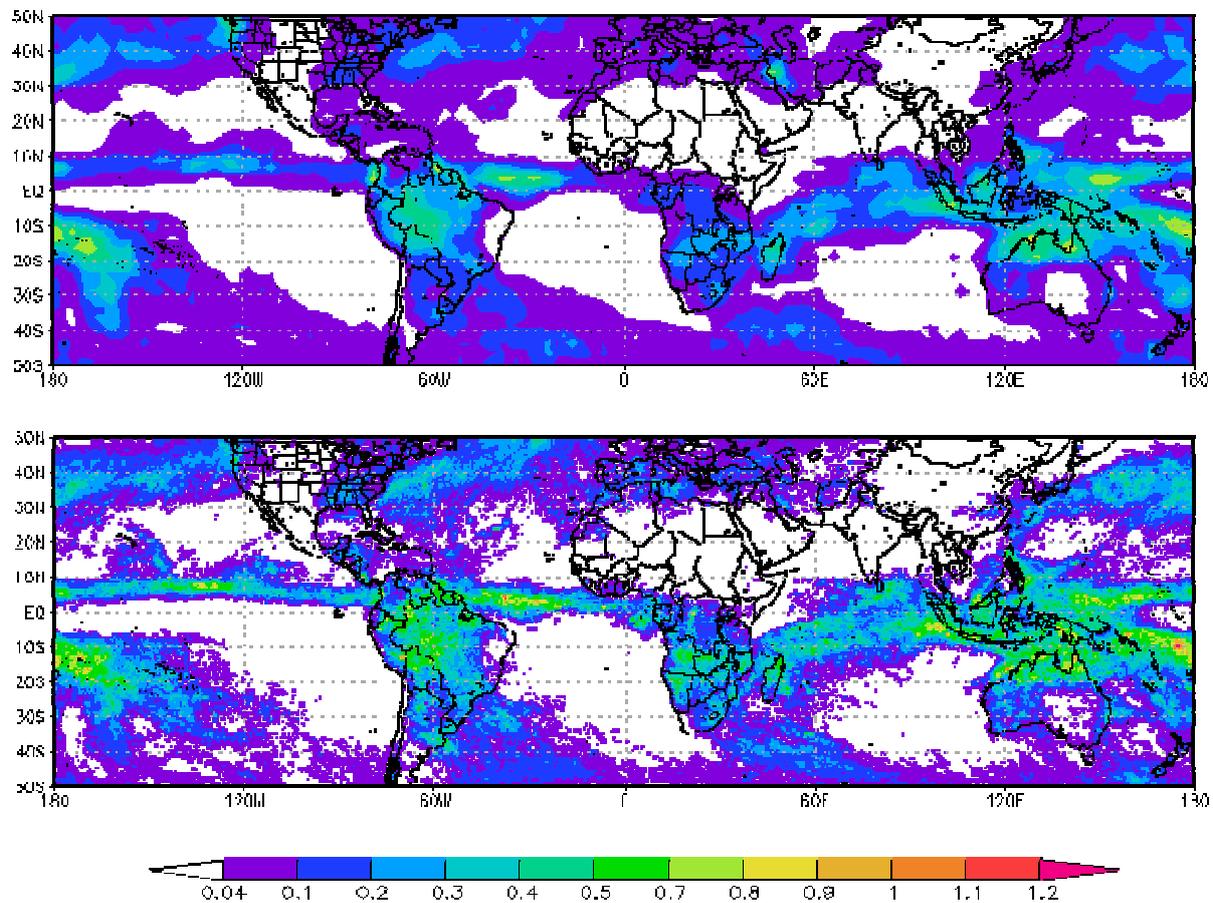


Fig. 4 Monthly rainfall distribution over the globe in January 2006 in mm/h: (upper) data for resolution $2.5^{\circ}\times 2.5^{\circ}$ from GPCP; (lower) data for resolution $0.25^{\circ}\times 0.25^{\circ}$ from the Tropical Rainfall Measuring Mission (TRMM) and Other Rainfall Estimate (3B42 V6) archive, made available by NASA (disc2.nascom.nasa.gov/Giovanni/tovas/TRMM_V6.3B42.shtml).

Fig. 6 to Fig. 8 focus on the temporal variability of precipitation. Fig. 6 depicts the monthly and annual variation of the average precipitation over the globe. We can see that at both scales the variability is remarkable. Thus, the annual precipitation in the last 30 years has varied between 957 and 996 mm and obviously much higher variation should have occurred in the past—but data of this type covering the entire globe do not exist for earlier periods. However, we can get an idea of earlier variation using raingauge data (see section 3.1) at certain locations.

Perhaps the oldest systematic observations of rainfall quantity in the world were made in Korea in the 15th century. Rainfall records for the city of Seoul (37.57°N, 126.97°E, 85 m) exist since 1770 and are considered to be reliable (Arakawa, 1956; Wang *et al.*, 2006, 2007). The recorded annual rainfall in Seoul is plotted in Fig. 7 along with running climatic averages at 10-year and 30-year time scales. The data are now available at a monthly scale from the climatic data base of the Dutch Royal Netherlands Meteorological Institute (KNMI), while the monthly data for 1770-1907 appear also in Arakawa (1956). Comparisons show that the two time series are generally consistent but not

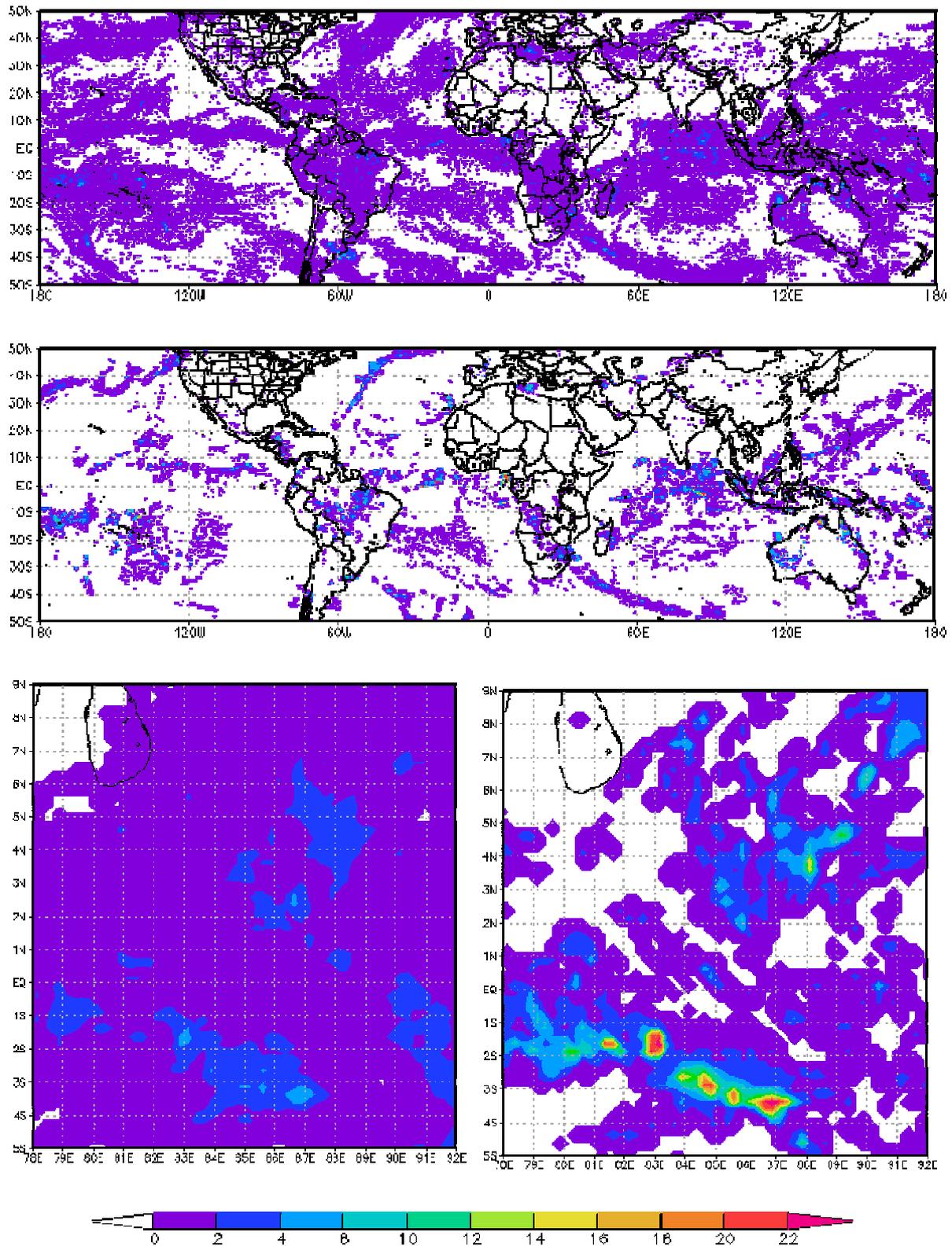


Fig. 5 Spatial rainfall distribution at daily and sub-daily scale: (upper) daily rainfall over the zone between 50°N and 50°S on 9 January 2006; (middle) 3-hourly rainfall at 09:00 in the same day; (lower left) zoom in of the upper panel for daily rainfall in the Indian ocean south-east of Sri Lanka (shown in figure); (lower right) zoom in of the middle panel for 3-hourly rainfall for the same area. Data in mm/h for resolution 0.25°×0.25° from the TRMM 3B42 V6 archive.

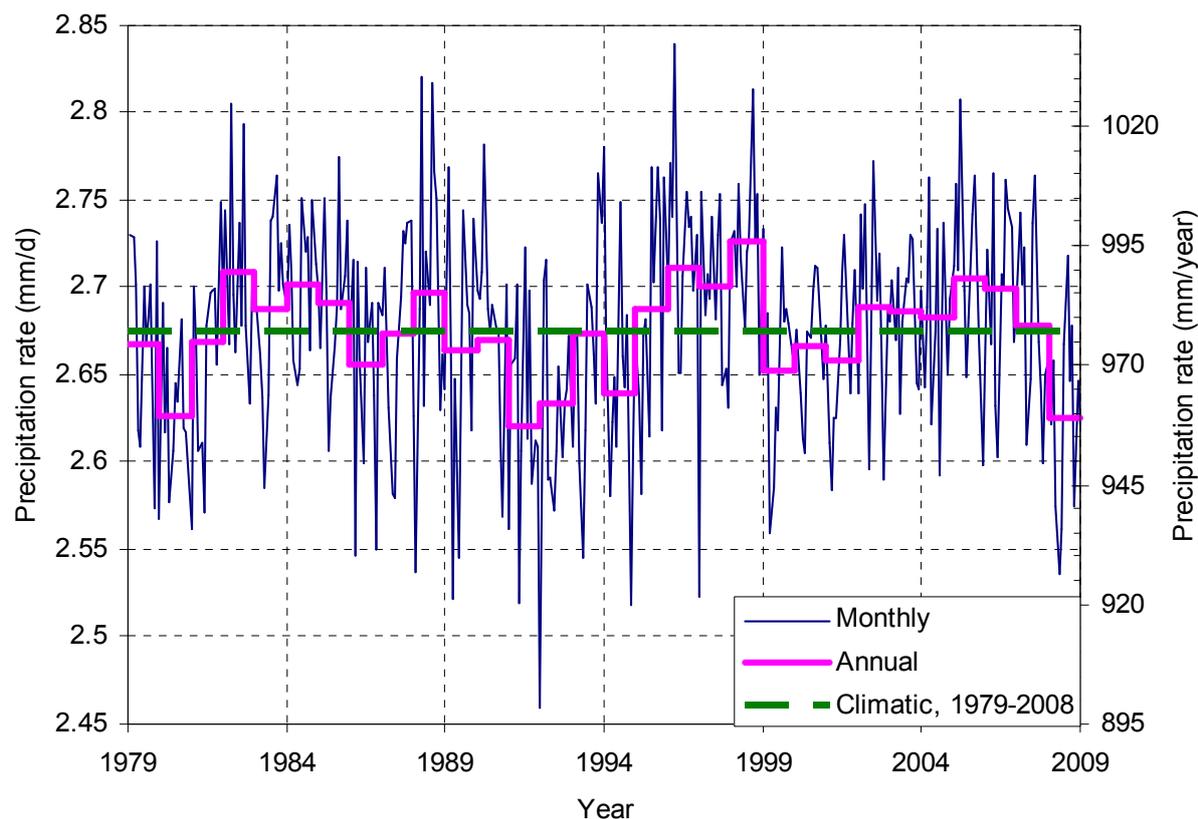


Fig. 6 Evolution of the globally averaged monthly precipitation in the 30-year period 1979-2008, based on data from GPCP (see caption of Fig. 2).

identical. The more modern data series has a few missing values, which generally correspond to high values of the older version (and it has been common practice in hydrometeorological data processing to delete very high values or outliers, which are regarded suspect, see section 3.1.2). In the time series plotted in Fig. 7 these gaps have been filled in using the values of the older time series, and a few other missing values have been filled in with the average of the four nearest monthly values of the same month (see justification in section 3.1.3). The plot shows that during the 238 years of record the annual rainfall varied between 634 and 3057 mm and the climatic 30-year average varied between 1139 and 1775 mm. These figures indicate a huge variability: the maximum observed annual rainfall is almost 5 times greater than the minimum and the maximum 30-year climatic rainfall is 55% higher than the minimum. Such observed changes underscore the ever-changing character of climate and render future changes of precipitation predicted by climate modellers (which typically vary within 10-20%; compare Fig. 10.12, upper left panel, in Meehl *et al.*, 2007, with Fig. 2 herein) to be unrealistically low and too unsafe to support planning.

Fig. 7 also includes a plot of another long time series, for the Charleston City, USA (32.79°N, 79.94°W, 3 m); the record starts in 1835. The time series is also available from the KNMI data base and a few missing monthly values have been filled in by the average of the four nearest monthly values of the same month. Here the annual rainfall varied between 602 and 1992 mm (3.3 times higher than minimum) and the climatic 30-year average varied between 1135 and 1425 mm (25% higher than minimum).

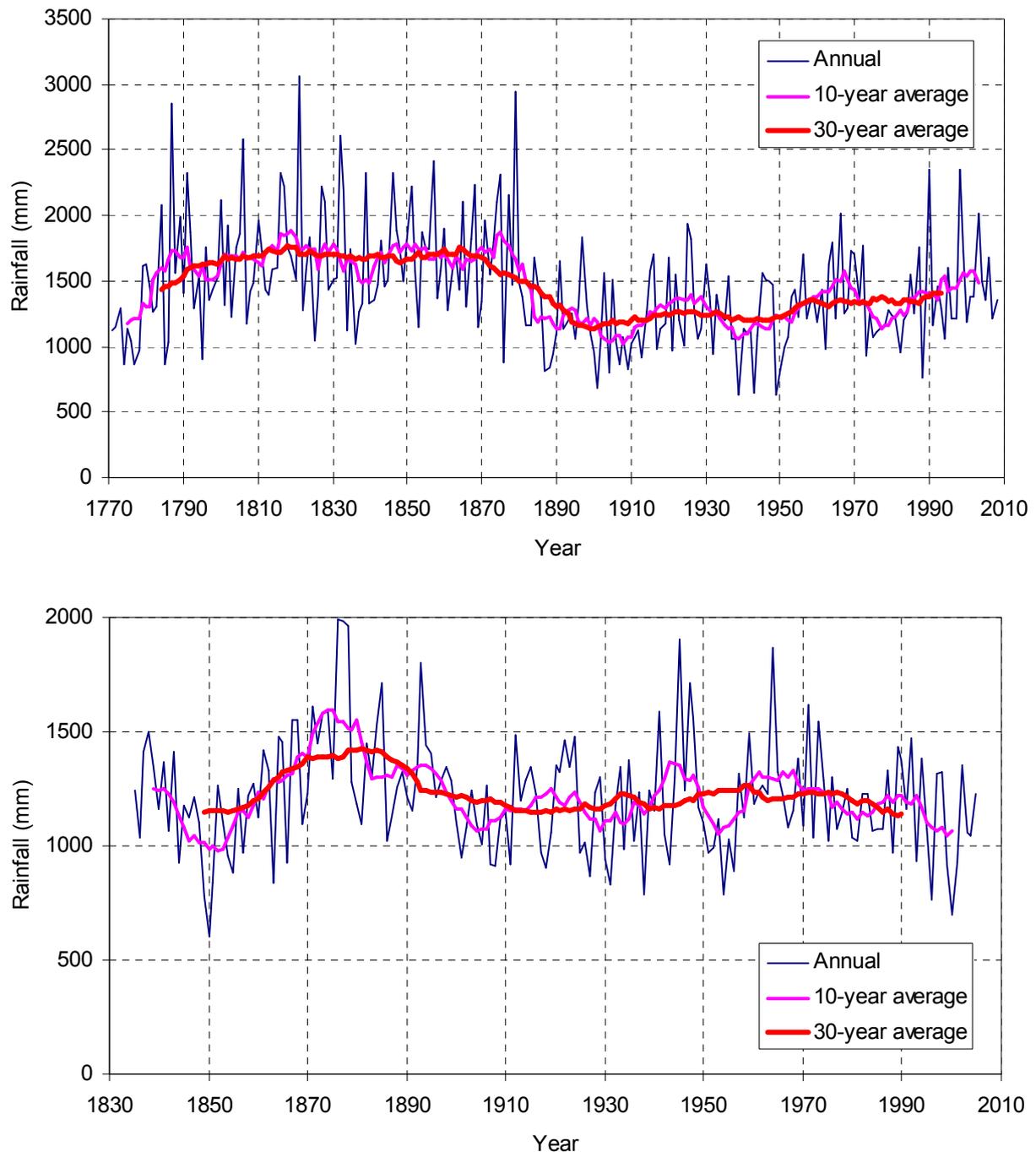


Fig. 7 Annual precipitation time series in two of the stations with the longest records worldwide: (upper) Seoul, Korea; (lower) Charleston City, USA. Data from the data base of the Dutch Royal Netherlands Meteorological Institute (KNMI; climexp.knmi.nl) and additional information as shown in text.

Finally, Fig. 8 depicts the time series of a storm measured at unusually high temporal resolution, that is 10 seconds. This storm, with duration 96 790 s or about 27 h starting at 1990-02-12T17:03:39, is one of several storms that were measured at the University of Iowa using devices that support high sampling rates (Georgakakos *et al.*, 1994). Fig. 8 also includes plots at 5-min and hourly time scales. The minimum intensity was virtually zero at all three scales, whereas the maximum rainfall intensity was 118.7, 38.9 and 18.1 mm/h at time scales of 10 s, 5 min and 1 h, respectively. As the mean intensity during the storm is 3.89 mm/h, these maximum values are 30, 10, and 4.6 times higher

than the mean. This example highlights the spectacular variability of rainfall, particularly at fine time scales (see also Uijlenhoet and Sempere-Torres, 2006). As the total rainfall amount of this storm event only slightly exceeds 100 mm, it could be thought of as a rather modest event. Storms with amounts much higher than this are often recorded even in semi-dry climates and, obviously, the variability of rainfall intensity during such storms is even higher.

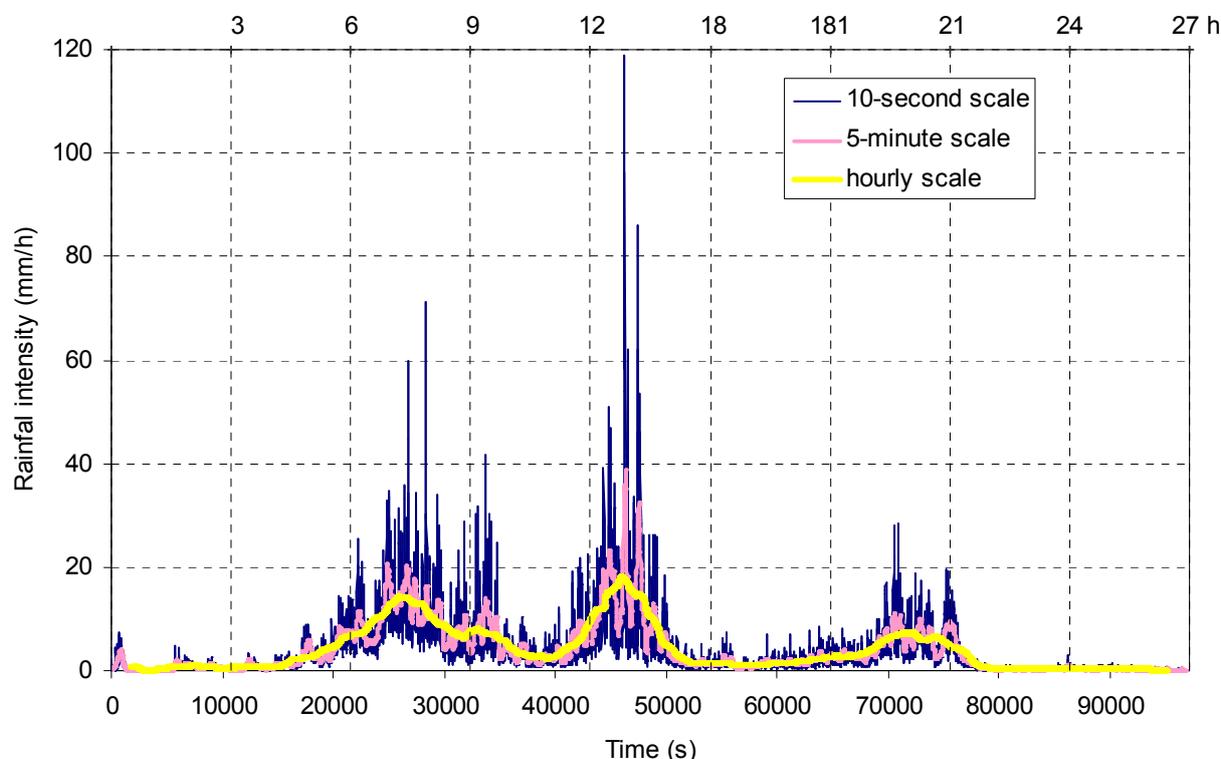


Fig. 8 Time series of a storm in Iowa, USA measured at the University of Iowa with temporal resolution of 10 seconds (Georgakakos *et al.*, 1994); time zero corresponds to 1990-02-12T17:03:39.

1.5 Probability and stochastic processes as tools for understanding and modelling precipitation

The high variability and the rough and irregular patterns in observed fields and time series are much more prominent in precipitation than in other meteorological variables such as atmospheric pressure or temperature. High variability implies high *uncertainty* and, unavoidably, this affects *predictability* in deterministic terms. Considering weather prediction as an example, it is well known that the forecasts of atmospheric pressure and temperature are much more reliable than those of precipitation. Numerical weather prediction (NWP) uses current weather conditions as input to mathematical models of the atmosphere, which solve the flow (Navier-Stokes) equations, the thermodynamic energy equation, the state equation of gases, and the equation of conservation of water vapour, over a grid covering the entire atmosphere. The processes related to cloud formation and precipitation (see section 2.2) are less accurately represented in these models. While the continuous improvement of NWP models resulted in a considerable reduction of forecast errors on pressure and temperature, the improvement in the so-called *quantitative precipitation forecast* (QPF) has been slower (Olson *et al.*, 1995). Further, although the advances in computing infrastructure permitted the increase in model resolution that leads generally to an improvement of precipitation forecasts, recently many authors have highlighted the limitations of such an approach

(e.g. Mass *et al.*, 2002; Lagouvardos *et al.*, 2003; Kotroni and Lagouvardos, 2004). The major advancement in QPF in the last decades was the abandonment of the pure deterministic approach, which seeks a unique prediction, and the adoption of a more *probabilistic* approach to precipitation forecast, based on earlier ideas of Epstein (1969) and Leith (1974). In this approach, known as *ensemble forecasting*, the same model produces many forecasts. To produce these forecasts, perturbations are introduced, e.g. in the initial conditions, and, because of the nonlinear dynamics with sensitive dependence on the initial conditions (e.g. Lorenz, 1963), these perturbations are magnified in time thus giving very different precipitation amounts in a lead time of one or more days. The different model outputs can then be treated in a probabilistic manner, thus assigning probabilities to rainfall occurrence as well as to the exceedence of a specified rainfall threshold. In this manner, although the model uses deterministic dynamics, the entire framework is of the *Monte Carlo* or *stochastic* type.

This method is satisfactory for a time horizon of forecast of a few days. In hydrology, this time horizon is relevant in real-time flood forecasting. However, in hydrological design, horizons as long as 50 or 100 years (the lifetimes of engineering constructions) are typically used. For such long horizons the use of deterministic dynamics and of the related laborious models would not be of any help. However, a probabilistic approach is still meaningful—in fact the only effective approach—and, in this case, it can be formulated irrespectively of the dynamics. Rather, the probabilistic approach should be based, in this case, on historical records of precipitation, such as those displayed in Fig. 7 and Fig. 8. The reasoning behind neglecting the deterministic dynamics is that, beyond a certain time horizon (which in precipitation is of the order of several days) even the simplest nonlinear systems tend to a *statistical equilibrium state*. In this state the probability distribution of the system properties, conditioned on the initial state, is practically equal to the marginal (i.e. unconditional) probability distribution of the same properties (Koutsoyiannis, 2009). This equilibrium, which is different from the typical thermodynamic equilibrium, corresponds to the maximization of the *entropy* of the vector of *random variables* defining the system state.

1.5.1 Basic concepts of probability

Probability is thus not only a mathematical tool to model precipitation uncertainty, but also a concept for understanding the behaviour of precipitation. Probabilistic thinking provides insights in to phenomena and their mathematical descriptions, which may not be achievable in deterministic terms. It should be recalled that, according to the Kolmogorov (1933) system, probability is a *normalized measure*, that is, a function P that maps sets (areas where unknown quantities lie) to real numbers (in the interval $[0, 1]$). Furthermore, a random variable \underline{x} is a single-valued function of the set of all elementary events (so that to each event it maps a real number) and is associated with a *probability distribution function*. The latter is defined as

$$F_{\underline{x}}(x) := P\{\underline{x} \leq x\} \quad (3)$$

where x is any real number, which should be distinguished from the random variable \underline{x} .^{*} $F_{\underline{x}}(x)$ is a non decreasing function of x with the obvious properties $F_{\underline{x}}(-\infty) = 0$ and $F_{\underline{x}}(+\infty) = 1$. For continuous

^{*} Distinction of random variables from their values is usually done by denoting them with upper case and lower case letters, respectively. This convention has several problems—e.g., the Latin x and the Greek χ , if put in

random variables (as is for instance the representation of a nonzero rainfall depth), the probability that a random variable \underline{x} would take any particular value x is $P\{\underline{x} = x\} = 0$. Thus, the question whether one particular value (say $x_1 = 10$ mm, assuming that \underline{x} denotes daily rainfall at a location) is more probable than another value (say $x_2 = 10$ m, which intuitively seems extremely improbable) cannot be answered in terms of the probability function P , as all particular values have probability equal to zero. The derivative of F , that is

$$f_{\underline{x}}(x) := dF_{\underline{x}}(x)/dx \quad (4)$$

termed *probability density function*, can provide this answer, as the quantity $f_{\underline{x}}(x) dx$ is the probability that rainfall will lie in an interval of length dx around x . Apparently then the ratio $f_{\underline{x}}(x_1) / f_{\underline{x}}(x_2)$ equals the ratio of the probabilities at points x_1 and x_2 .

These rather simple notions allow quantification of uncertainty and enable producing a different type of predictions, which offer a concrete foundation of rational decisions for the design and management of water resources projects. This quantification is sometimes (mostly in Bayesian statistics) referred to as *probabilization of uncertainty* that is meant to be the axiomatic reduction from the notion of unknown to the notion of a random variable (Robert, 2007).

1.5.2 Stochastic processes

In the study of rainfall variation in time, the notion and the theory of *stochastic processes* provides the necessary theoretical framework. A stochastic process is defined as an arbitrarily (usually infinitely) large family of random variables $\underline{x}(t)$ (Papoulis, 1991). In most hydrological applications time is discretized using an appropriate time step δ ; for integer i the average of the continuous time process $\underline{x}(t)$ from $t = (i - 1)\delta$ to $t = i\delta$, is usually denoted \underline{x}_i and forms a discrete time stochastic process. The index set of the stochastic process (i.e. the set from which the index t or i takes its values) can also be a vector space, rather than the real line or the set of integers. This is the case, for instance, when we assign a random variable (e.g. rainfall depth) to each geographical location (a two dimensional vector space) or to each location and time instance (a three-dimensional vector space). Stochastic processes with a multidimensional index set are also known as *random fields*.

A realization $x(t)$ (or x_i) of a stochastic process $\underline{x}(t)$ (or \underline{x}_i), which is a regular (numerical) function of the time t (or a numerical sequence in time i) is known as a *sample function*. Typically, a realization is observed at countable time instances (and not in continuous time, even if the process is of continuous-time type). This sequence of observations is also referred to as a *time series*. Clearly then, a time series is a sequence of numbers, whereas a stochastic process is a family of random variables.[†]

The distribution and the density functions of the random variable \underline{x}_i , that is

$$F_i(x) := P\{\underline{x}_i \leq x\}, f_i(x) := dF_i(x)/dx \quad (5)$$

upper case, are the same symbol X —other texts do not distinguish the two at all, thus creating another type of ambiguity. Here we follow a different convention, in which random variables are underscored and their values are not.

[†] Unfortunately, a large body of literature does not make this distinction and confuses stochastic processes with time series.

are called, respectively *first order distribution function* and *first order density function* of the process. Likewise, the *second order distribution function* is $F_{i_1 i_2}(x_1, x_2) = P\{\underline{x}_{i_1} \leq x_1, \underline{x}_{i_2} \leq x_2\}$ and this can be generalized to define the n^{th} *order distribution function*. It should be recalled that the expected value of a function g of one, two or more random variables is the integral of g multiplied by the density f , that is

$$E[g(\underline{x}_i)] := \int_{-\infty}^{\infty} g(x) f_i(x) dx, \quad E[g(\underline{x}_{i_1}, \underline{x}_{i_2})] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x_1, x_2) f_{i_1 i_2}(x_1, x_2) dx_1 dx_2 \quad (6)$$

The use of square brackets in $E[]$ and the random variables \underline{x}_i rather than their values x signifies the fact that the expected value is not a function of the real number x ; rather it depends solely on the distribution function associated with the random variable \underline{x}_i . Of particular interest are the cases where $g(\underline{x}_i) = \underline{x}_i$, whence $E[\underline{x}_i] =: \mu_i$ is the mean value of \underline{x}_i , and $g(\underline{x}_{i_1}, \underline{x}_{i_2}) = (\underline{x}_{i_1} - \mu_{i_1})(\underline{x}_{i_2} - \mu_{i_2})$, whence $E[(\underline{x}_{i_1} - \mu_{i_1})(\underline{x}_{i_2} - \mu_{i_2})] =: C_{i_1 i_2}$ is the process *autocovariance*, that is, the covariance of the random variables \underline{x}_{i_1} and \underline{x}_{i_2} . The process *variance* (the variance of the variable \underline{x}_j), is a special case of the latter, that is, $\text{Var}[\underline{x}_j] = C_{jj}$, whereas the standard deviation is the square root of the latter, that is, $\sigma_j := \sqrt{C_{jj}}$. Consequently, the process *autocorrelation* (the correlation coefficient of the random variables \underline{x}_{i_1} and \underline{x}_{i_2}) is $\rho_{i_1 i_2} := C_{i_1 i_2} / (\sigma_{i_1} \sigma_{i_2})$.

1.5.3 Stationarity

As implied by the above notation, in the general setting, the statistics of a stochastic process, such as the mean and autocovariance, depend on time i and thus vary with time. However, the case where these statistical properties remain constant in time is the most interesting. A process with this property is referred to as a *stationary* process. More precisely, a process is called (*strict-sense*) *stationary* if all its statistical properties are invariant to a shift of time origin. That is, the distribution function of any order of \underline{x}_{i+j} is identical to that of \underline{x}_i . A process is called *wide-sense stationary* if its mean is constant and its autocovariance depends only on time differences (lags), that is

$$E[X_i] = \mu, \quad E[(X_{i+j} - \mu)(X_i - \mu)] = C_j \quad (7)$$

Evidently, the standard deviation is constant too, i.e., $\sigma_i = \sigma$, and the autocorrelation is a function of the time lag only, i.e., $\rho_{i+j, i} = \rho_j$. A strict-sense stationary process is also wide-sense stationary but the reverse is not true.

A process that is not stationary is called *nonstationary*. In a nonstationary process one or more statistical properties depend on time. A typical case of a nonstationary process is the cumulative rainfall depth whose mean obviously increases with time. For instance, let us assume that the instantaneous rainfall intensity $i(t)$ at a geographical location and period of the year is a stationary process, with a mean μ . Let us further denote by $h(t)$ the rainfall depth collected in a large container (a cumulative raingauge) at time t and assume that at the time origin, $t = 0$, the container is empty. Clearly $E[h(t)] = \mu t$. Thus $h(t)$ is a nonstationary process.

It should be stressed that stationarity and nonstationarity are properties of a stochastic process, not of a sample function or time series. There is some confusion in the literature about this, as there are

several studies that refer to a time series as stationary or nonstationary. As a general rule, to characterise a process nonstationary, it suffices to show that some statistical property is a *deterministic* function of time (as in the above example of the cumulative rainfall), but this cannot be directly inferred merely from a time series. To understand this, let us consider the time series of annual rainfall in Seoul plotted in the upper panel of Fig. 7. Misled by the changing regime of precipitation at climatic scale, as manifested in the plot of the 30-year average, it would be tempting to note (a) an increasing trend in the period 1770-90; (b) a constant climate with high precipitation during the period 1790-1870; (c) a decreasing trend between 1870-1900; and (d) a constant climate with low precipitation thereafter. It is then a matter of applying a fitting algorithm to determine, say, a broken-line type of function to the time series, which would be called a “deterministic function of time”. The conclusion would then be that the time series is nonstationary. However, this is a wrong *ex post* argument, which interprets the long-term variability of the processes as a deterministic function. Had the function been indeed deterministic, it would also apply to future times, which obviously is not the case. Comparison with the previous example (cumulative rainfall), where the deterministic function $E[x(t)] = \mu t$ was obtained by theoretical reasoning (deduction) rather than by inspection of the data, demonstrates the real basis of nonstationarity. Koutsoyiannis (2006b) has provided a more detailed study of this issue.

Stochastic processes describing periodic phenomena, such as those affected by the annual cycle of the Earth, are clearly nonstationary. For instance, the daily rainfall at a mid-latitude location could not be regarded as a stationary process. Rather, a special type of a nonstationary process, whose properties depend on time in a periodical manner (are periodic functions of time) should be used. Such processes are called *cyclostationary* processes.

1.5.4 Ergodicity

The concept of *ergodicity* (from the Greek words *ergon* – work – and *odos* – path) is central to the problem of determining the distribution function of a process from a single sample function (time series). A stationary stochastic process is ergodic if any statistical property can be determined from a sample function. Given that, in practice, the statistical properties are determined as time averages of time series, the above statement can be formulated alternatively - a stationary stochastic process is ergodic if time averages equal ensemble averages (i.e. expected values). For example, a stationary stochastic process is *mean ergodic* if

$$E[x_j] := \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N x_i \quad (8)$$

The left-hand side in the above equation represents the ensemble average whereas the right-hand side represents the time average, for the limiting case of infinite time. While the left-hand side is a parameter, rather than a random variable, the right-hand side is a random variable (as a sum of random variables). Equating a parameter with a random variable implies that the random variable has zero variance. This is precisely the condition that makes a process ergodic, a condition that does not hold true for every stochastic process.

1.5.5 Some characteristic stochastic properties of precipitation

It has been widely thought that rainfall exhibits some autocorrelation (or time dependence) if the time scale of study is daily or sub-daily, but this dependence vanishes at larger time scales, such as monthly or yearly. Thus, for time scales monthly and above, rainfall data series have been traditionally treated as independent samples. Mathematically, such a perception corresponds to a Markovian dependence at fine time scales, in which the autocorrelation decreases rapidly with time lag in an exponential manner, that is

$$\rho_j = \rho^j \quad (9)$$

where $\rho := \rho_1$. Then for a large lag j , or for a large scale of aggregation and even for the smallest lag (one), the autocorrelation is virtually zero (e.g. Koutsoyiannis, 2002). If \underline{x}_j denotes the stochastic process at an initial time scale, which is designated as scale 1, then the averaged process at an aggregated time scale $k = 2, 3, \dots$, is

$$\underline{x}_j^{(k)} := \frac{\underline{x}_{(j-1)k+1} + \dots + \underline{x}_{jk}}{k} \quad (10)$$

(with $\underline{x}_i^{(1)} \equiv \underline{x}_i$). Let $\sigma^{(k)}$ be the standard deviation at scale k . In processes \underline{x}_j independent of time, $\sigma^{(k)}$ decreases with scale according to the well known classical statistical law of inverse square-root, that is

$$\sigma^{(k)} = \frac{\sigma}{\sqrt{k}} \quad (11)$$

However, this law hardly holds in geophysical time series including rainfall time series, whatever the scale is. This can be verified based on the examples presented in section 1.4. A more plausible law is expressed by the elementary scaling (power-law) property

$$\sigma^{(k)} = \frac{\sigma}{k^{1-H}} \quad (12)$$

where H is the so called Hurst exponent, after Hurst (1951) who first studied this type of behaviour in geophysical time series. Earlier, Kolmogorov (1940), when studying turbulence, had proposed a mathematical model to describe this behaviour. The behaviour has been known by several names including the *Hurst phenomenon*, *long-term persistence* and *long range dependence*, and a simple stochastic model that reproduces it is known as a *simple scaling stochastic model* or *fractional Gaussian noise* (due to Mandelbrot and van Ness, 1968). Here the behaviour is referred to as the *Hurst-Kolmogorov (HK) behaviour* or *HK (stochastic) dynamics* and the model as the *HK model*.

This behaviour implies that the autocorrelation decreases slowly, i.e., according to a power-type function, with lag j :

$$\rho_j^{(k)} = \rho_j = (1/2) [(|j+1|)^{2H} + (|j-1|)^{2H}] - |j|^{2H} \approx H(2H-1)j^{2H-2} \quad (13)$$

so that independence virtually never holds, unless $H = 0.5$, a value which reinstates classical statistics including the law in equation (11). Most often, natural processes including rainfall are positively correlated and H varies in the range (0.5, 1).

The above framework is rather simple and allows easy exploration of data to detect whether they indicate consistence with classical statistics or with the HK behaviour. A simple exploration tool is a double logarithmic plot of the estimates of standard deviation $\sigma^{(k)}$ vs. scale k , which is known as a *climacogram*.[‡] In such a plot, the classical law and the HK law are manifested by a linear arrangement of points with slopes -0.5 and $H - 1$, respectively. We must bare in mind, however, that a consequence of the HK law in equation (12) is that the classical estimator of the variance

$$\underline{s}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (14)$$

where n is the sample size, $\bar{x} \equiv \bar{x}_1^{(n)}$ is the estimator of the mean, and \underline{s} the estimator of standard deviation, implies negative bias if there is temporal dependence. The bias becomes very high for HK processes with H approaching 1. Apparently then, \underline{s} could be a highly biased estimator of σ ; an approximately unbiased estimator is (Koutsoyiannis, 2003a; Koutsoyiannis and Montanari, 2007):

$$\underline{s} \approx \sqrt{\frac{n'}{n'-1}} \underline{s} \quad (15)$$

where n' is the “equivalent” (or “effective”) sample size, i.e., the sample size that in the framework of classical statistics would lead to the same uncertainty (in the estimation of μ by $\bar{x} \equiv \bar{x}_1^{(n)}$) as yields an HK series with sample size n . For an HK process, n' is related to n by

$$n' = n^{2(1-H)} \quad (16)$$

It can be seen that n' can be very small even for high n if H is high, and thus the correcting factor $\sqrt{n'/(n'-1)}$ in equation (15) can be very large (see Koutsoyiannis and Montanari, 2007).

Returning to the time series of globally averaged monthly precipitation in the 30-year period 1979-2008, which has been discussed earlier and is displayed in Fig. 6, we may now study its statistical properties for several time scales. As the precipitation amounts are averaged over the entire globe, the effect of seasonality is diminished and the time series can be modelled by a stationary process rather than a cyclostationary one. Fig. 9 depicts the climacogram, that is, a logarithmic plot of standard deviation vs. scale. Empirical estimates of standard deviations have been calculated using both the classical estimator in equation (14) and the HK estimator in equation (15). Theoretical curves resulting from the classical statistical model (assuming independence), the Markovian model and the HK model have also been plotted. For the Markovian model, the lag one autocorrelation coefficient, estimated from the monthly data, is $\rho = 0.256$ and for the HK model the estimate of the Hurst coefficient is $H = 0.70$. This can be obtained readily from the slope of the straight line fitted to

[‡] Climacogram < Greek Κλιμακόγραμμα < [climax (κλιμαξ) = scale] + [gamma (γράμμα) = written].

the group of empirical points in Fig. 9, which should be $H - 1$. Here a slightly modified algorithm from Koutsoyiannis (2003a) has been used for the estimation of H . Overall, Fig. 9 clearly demonstrates that the empirical points are inconsistent with the classical and Markovian models and justify an assumption of HK behaviour.

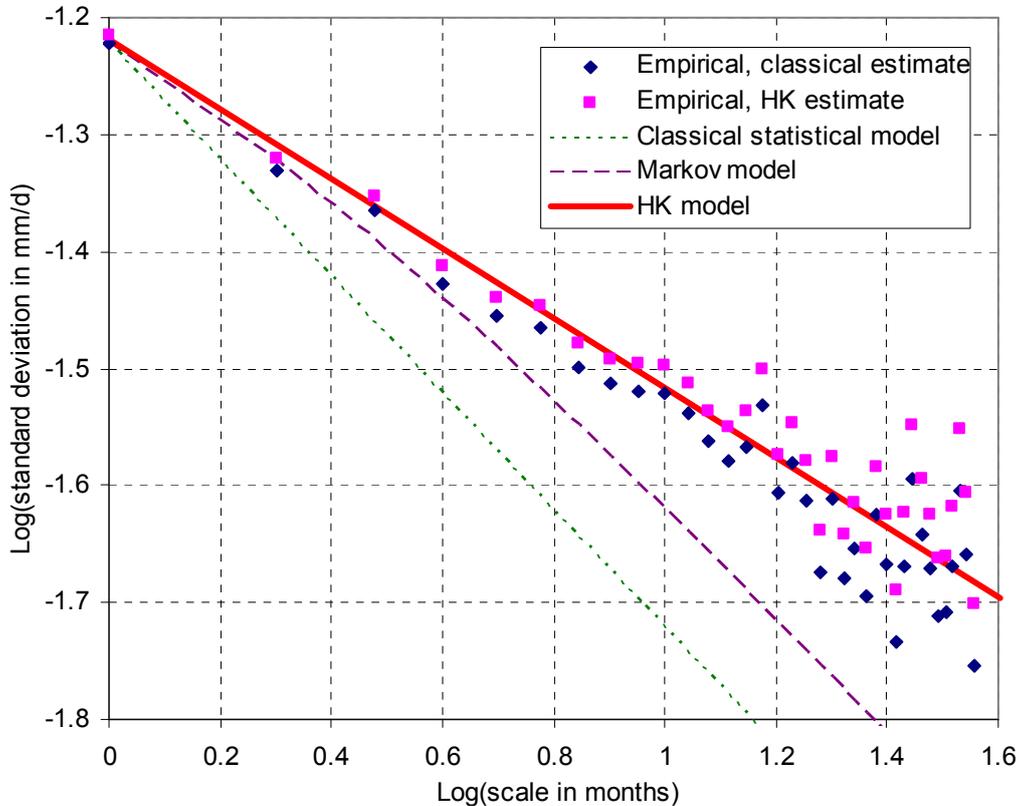


Fig. 9 Climacogram of the time series of globally averaged monthly precipitation in the 30-year period 1979-2008 shown in Fig. 6. The estimate of the Hurst coefficient for the HK model is $H = 0.70$.

Similar plots have been constructed, and are shown in Fig. 10, for the annual precipitation time series from Seoul, Korea and Charleston City, USA, displayed in Fig. 7. Again the empirical evidence from data precludes the applicability of the classical statistical model and favours the HK statistics. An additional plot for the ten-second precipitation time series in Iowa, USA, displayed in Fig. 8, is depicted in Fig. 11. Here the Hurst coefficient is very high, $H = 0.96$. The difference between the empirical points based on classical statistics on the one hand and the HK statistics on the other hand is quite impressive. Apparently, the classical model is completely inappropriate for the rainfall process.

The HK stochastic processes can be readily extended in a two-dimensional (2D) setting (or even multidimensional). The 2D version of equation (12) is

$$\sigma^{(k)} = \frac{\sigma}{k^{2-2H}} \quad (17)$$

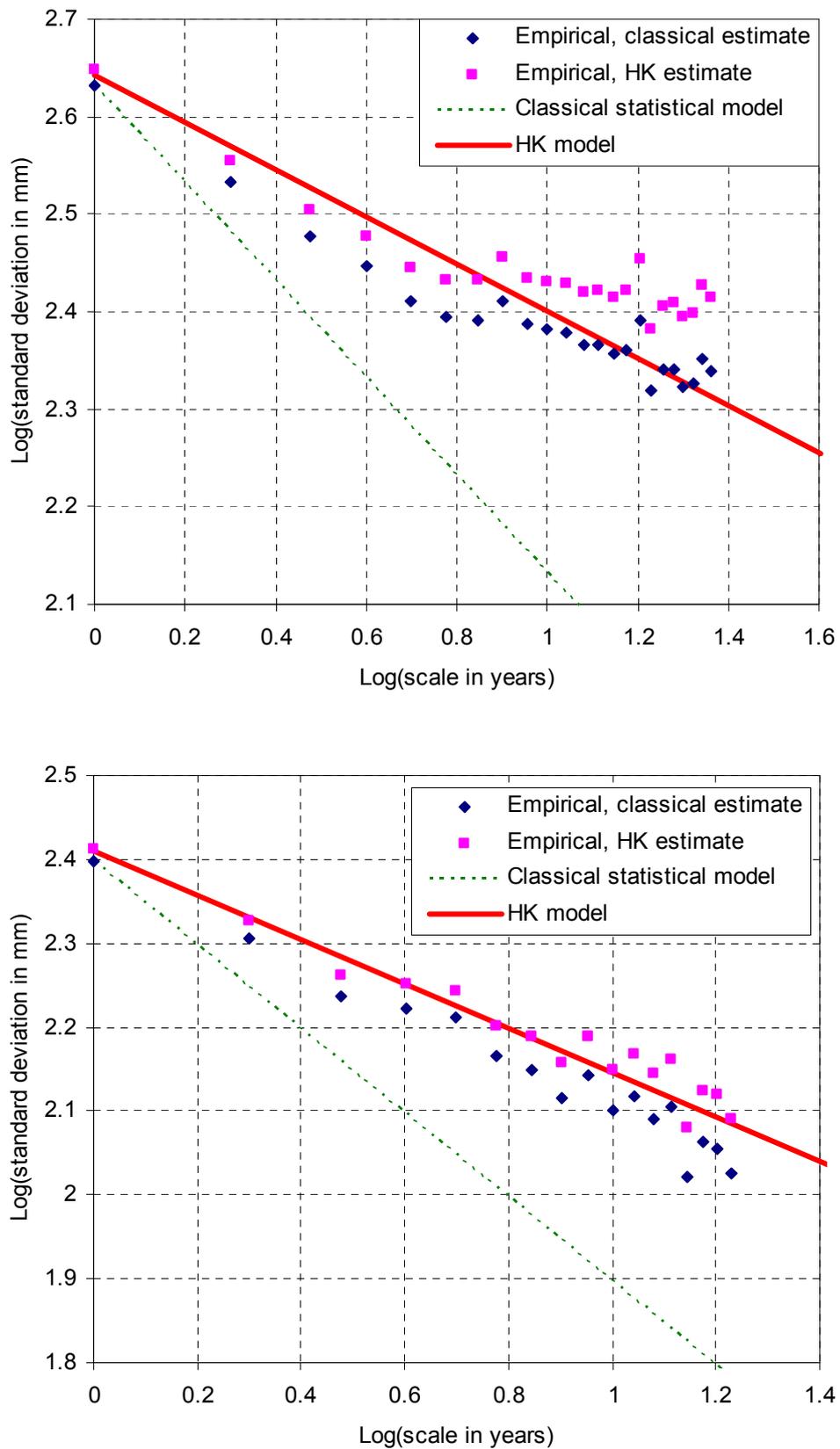


Fig. 10 Climacogram of the annual precipitation time series at: (upper) Seoul, Korea and (lower) Charleston City, USA, which are shown in Fig. 7; the estimated Hurst coefficients are 0.76 and 0.74, respectively.

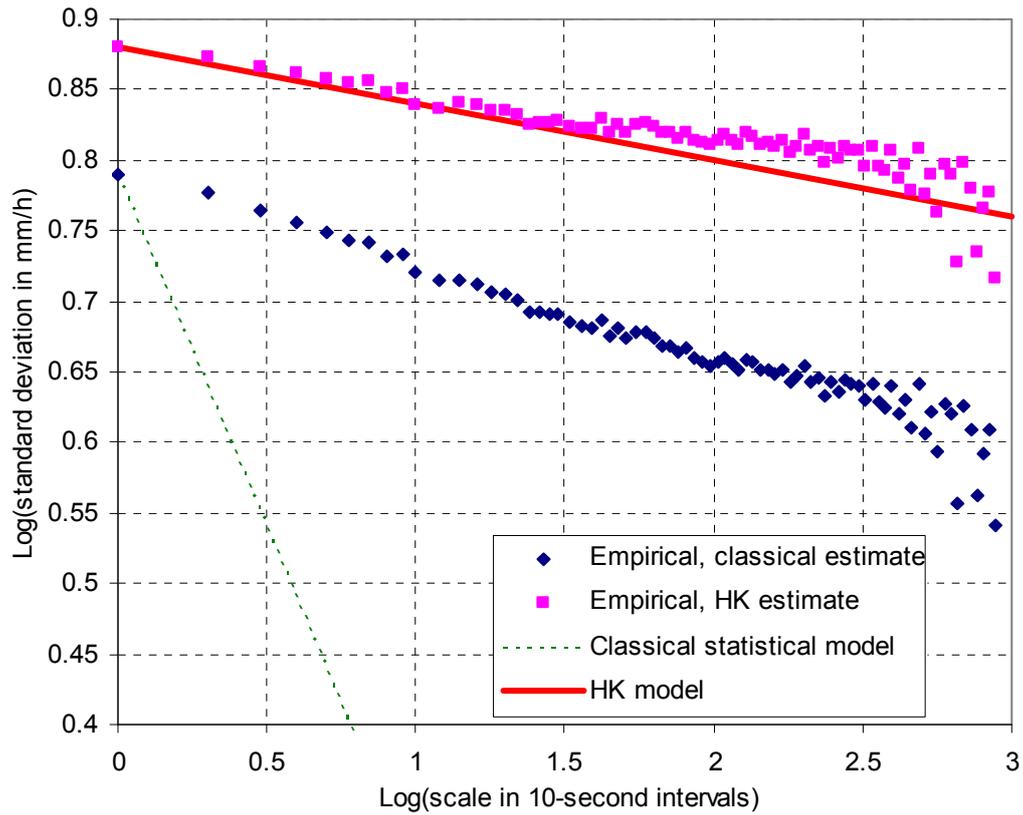


Fig. 11 Climacogram of the ten-second precipitation time series in Iowa, USA displayed in Fig. 8; the estimated Hurst coefficient is 0.96.

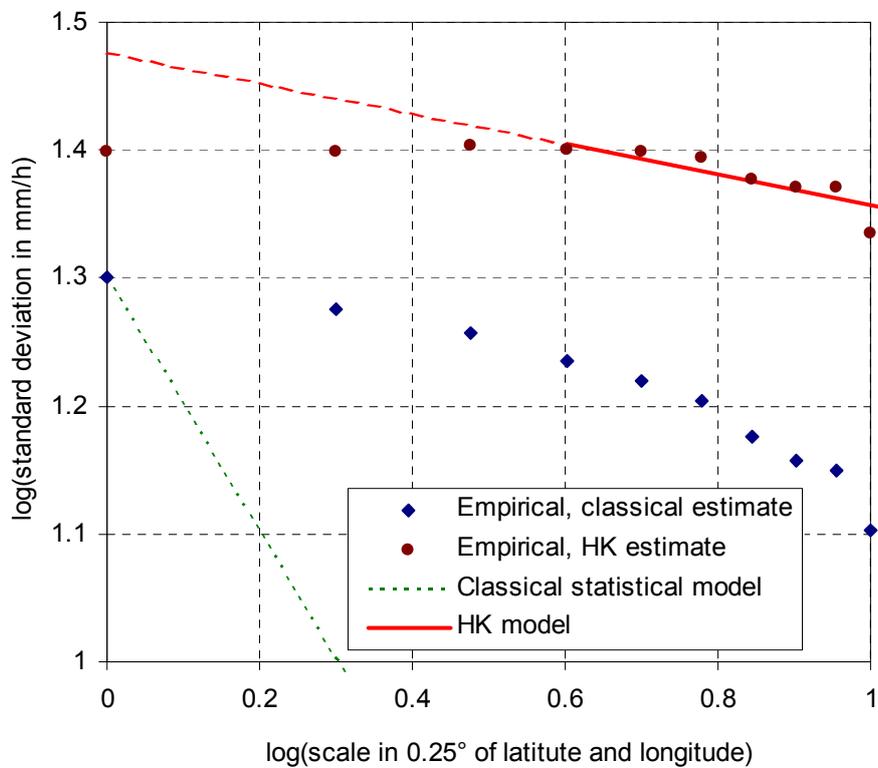


Fig. 12 Climacogram of the spatial daily rainfall over the area 9°N-5°S and 78-92°E (Indian ocean south-east of Sri Lanka) on 9 January 2006, as shown in the lower-left panel of Fig. 5; the estimated Hurst coefficient is 0.94.

This can be obtained by substituting k^2 for k in equation (12). Equations (15) and (16) still hold provided that n is the number of points, which is inversely proportional to k^2 . Fig. 12 demonstrates this behaviour by means of a climacogram for the spatial daily rainfall over the area 9°N-5°S and 78-92°E (Indian ocean south-east of Sri Lanka) on 9 January 2006, displayed in the lower-left panel of Fig. 5. Here the estimated Hurst coefficient is again very high, $H = 0.94$. As in all previous cases the classical model is completely inappropriate, while the HK model seems reasonable for scales ≥ 4 , which correspond to a resolution of $1^\circ \times 1^\circ$ and beyond.

Thus, the evidence presented using several examples of different spatial and temporal scales indicates that the HK dynamics is consistent with the nature of rainfall. This dynamics appear as a scaling behaviour, either in time or in space, which is either full, applicable to the entire range of scales, or asymptotic, applicable to large scales. Both these scaling behaviours are manifested as power laws of the standard deviation vs. temporal or spatial scale and of the autocorrelation vs. lag. There exists another type of scaling behaviour in precipitation, the scaling in state, which is sometimes confused with the other two scaling behaviours, but is fundamentally different. Scaling in state is a property of the marginal distribution function of rainfall (it has no relation to the dependence structure of the process unlike other types of scaling) and is expressed by power laws of the tails of (a) the probability density function $f(x)$, (b) the survival function (or exceedence probability) $F^*(x) := P\{\underline{x} > x\} = 1 - F(x)$ and (c) the return period $T = \delta / F^*(x)$ where δ is the length of the time scale examined. These scaling properties are expressed as

$$x \propto T^\kappa, \quad F^*(x) \propto x^{-1/\kappa}, \quad f(x) \propto x^{-1-1/\kappa} \quad (18)$$

and are equivalent to each other. All these are asymptotic, i.e., they hold only for large values of x or, in other words for the distribution tails. Such tails are known by several names, such as long, heavy, strong, power-type, overexponential, algebraic, or Pareto tails. The latter name comes from the Pareto distribution, which in its simplest form is given in equation (18), although its generalized form is applicable to rainfall (see section 5.2). As this is an asymptotic behaviour, long records are needed to observe it. Fig. 13 shows a logarithmic plot of the empirical distribution (expressed in terms of return period T) of a large data set of daily rainfall. This data set was formed from records of 168 stations worldwide, each of which contained 100 years of data or more (Koutsoyiannis, 2004b). For each station with n years of record, n annual maximum values of daily rainfall were extracted. These values were standardized by their mean and merged in one sample of length 17 922 station-years. From the theoretical distributions, also plotted in the figure, it is observed that the Pareto distribution (whose right tail appears as a straight line in the logarithmic plot; see section 5.2) with $\kappa = 0.15$ provides the best fit, thus confirming the applicability of asymptotic scaling in state and the inappropriateness of the exponential-type tail. This has severe consequences particularly in hydrological design, as distributions with exponential tails have been most common in hydrological practice, whereas it is apparent that the power-type tails are more consistent with reality. As shown in Fig. 13 the difference between the two types can be substantial.

Koutsoyiannis (2005a,b) produced the foresaid different types of scaling from the principle of maximum entropy. As entropy is a measure of uncertainty, the applicability of the principle of maximum entropy and its consistence with observed natural behaviours characterizing the precipitation process underscores the dominance of uncertainty in precipitation.

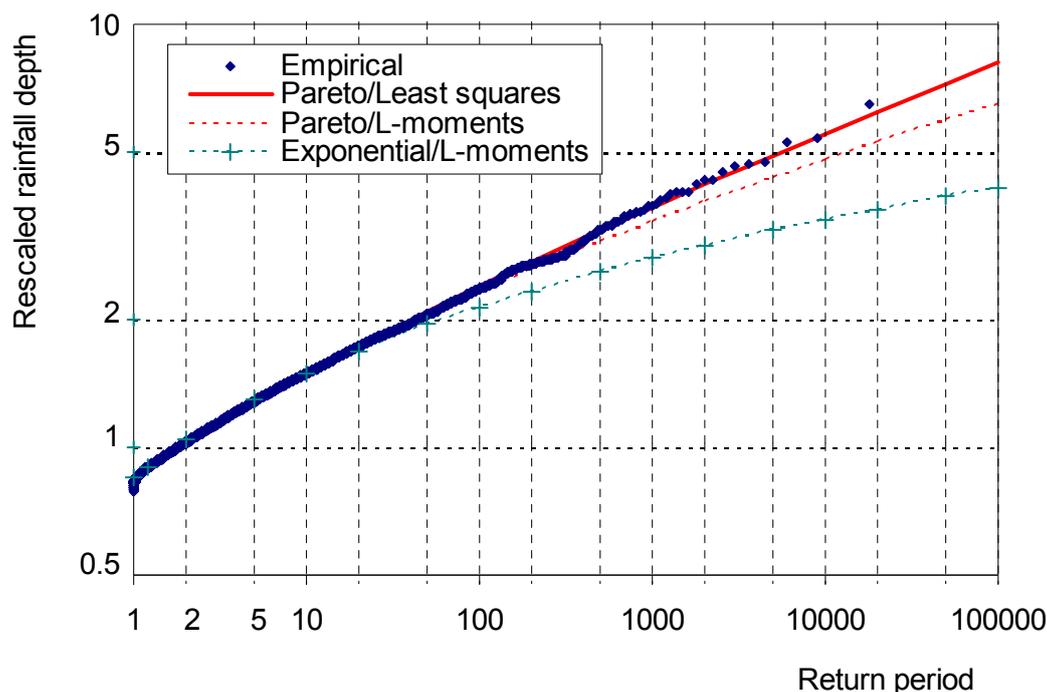


Fig. 13 Logarithmic plot of rescaled daily rainfall depth vs. return period: empirical estimates from a unified sample over threshold, formed using rainfall data from 168 stations worldwide (17922 station-years). The unified sample was rescaled by the mean of each station, and fitted using a Pareto and an exponential distribution models (adapted from Koutsoyiannis, 2004b).

2 Physical and meteorological framework

Atmospheric air is a heterogeneous mixture of gases, also containing suspended particles in liquid and solid phase. The most abundant gases are the Nitrogen (N_2) and Oxygen (O_2) that account for about 78% and 21%, respectively, by volume of the atmospheric permanent gases, followed by Argon (Ar) and traces of other noble gases. Their concentrations are almost constant worldwide and up to an altitude of about 90 km. Water vapour (H_2O) appears in relatively low concentrations, which are highly variable. However, water vapour is very important for the energy exchange on earth (it accounts for 65% of the radiative transfer of energy in the atmosphere; Hemond and Fechner-Levy, 2000), as well as the mass transfer processes in the hydrological cycle. Under certain conditions (i.e. pressure and temperature) water vapour can transform to droplets or ice crystals with subsequent release of latent heat (see sections 2.1 and 2.2). More generally, the water vapour content of atmospheric air affects its density and it is of central importance in atmospheric thermodynamics (section 2.1). The varying content and high importance of water vapour in precipitation processes and thermodynamics, has led to the study of atmospheric air as a mixture of two (ideal) gases: *dry air* and *water vapour*. This mixture is usually referred to as *moist air* and has thermodynamic properties determined by its constituents (e.g. Rogers and Yau, 1996; Cotton and Anthes, 1989).

The particles of solid and liquid material suspended in air are called *aerosols*. Common examples of aerosols are water droplets and ice crystals (called *hydrometeors*), smoke, sea salt (NaCl), dust and pollen. The size distribution of solid aerosols depends strongly on their location. For example, the

size spectrum of aerosols over land is narrow with high concentrations of small particles (e.g. kaolinite, dust, pollen etc.), whereas the size spectrum of aerosols over sea is wider with small concentrations of larger particles (e.g. sea salt; e.g. Ryan *et al.*, 1972). Existence of aerosols in the atmosphere is of major importance, since a select group of aerosols called *hygroscopic nuclei* is crucial for the nucleation of liquid water and initiation of rain (e.g. Brock, 1972, and section 2.2 below).

When moist air is cooled (i.e. below its dew point; see section 2.1), an amount of water vapour condenses and a *cloud* forms, but precipitation may or may not occur. Initiation of rain requires the formation of hydrometeors (i.e. water droplets and ice crystals) of precipitable size (e.g., Gunn and Kinzer, 1949; Twomey, 1964, 1966; Brock, 1972). Formation and growth of these particles are governed by processes that take place at scales comparable to their size (μm - mm). The latter processes form the core of *cloud microphysics*, whereas large scale processes related to thermodynamics of moist air and motion of air masses form the core of *cloud dynamics*. Importantly, precipitation is the combined effect of both large- and micro-scale processes, and both processes are equally important and necessary for precipitation to occur.

2.1 Basics of moist air thermodynamics

In a parcel of moist air at temperature T with volume V and mass $M = M_d + M_v$, with the two components denoting mass of dry air and water vapour, respectively, the density is $\rho = M/V$ and the concentration of water vapour, known as *specific humidity*, is $q := M_v/M$. The quantity $r := M_v/M_d = q/(1-q)$ is usually referred to as the *mixing ratio*. The total pressure of the moist air in the parcel, p (the atmospheric pressure), equals the sum of the partial pressure of dry air p_d and that of water vapour e (i.e. $p = p_d + e$). Specific humidity and vapour pressure are interrelated through

$$q = \frac{\varepsilon e}{p - (1 - \varepsilon) e} \quad (19)$$

where $\varepsilon = 0.622$ is the ratio of molar masses of water vapour and dry air. Air cannot hold an arbitrarily high quantity of vapour. Rather, there is an upper limit of the vapour pressure e^* , called the *saturation vapour pressure*, which depends on the temperature T and is given by the Clausius-Clapeyron equation. A useful approximation to this equation is:

$$e^*(T) = 6.11 \exp\left(\frac{17.67T}{T + 237.3}\right) \quad (20)$$

where e^* is in hPa and T is in $^{\circ}\text{C}$. Consequently, from equations (19) and (20) we can calculate the saturation specific humidity q^* , which is a function of T , and expresses the water vapour holding capacity of air. As shown in Fig. 14, this capacity changes drastically, almost exponentially, with temperature, so that a change of temperature from -40 to 40°C increases this capacity by 2.5 orders of magnitude.

The ratio of the actual to saturation vapour pressure, i.e., $e/e^* =: U$, called the *relative humidity*, is normally smaller than 1. When an air parcel cools, while e remains constant, e^* decreases and hence U increases, up to the saturation value 1 or 100%. The temperature T_d at which saturation occurs is called the *dew point* temperature and is calculated from equation (20) by setting $e^*(T_d) = e$. Therefore, cooling of the air parcel below the dew point temperature will result in condensation, or

transformation of the excess water vapour into liquid water in the form of droplets. During this change of phase the relative humidity remains 100%. Condensation releases heat at a fairly constant rate ($L \approx 2.5 \text{ MJ/kg}$); this rate equals that of evaporation of water at a constant temperature and is thus called *latent heat*.

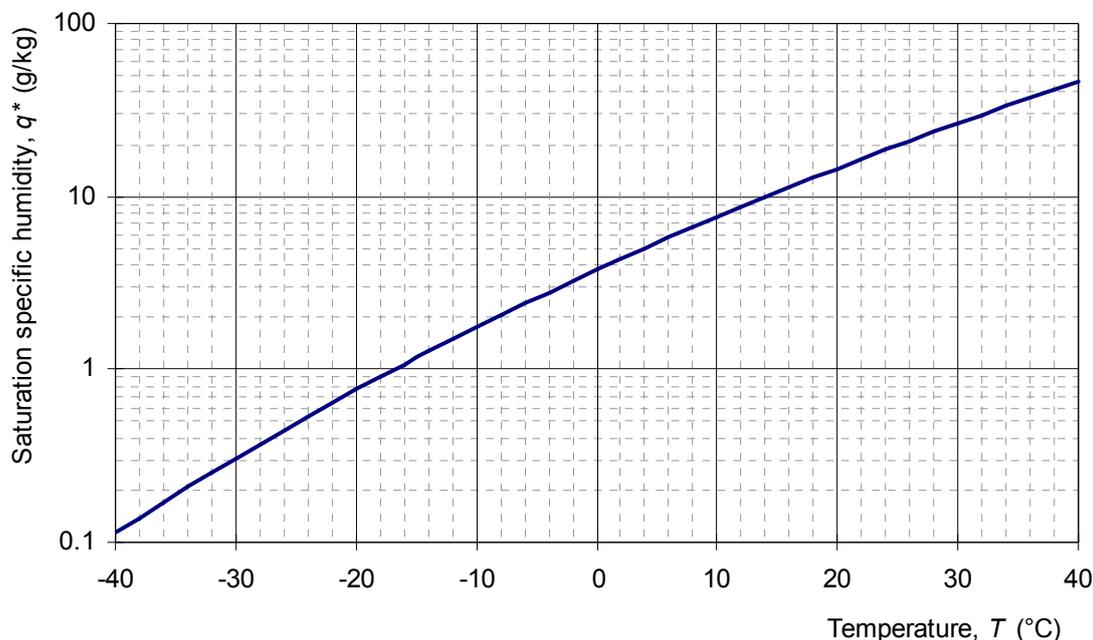


Fig. 14. Saturation specific humidity as a function of air temperature.

For an air parcel to ascend and expand spontaneously, so that condensation and cloud formation can occur, the ambient (atmospheric) temperature gradient $\gamma := -dT/dz$, where z denotes altitude, also known as *lapse rate*, must be high (otherwise an uplifted air parcel will sink again). While the parcel ascends and expands adiabatically (i.e. in a way that no heat transfer takes place between the air parcel and its ambient air), its own lapse rate is $\gamma_d = 9.8^\circ\text{C/km}$ if the expansion is dry adiabatic (i.e. if it takes place without condensation of water vapor) and somewhat smaller, γ^* , if the expansion is moist adiabatic (i.e., if the temperature has fallen below dew point, so that some of the water vapour in the parcel condenses to liquid form). The gradient γ^* is not constant but varies with temperature T and air pressure p so that $\gamma^* = 4^\circ\text{C/km}$ for $T = 25^\circ\text{C}$ and $p = 1000 \text{ hPa}$, whereas $\gamma^* = 9^\circ\text{C/km}$ for $T = -25^\circ\text{C}$ and $p = 1000 \text{ hPa}$; an average value is $\gamma^* = 6.5^\circ\text{C/km}$ (Koutsoyiannis, 2000a; see also Wallace and Hobbs, 1977). When the ambient lapse rate γ is smaller than γ^* the atmosphere is stable, and no spontaneous lift occurs and no clouds are formed. When $\gamma > \gamma_d$ the atmosphere is unstable and favours air lift and formation of clouds. The case $\gamma^* < \gamma < \gamma_d$ is known as *conditional instability* and it serves as an important mechanism for mesoscale precipitation processes (see sections 2.4 and 2.5).

2.2 Formation and growth of precipitation particles

The Clausius-Clapeyron equation describes the equilibrium condition of a thermodynamic system consisting of bulk water and vapour. A state out of the equilibrium, in which $e > e^*$ ($U > 1$) is possible, but thermodynamically unstable, and is called *supersaturation*. Detailed study of the transition of water vapour to liquid or ice at or above saturation is associated with certain *free energy barriers*. An

example of such an energy barrier is the dynamic energy associated with the surface tension, σ , of a water droplet. For a spherical droplet, σ is proportional to the pressure of water within the droplet p and inversely proportional to its radius r (i.e. $\sigma = p/2r$). This means that a high vapour pressure is needed for a very small droplet to be maintained and not evaporate. In essence, the free energy barrier of surface tension makes droplet formation solely by condensation of water vapour (a process usually referred to as *homogeneous nucleation*) almost impossible in nature. However, if the surface tension barrier is bypassed, common supersaturations of the order of 1-2% (i.e. $U = 1.01-1.02$) are sufficient for water vapour to diffuse toward the surface of the droplet. The rate of *diffusional growth* is proportional to the supersaturation $U - 1$ of the ambient air, and inversely proportional to the radius r of the droplet, i.e., $dr/dt \propto (U - 1)/r$ (Mason, 1971; Rogers and Yau, 1996).

While homogeneous nucleation requires large supersaturations, formation of droplets is drastically facilitated by particulated matter of the size of micrometres or lower, the *aerosols*, some of which, called *condensation nuclei*, are *hydrophilic* and serve as centres for droplet condensation (Brock, 1972; Slinn, 1975; Hobbs *et al.*, 1985). This process is usually referred to as *heterogeneous nucleation* and it is almost exclusively the process that governs water vapour condensation in the atmosphere (Houze, 1993).

When the temperature in the cloud drops below the freezing point, water droplets are said to be *supercooled*, and they may or may not freeze. For pure water droplets, *homogeneous freezing* does not occur until the temperature drops below -40°C (Rogers and Yau, 1996). However, the presence of certain condensation nuclei, called *ice nuclei*, may allow freezing of water droplets at temperatures a few degrees below 0°C . These nuclei are particles of the size of micrometres or lower that form strong bonds with water and closely match the crystalline structure of ice. Different particles serve as condensation nuclei at different subfreezing temperatures. For example, silver iodide (AgI) serves as an ice nucleator at -4°C and kaolinite at -9°C (e.g. Houghton, 1985).

Evidently, a cloud is an assembly of tiny droplets with usually met concentrations of several hundred per cubic centimetre, and radii of several micrometres. This structure is very stable and the only dominant process is vapour diffusion, which accounts for the size growth evolution of the whole droplet population (Telford and Chai, 1980; Telford and Wagner, 1981). Precipitation develops when the cloud population becomes unstable and some droplets grow faster relative to others.

In general, two main mechanisms account for the cloud microstructure to become unstable. The first mechanism is the *collision* and *coalescence* (i.e. sticking) of larger (and faster moving) collector drops with smaller (and slower moving) collected droplets. This mechanism is particularly important for precipitation development in *warm clouds* (i.e. at temperatures in excess of 0°C ; see e.g. Houze, 1993) and, for a long time, it has formed an active research area in cloud and precipitation physics (e.g. Langmuir, 1948; Bowen, 1950; Telford, 1955; Scott, 1968, 1972; Long, 1971; Drake, 1972a,b; Gillespie, 1972, 1975; Robertson, 1974; Berry and Reinhardt, 1974a,b; Vohl *et al.*, 1999; Pinsky *et al.*, 1999, 2000; Pinsky and Khain, 2004; and review in Testik and Barros, 2007). Its significance for precipitation processes depends considerably on the droplet size spectra, with larger effectiveness for wider spectra with small concentrations of larger particles (Berry and Reinhardt, 1974a,b).

The second mechanism is related to interaction between water droplets and ice crystals, and is limited to clouds with tops that extend to subfreezing temperatures (i.e. *cold clouds*). In particular,

when an ice crystal develops in the presence of a large number of supercooled droplets, the situation becomes immediately unstable and the ice crystal grows due to diffusion of water vapour from the droplets toward the crystal. This is due to the fact that the equilibrium vapour pressure over ice is less than that over water at the same subfreezing temperature. Thus, the ice crystal grows by diffusion of water vapour and the supercooled droplets evaporate to compensate for this. The transfer rate of water vapour depends on the difference between the equilibrium vapour pressure of water and ice, a quantity that becomes sufficiently large at about -15°C (Uijlenhoet, 2008). The latter process is called the *Bergeron-Findeisen mechanism* after the scientists who first studied it (Bergeron, 1935; Findeisen, 1938).

Once the ice crystals have grown by vapour diffusion to sizes sufficiently large for gravitational settling to dominate, they start falling and colliding with their ambient droplets and ice crystals, a process usually referred to as *accretional growth*. In the first case (i.e. when ice crystals collide with droplets) graupel or hail may form, whereas in the second case snowflakes are likely to form.

As the frozen particles fall, it is possible to enter layers with temperature higher than 0°C and start melting. If the particles have relatively small terminal velocities (or equivalently small size; see section 2.3), they may reach the ground as raindrops indistinguishable from those formed by coalescence. Alternatively, in cold weather or when large hailstones are formed, the precipitation particles may reach the ground unmelted.

Additional discussion on the mechanisms of formation and growth of precipitation particles, and the potential human intervention on the mechanisms by technological means are discussed in Chapter 79, "Abstraction of atmospheric humidity" of the present volume.

2.3 Properties of precipitation particles

2.3.1 Terminal velocity

The terminal velocity $U_X(D)$ of a precipitable particle of type $X = R$ (rain), H (hail), S (snow), and effective diameter D is the maximum velocity this particle may develop under gravitational settling relative to its ambient air. In theory, $U_X(D)$ can be obtained by balancing the weight of the particle with the sum of the static and dynamic buoyancy (i.e. drag forces) on the particle. For a rigid spherical raindrop, one obtains $U_R(D) \propto \sqrt{D}$ (e.g. Rogers and Yau, 1996).

Theoretical calculation of $U_X(D)$ becomes more complicated when the dynamical characteristics of the falling particles depend on their linear size D and the ambient temperature T . For example, droplets with diameters D smaller than about 0.35 mm are approximately spherical, drops with diameters in the range 0.35-1 mm tend to deform by the aerodynamic shear receiving a more elliptical shape, whereas larger drops frequently break down into smaller droplets due to excessive elongation or surface vibrations (e.g. Testik and Barros, 2007; Uijlenhoet, 2008). Moreover, the crystalline structure, shape, size and, hence, the aerodynamic properties of snowflakes depend on the ambient temperature T (Fletcher, 1962; Locatelli and Hobbs, 1974; Houghton, 1985; and Rogers and Yau, 1996).

In the absence of exact theoretical solutions for the terminal velocity $U_x(D)$ of precipitation particles under complex atmospheric conditions, several empirical formulae have been developed (e.g. Gunn and Kinzer, 1949; Liu and Orville, 1969; Wisner et al., 1972; Locatelli and Hobbs, 1974; Atlas and Ulbrich, 1977; Lin et al., 1983). According to Liu and Orville (1969), who performed a least squares analysis of Gunn and Kinzer's (1949) data, the terminal velocity of raindrops of diameter D can be approximated by a power-law type relationship:

$$U_R(D) = a D^b \quad (21)$$

where $a = 2115 \text{ cm}^{1-b}/\text{s}$ and $b = 0.8$ are empirical constants. For raindrops with diameters in the range $0.5 \leq D \leq 5 \text{ mm}$, Atlas and Ulbrich (1977) (see also Uijlenhoet, 2008) suggest the use of equation (21) with parameters $a = 1767 \text{ cm}^{1-b}/\text{s}$ and $b = 0.67$.

For hail, Wisner et al. (1972) suggest:

$$U_H(D) = D^{1/2} \left(\frac{4g\rho_H}{3C_D\rho} \right)^{1/2} \quad (22)$$

where $g = 9.81 \text{ m/s}^2$ is the acceleration of gravity, $\rho \approx 1.2 \text{ kg/m}^3$ is the density of air, $\rho_H = 800\text{-}900 \text{ kg/m}^3$ is the density of the hailstone and $C_D = 0.6$ is a drag coefficient.

For graupel-like snow of hexagonal type, Locatelli and Hobbs (1974) suggest:

$$U_S(D) = c D^d \quad (23)$$

where $c = 153 \text{ cm}^{1-d}/\text{s}$ and $d = 0.25$ are empirical constants that, in general, depend on the shape of the snowflakes (e.g. Stoelinga et al., 2005).

$U_x(D)$ relationships other than power laws have also been suggested (e.g. Beard, 1976, and review by Testik and Barros, 2007). However, the power-law form in equations (21)-(23) is the only functional form that is consistent with the power law relations between the radar reflectivity factor Z (see section 3.2) and the rainfall intensity i (Uijlenhoet 1999, 2008).

2.3.2 Size distribution

A commonly used parameterization for the size distributions of precipitation particles is that introduced by Marshall and Palmer (1948). According to this parameterization, precipitation particles have exponential size distributions of the type

$$n_x(D) = n_{0x} \exp(-b_x D), \quad X = R, H, S \quad (24)$$

where the subscript X denotes the type of the particle: rain (R), hail (H) or snow (S), D is the effective diameter of the particle, b_x is a distribution scale parameter with units of $[\text{Length}^{-1}]$ (see below) and n_{0x} is an intercept parameter that depends on the type of the particle with units of $[\text{Length}^{-4}]$: that is, number of particles per unit diameter and per unit volume of air (see below).

To determine the parameters n_{0R} and b_R in equation (24) for rainfall, Marshall and Palmer (1948) used observations from summer storms in Canada. The study reported a constant value of the intercept parameter $n_{0R} = 8 \times 10^{-2} \text{ cm}^{-4}$, whereas the scale parameter b_R was found to vary with the

rainfall intensity i at ground level as: $b_R = 41 i^{-0.21} \text{ cm}^{-1}$, where i is in mm/h. Clearly, the mean raindrop size $1/b_R$ increases with increasing rainfall intensity i .

Gunn and Marshall (1958) used snowfall observations from Canada to determine the parameters n_{0S} and b_S for snow. The study concluded that both n_{0S} and b_S depend on the precipitation rate as:

$$n_{0S} = 0.038 i^{-0.87} \text{ cm}^{-4}, \quad b_S = 25.5 i^{-0.48} \text{ cm}^{-1} \quad (25)$$

where i is the water equivalent (in mm/h) of the accumulated snow at ground level. Similar to the mean raindrop size, the mean snowflake size $1/b_S$ increases with increasing i . A modification to the distribution model of Gunn and Marshall (1958) has been proposed by Houze *et al.* (1979) and Ryan (1996). According to these authors the intercept parameter for snow, n_{0S} , is better approximated as a decreasing function of the temperature T of the ambient air. The latter is responsible for the properties and structure of ice crystals (see section 2.2).

Federer and Waldvogel (1975) used observations from a multicell hailstorm in Switzerland to determine the parameters n_{0H} and b_H for hail. The study showed pronounced variability of the intercept parameter $n_{0H} = 15 \times 10^{-6}$ to $5.2 \times 10^{-4} \text{ cm}^{-4}$, moderate variability of the scale parameter $b_H = 3.3 - 6.4 \text{ cm}^{-1}$, and concluded to an exponential mean size distribution for hailstones with constant parameters: $n_{0H} \approx 1.2 \times 10^{-4} \text{ cm}^{-4}$ and $b_H \approx 4.2 \text{ cm}^{-1}$.

Alternative models, where the size distributions of precipitation particles are taken to be either gamma or lognormal, have also been suggested (e.g. Ulbrich, 1983; Feingold and Levin, 1986 and Joss and Waldvogel, 1990). However, the exponential distribution model introduced by Marshall and Palmer (1948) has been empirically validated by a number of studies (see e.g. Kessler, 1969; Federer and Waldvogel, 1975; Joss and Gori, 1978; Houze *et al.*, 1979; Ryan, 1996; Ulbrich and Atlas, 1998; Hong *et al.*, 2004), and has found the widest application by being used in the cloud resolving schemes of many state-of-the-art numerical weather prediction (NWP) models (e.g. Cotton *et al.*, 1994; Grell *et al.*, 1995; Reisner *et al.*, 1998; Thompson *et al.*, 2004; Skamarock *et al.*, 2005).

A more general formulation for the size distribution of precipitation particles, which includes the exponential model of Marshall and Palmer (1948), and the gamma and lognormal models as special cases, was suggested by Sempere-Torres *et al.* (1994, 1998). According to their formulation, the size distribution of precipitation particles can be parameterized as

$$n_X(D) \propto i^f g(D_X/i^z), \quad X = R, H, S \quad (26)$$

where f and z are constant exponents, i is the precipitation rate and $g(x)$ is a scalar function with parameter vector \mathbf{a} . For a certain form of g , the functional dependence of the parameters f , z and \mathbf{a} is obtained by satisfying the equation for the theoretical precipitation rate originating from particles with size distribution $n_X(D)$ (Sempere-Torres *et al.*, 1994, 1998; Uijlenhoet, 2008):

$$i = \frac{\pi}{6} \int_0^{\infty} n_X(D) U_X(D) D^3 dD, \quad X = R, H, S \quad (27)$$

where $U_X(D)$ is given by equations (21) - (23). Note, however, that the units of n_X depend on those used for D and i and, of course, the functional form of $g(x)$.

2.4 Clouds and precipitation types

Clouds owe their existence to the process of condensation, which occurs in response to several dynamical processes associated with motions of air masses, such as orographic or frontal lifting (see section 2.5), convection and mixing. At the same time, clouds and the resulting precipitation influence the dynamical and thermodynamic processes in the atmosphere. For example, clouds affect air motions through physical processes, such as the redistribution of the atmospheric water and water vapour, the release of latent heat by condensation and the modulation of the transfer of solar and infrared radiation in the atmosphere.

A cloud system is formed by a number of recognizable isolated cloud elements that are identifiable by their shape and size (e.g. Scorer and Wexler, 1963; Austin and Houze, 1972; Orlanski, 1975). On the lowest extreme, cloud systems with scale of about 1 km or less are classified as microscale systems. On the highest extreme, atmospheric phenomena of linear extent of 1000 km and upwards are classified in the synoptic scale and include the cloud systems associated with baroclinic instabilities, and extratropical cyclones (i.e. low-pressure centres). In between those two extreme scales, atmospheric phenomena with linear extent between a few kilometres to several hundred kilometres are the so-called mesoscale phenomena. These phenomena are more likely associated with atmospheric instabilities, as well as frontal and topographic lifting. Mesoscale phenomena include many types of clouds and cloud systems that are usually classified in two main categories: *stratiform* and *convective* (cumulus) cloud systems. In general, stratiform cloud systems have the shape of a flat appearing layer and produce widespread precipitation associated with large scale ascent, produced by frontal or topographic lifting, or large scale horizontal convergence. By contrast, convective cloud systems have large vertical development, produce localized showery precipitation and are associated with cumulus-scale convection in unstable air. Next we focus on the structure of these systems and the forms of precipitation they produce.

2.4.1 Cumulus cloud systems

Cumulus clouds are formed by small thermals (upward moving air parcels heated by contact to the warm ground) where condensation occurs and they grow to extend vertically throughout the troposphere. Their vertical extent is controlled by the depth of the unstable layer, while their horizontal extent is comparable to their vertical extent. A typical linear dimension of a cumulus cloud is 3-10 km, with updraft velocities of a few metres per second (Rogers and Yau, 1996).

Observations performed by Byers and Braham (1949; see also Weisman and Klemp, 1986) revealed that convective storms are formed by a number of cells, each one of which passes through a characteristic cycle of stages (Fig. 15). The *cumulus stage* of a cell is characterized by an updraft throughout most of the cell. At this stage, which lasts approximately 10-20 min, the cell develops and expands vertically while the air becomes saturated and hydrometeors grow due to vapour condensation and turbulent coalescence (see section 2.2).

Some ice and water particles grow large enough to fall relative to the ambient updraft and initiate a downdraft within the cell. The downdraft is initially in saturated condition, but as it moves toward the lower troposphere and mixes with sub-saturated air, evaporational cooling occurs, which introduces negative buoyancy and accelerates the downdraft. This is the start of the *mature stage* of the cell, which lasts for approximately 15-30 min. The air of the downdraft reaches the ground, as a

cold core, and changes the surface wind pattern. This change may initiate a new thermal at a neighbouring location which might grow to a new cell. The downdraft interferes with the updraft at the lower levels of the cloud and finally cuts off the updraft from its source region. At this point, the cell enters its *dissipating stage*. At this stage, which lasts for about 30 min, the updraft decays and consequently the precipitation source is eliminated.

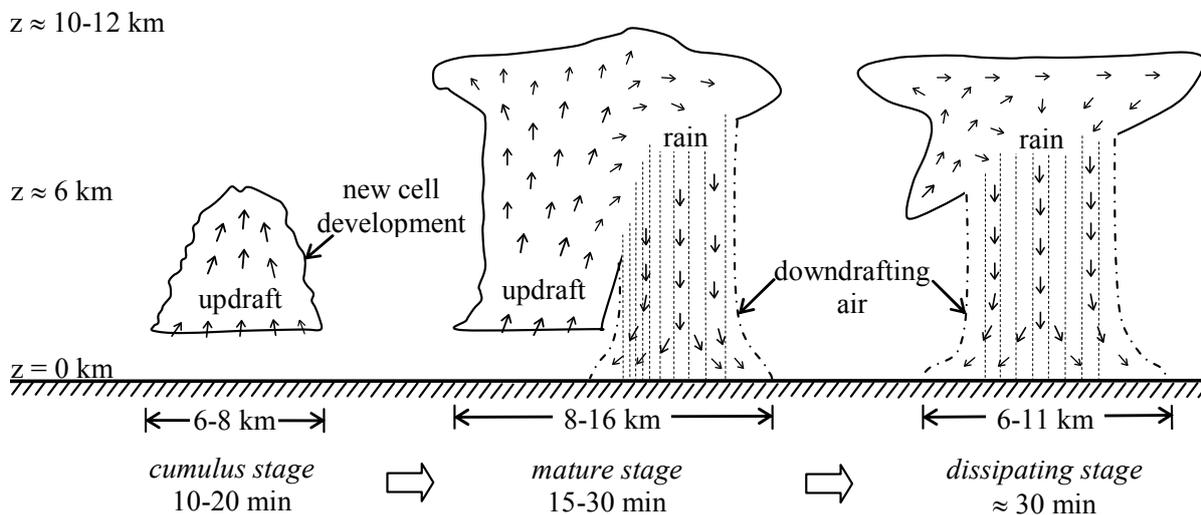


Fig. 15 Stages of development of convective cells (adapted from Weisman and Klemp, 1986).

2.4.2 Stratus cloud systems

Stratus clouds are associated with mesoscale, or even synoptic, vertical air motions that arise from large scale horizontal convergence and frontal or orographic lifting of moist air masses. The ascending motion of air is weak (i.e. a few tens of centimetres per second) relative to cumulus convection, but it extends over large areas and durations to produce widespread rain or snow.

The life time of a stratus formation is of the order of days, and its size may extend over hundreds of kilometres horizontally. The ascended air masses, having the form of a flat appearing layer, remain convectively stable even after they are lifted to higher altitudes. Since atmospheric turbulence is not intense, initiation of rain is mainly dominated by the ice particle growth due to vapour deposition (the Bergeron-Findeisen mechanism; section 2.2), when the ascended air masses are thick enough to reach subfreezing temperatures. In general, thin stratus clouds are usually non-precipitating, whereas thick stratus clouds (i.e. 1-2 km vertical extent) are capable of producing substantial widespread rain or snow.

Although the classification of cloud systems in stratiform and convective is useful for observation purposes, it cannot be considered sharp (Harrold and Austin, 1974). Observations from radars or raingauges show that widespread precipitation has fine-scale structure with intense precipitation regions confined to elements with size of a few kilometres, while rainfall features of convective origin (e.g. cells) can grow and/or cluster over a large region producing continuous precipitation similar to that of stratiform formations.

In general, convective rainfall patterns are non uniform and are associated with locally intense rainfall regions ranging in size from 3-10 km. The latter evolve rapidly in time and are separated by

areas free of precipitation. By contrast, stratiform patterns are associated with less pronounced small scale structure and a wider overall extent that persists in time.

2.5 Precipitation generating weather systems

2.5.1 Fronts

Atmospheric circulation is formed by advecting air masses with fairly uniform characteristics. Depending on their source of origin, different air masses may have different temperatures and moisture contents. For example, continental air masses are dryer and their temperatures vary in a wider range relative to maritime air masses. The interface of two opposing air masses with different temperatures and moisture contents is usually referred to as a *front*. Along this interface, the warmer and lighter air rises above the colder and denser air. The vertical lifting causes the warmer air to cool adiabatically, the water vapour to condense and, hence, precipitation to form.

A *cold front* occurs when advancing cold air wedges itself under warmer air and lifts it (Fig. 16.a), whereas a *warm front* develops when faster moving warm air overrides a colder and denser air mass (Fig. 16.b). An *occluded front* forms when warm air is trapped between two colder and denser air masses. An example of an occluded front is shown in Fig. 16.c, where a cold front catches up a slower moving warm front.

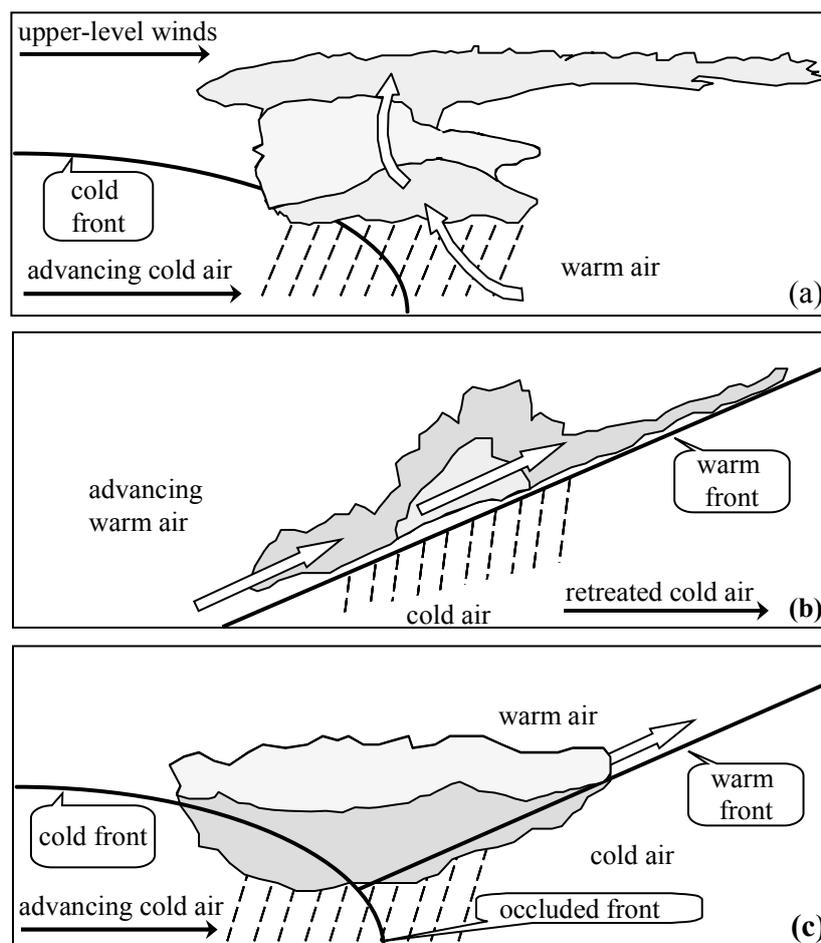


Fig. 16 Schematic illustration of different types of fronts: (a) cold front, (b) warm front and (c) occluded front (adapted from Koutsoyiannis and Xanthopoulos, 1999).

Fronts may extend over hundreds of kilometres in the horizontal direction and are associated with vertical wind speeds of the order of a few tens of centimetres per second. This range of values is in accordance with vertical motions caused by the horizontal wind convergence of synoptic-scale low-level flow. Hence, frontal precipitation is mostly stratiform with widespread rain or snow over large areas and durations. Note, however, that embedded within the areas of frontal precipitation there are mesoscale regions that exhibit cellular activity.

2.5.2 Mechanical lifting and orographic precipitation

Orographic precipitation occurs when horizontally moving warm and humid air meets a barrier such as a mountain range. In this case, the barrier causes uplift of the incoming air. As the moist air moves upslope, it cools adiabatically, water vapour condenses to liquid water or ice (depending on the altitude where the dew point temperature occurs) and precipitation is likely to form (e.g. Smith, 1993; Hemond and Fechner-Levy, 2000). In general, orographic precipitation (unless combined with other mechanisms such as cyclonic activity and fronts) is narrow banded since it occurs in association with water vapour condensation by mechanical lifting, a process that becomes effective at a certain elevation along the topography. After surpassing the top of the mountain range, on the lee side, the air moves downward and this causes adiabatic warming, which tends to dissipate the clouds and stop the precipitation, thus producing a *rain shadow*.

2.5.3 Extratropical cyclones

Extratropical cyclones are synoptic scale low pressure systems that occur in the middle latitudes (i.e. pole-ward of about 30° latitude) and have length scales of the order of 500-2500 km (e.g. Hakim, 2003). They usually form when two air masses with different temperatures and moisture contents that flow in parallel or are stationary become coupled by a pre-existing upper level disturbance (usually a low pressure centre) near their interface.

An example is the formation of extratropical cyclones along the interface of midlatitude westerlies (i.e. winds that flow from West to East; e.g. Lutgens and Tarbuck, 1992) with the equator-ward moving polar, and thus colder, air masses (i.e. polar easterlies). As shown in Fig. 17, which refers to the Northern Hemisphere, the motion of both warm and cold air masses is caused by pressure gradients and their direction is south-north and north-south, respectively. However, these directions are diverted to the right (in the Northern Hemisphere) by Coriolis forces. The initial disturbance formed by the shear along the interface of the two air masses (Fig. 17.a) grows as the warmer and lighter air rises above the colder air and starts rotating in an emerging spiral called the *cyclone* (Fig. 17.b). As the cyclone evolves, the cold front approaches the slower moving warm front (Fig. 17.c) and then catches up with it forming an *occluded front* (Fig. 17.d). Finally, mixing between the two air masses causes the fronts to lose their identities and the cyclone to dissipate. The adiabatic cooling of the warm and moist air results in a widespread region of stratiform precipitation that propagates with the upper-level flow far beyond the fronts (Fig. 17.e).

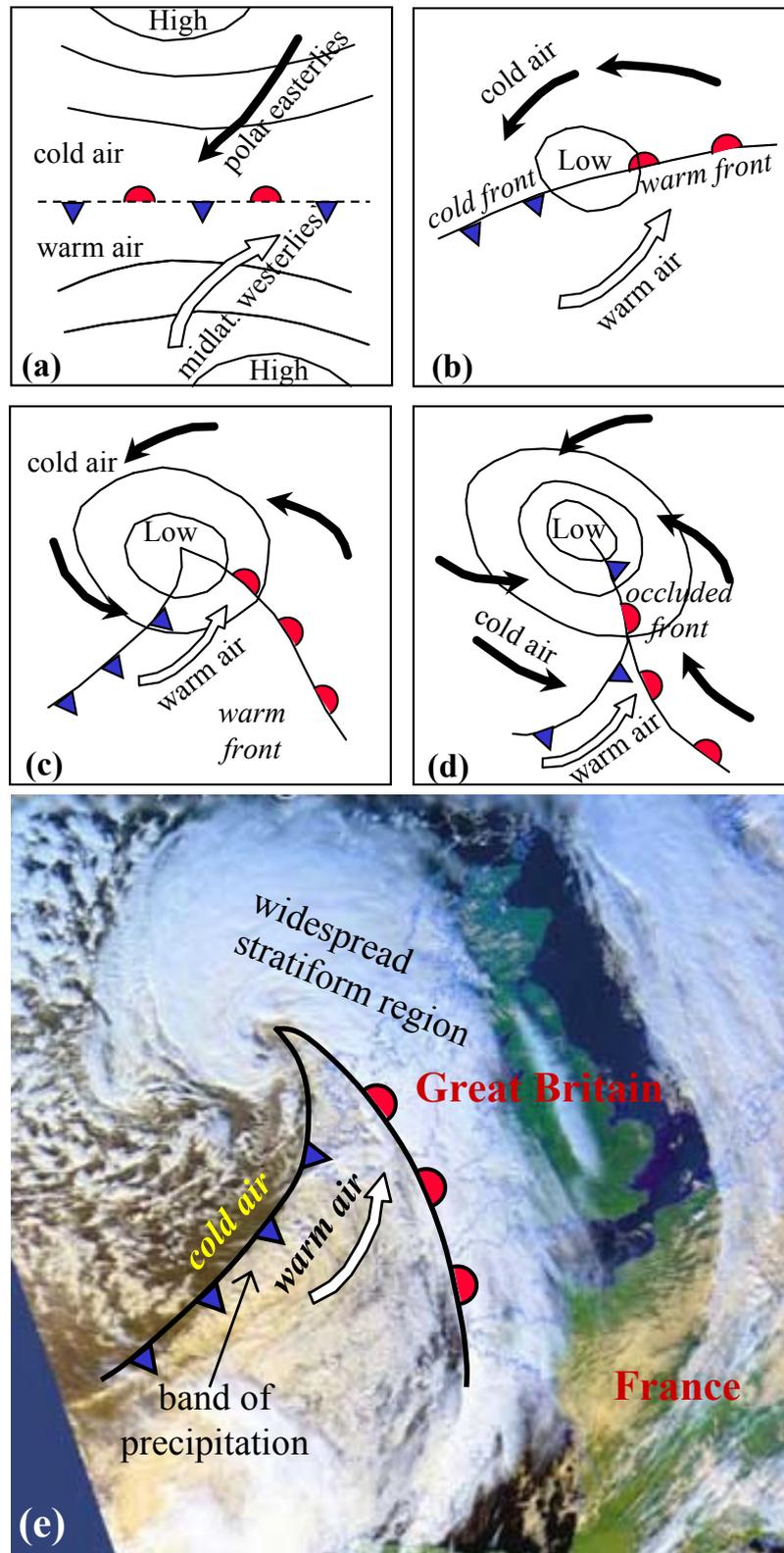


Fig. 17 (a)-(d) Schematic illustration of the evolution of an extratropical cyclone at the interface of midlatitude westerlies and the equator-ward moving polar easterlies (adapted from Koutsoyiannis and Xanthopoulos, 1999). (e) Extra-tropical cyclone over the British Isles on 17 January 2009: motion of air masses, fronts and characteristic precipitation regions (www.ncdc.noaa.gov/sotc/index.php?report=hazards&year=2009&month=jan).

2.5.4 Isolated extratropical convective storms

A short-lived *single-cell* is the simplest storm of convective origin. Single cells have horizontal cross-sections of the order of 10-100 km² and move with the mean environmental flow over the lowest 5-7 km of the troposphere. The stages of development of a single-cell storm were discussed in section 2.4. The *multi-cell* storm is a cluster of short-lived single cells with cold outflows (i.e. downdrafts) that combine to form a large gust front (Weisman and Klemp, 1986). The convergence along the leading edge of the front triggers new updraft development and subsequent formation of new cells. Because of the new cell development, multi-cell storms may last several days and span over large areas with linear extents of hundreds of kilometres.

The *super-cell* storm is the most intense of all isolated convective storms. It has a lifetime of several hours, it exhibits large vertical development and produces strong winds, heavy rainfall or hail and long-lived tornadoes, that is, intense vortices with diameter of the order of 100-500 m (e.g. Browning and Ludlam, 1962; Rotunno, 1986; Weisman and Klemp, 1986; Bluestein, 2003), where the updrafts and downdrafts are displaced horizontally and interact mutually to sustain a long-lived circulation (Fig. 18). The updraft enters at low levels and ascends in a region called the *vault*, which might penetrate into the stratosphere. Super-cell storms usually evolve from multi-cell formations when the magnitude of the vertical wind shear, defined as the difference between the density weighted mean wind over the lowest 6 km and a representative surface layer wind (e.g. 500 m mean wind), suffices to produce a long-lived rotating updraft that mutually interacts with the downdraft (e.g. Weisman and Klemp, 1982, 1984, 1986).

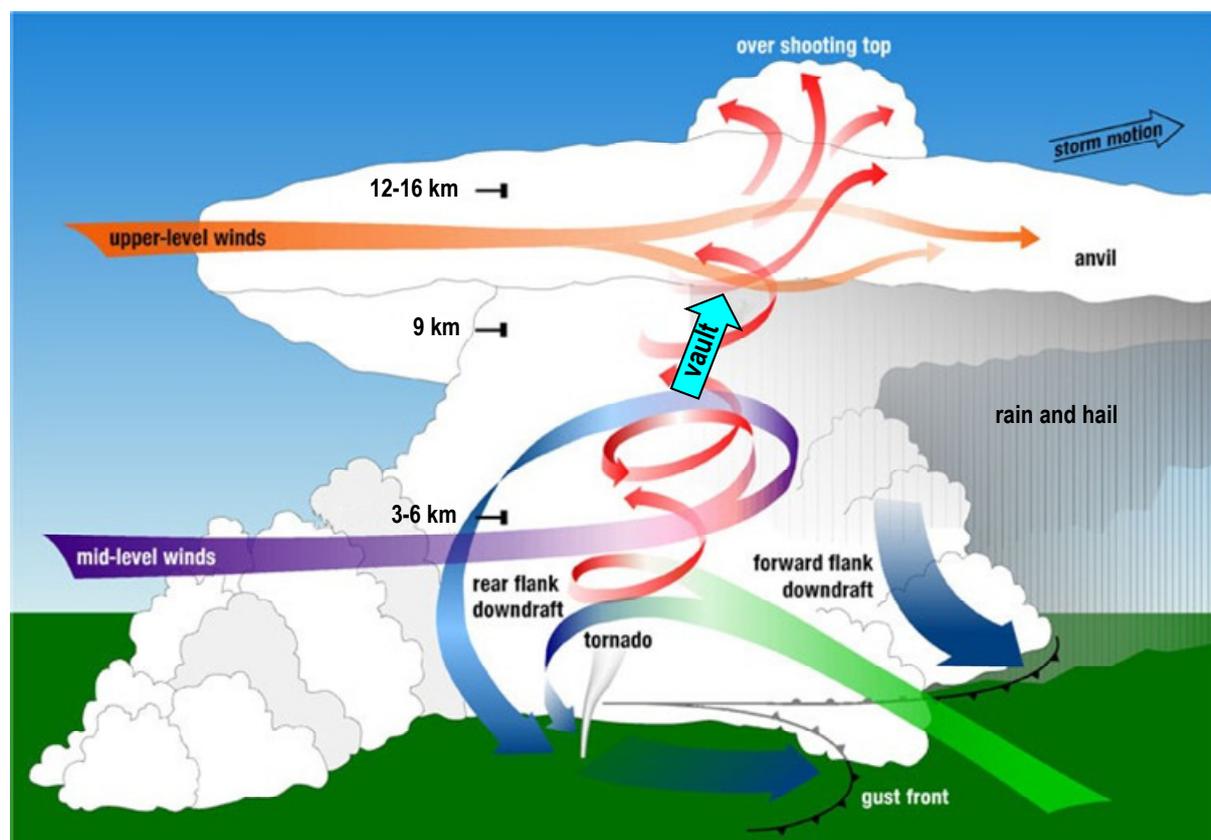


Fig. 18 Schematic illustration of the wind circulation in a super-cell storm (adapted from www.nssl.noaa.gov/primer/tornado/).

2.5.5 Extratropical squall lines and rainbands

Intense rainfall events are usually organized in lines (i.e. squall lines) and bands (i.e. rainbands) with characteristic scales of hundreds of kilometres. According to Hane (1986), *rainbands* are sufficiently elongated rainfall areas that are non-convective or weakly convective, and *squall lines* include all linear convective structures stronger than rainbands. Those large scale features are considered to be manifestations of the large-mesoscale horizontal circulation in association with spatial fluctuations of the surface temperature and moisture content of atmospheric air masses.

The conditions for squall line formation are: (1) a convectively unstable near-surface environment (i.e. moist and warm near-surface air with relatively cold air aloft) to maintain the development of convective cells, (2) a layer of dry air directly above the near-surface moist air to enhance development of an intense and wide cold downdraft by evaporative cooling (i.e. the dry middle-level air causes precipitation particles to evaporate and a negatively buoyant cold front to form), and (3) a triggering mechanism for release of the convective instability (e.g. frontal or orographic lifting). Once the squall line has formed, it feeds itself through convergence along the cold gust front. This convergence produces strong ascent and forms new cells ahead of the storm.

Rainbands in extratropical regions occur primarily in association with well organized extratropical cyclones (Hane, 1986). In this case, precipitation is maintained by the ascent resulting from the warm advection of the advancing cyclone, with subsequent formation of a widespread region of stratiform precipitation (section 2.5.3). Extratropical rainbands can also be formed in synoptic-scale environments other than those associated with cyclonic circulation. An example is the environments associated with the development of symmetric instabilities (e.g. Bennetts and Sharp, 1982; Seltzer *et al.*, 1985).

2.5.6 Monsoons

The term *monsoon* generally applies to climates that exhibit long, distinct and remarkably regular rainy and dry periods associated with the spatial distribution of solar heating during summer and winter. According to a definition proposed by Ramage (1971), a monsoon climate is characterized by: (1) prevailing wind directions that shift by at least 120° between January and July, (2) prevailing wind direction that persists at least 40% of the time in January and July, (3) mean wind speeds that exceed 3 m/s in either January or July, and (4) fewer than one cyclone-anticyclone alternation every 2 years in either January or July in a 5° latitude-longitude rectangle. In essence, Ramage's (1971) criteria exclude most extratropical regions with prevailing synoptic-scale cyclonic and anticyclonic circulations and, in addition, require the mean wind direction to be driven and sustained exclusively by the seasonally varying temperature contrast between continental and oceanic masses. Under these constraints, only India, South-Eastern Asia, Northern Australia and West and central Africa have monsoon climates (Slingo, 2003). For example, in India, about 80% of the mean annual rainfall accumulation (about 2 m) occurs during the months of June, July and August (Smith, 1993).

The main driving mechanism for monsoons is the temperature contrast between continental and oceanic masses due to the seasonal cycle of solar heating. More precisely, the lower thermal inertia of continental masses relative to oceans causes the former to heat up more rapidly during spring and summer by the solar radiation. This results in a sharp temperature gradient, which causes a humid flow of oceanic near-surface air to direct toward the land (something similar to a massive sea

breeze). As it reaches the land, the humid air warms up and rises, water vapour condenses to liquid water and rain falls. A similar process occurs during winter, when the continental air masses cool up more rapidly than the surrounding ocean water, with subsequent formation of a cold and dry massive low-level flux toward the ocean.

An important factor that determines the intensity of monsoon rainfall is the geographical orientation of continents and oceans relative to the equator (Slingo, 2003). For example, the North-South orientation of the South-Eastern Asian and Northern Australian monsoon system allows the dry outflow from the winter continent to warm up and load moisture from the ocean, flow across the equator toward the summer hemisphere and, eventually, feed the monsoon rains over the summer continent. This is also the reason why the largest rainfall accumulations for durations larger than 24 hours are associated with the Asian-Australian monsoon system (Smith, 1993).

2.5.7 Tropical cyclones

Tropical cyclones (TCs) form a particular class of synoptic scale low pressure rotating systems that develop over tropical or subtropical waters (Anthes, 1982; Landsea, 2000). These systems have linear extent of the order of 300-500 km and are characterized by well organized convection and cyclonic (counter-clockwise in the Northern Hemisphere) surface wind circulation around a relatively calm low-pressure region, called the eye of the storm (Figs. 19 and 20).

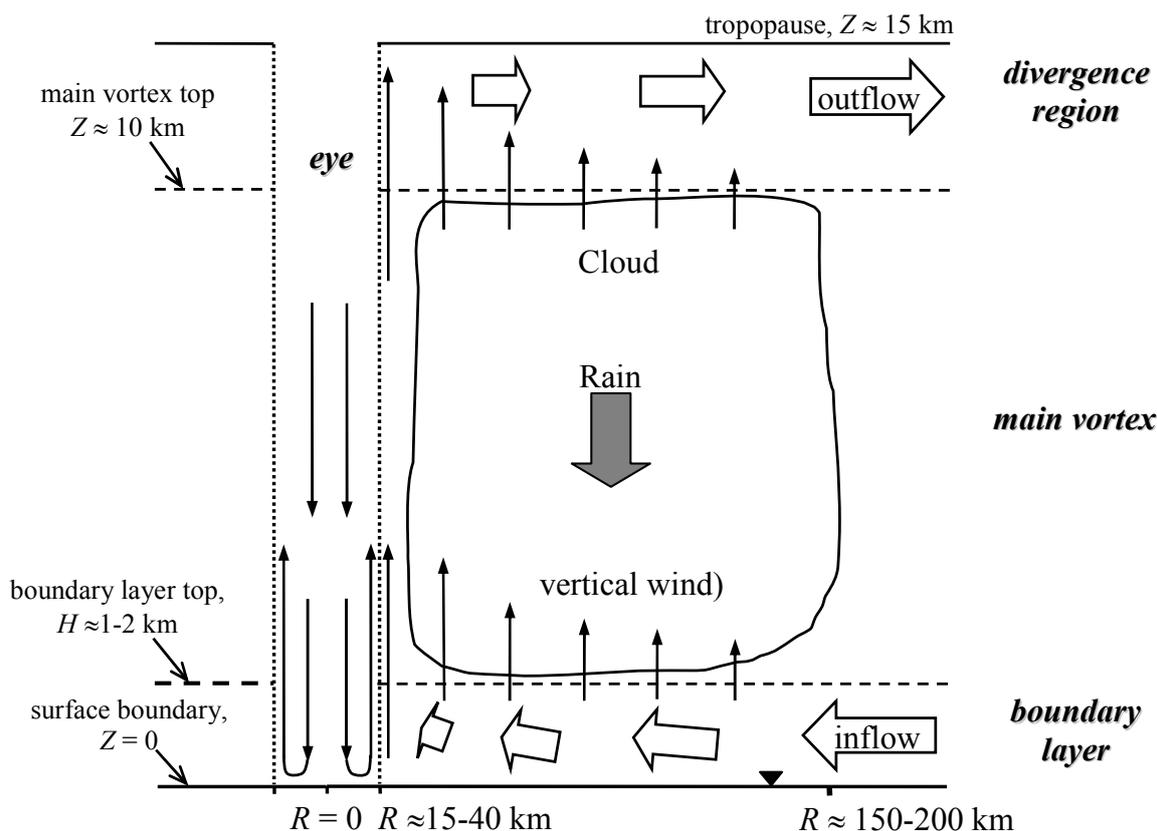


Fig. 19 Schematic representation of the structure of a mature hurricane.

Tropical cyclones with sustained wind speeds in the range 17-32 m/s are called *tropical storms* (TS), whereas stronger TCs are usually referred to as *hurricanes* (i.e. when observed in the North Atlantic Ocean, in the Northeast Pacific Ocean east of the dateline and in the South Pacific Ocean east of

160°E) or *typhoons* (i.e. when observed in the Northwest Pacific Ocean west of the dateline). Note, however, that extreme rainfall accumulations for durations of the order of a day or higher are usually produced by moderate or even low intensity TCs (Langousis and Veneziano, 2009b). An example is the tropical storm (TS) Allison in 2001, which looped over the Houston area causing rainfall accumulations in excess of 850 mm. According to the US National Oceanic and Atmospheric Administration (NOAA; Stewart, 2002), TS Allison (2001) ranks as the costliest and deadliest TS in the history of the US with 41 people killed, 27 of which were drowned, and more than \$6.4 billion (2007 USD) in damage.

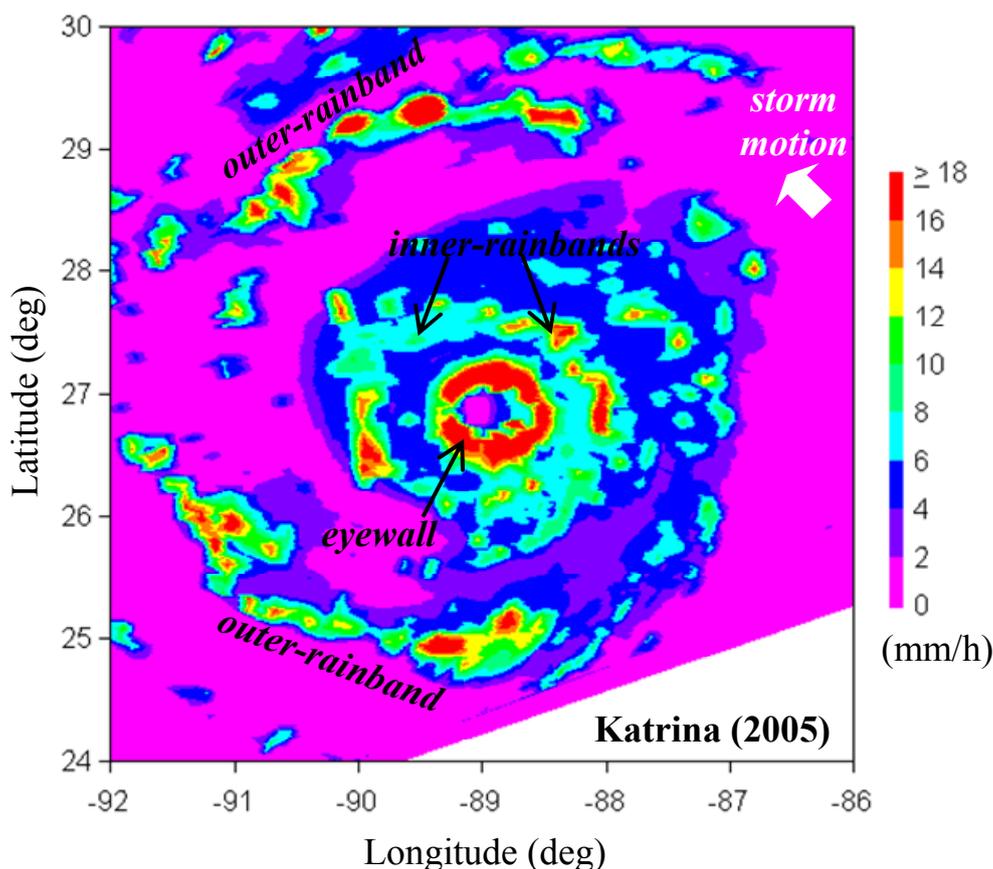


Fig. 20 TRMM microwave imager (TMI) rainfall retrievals for hurricane Katrina on 28 August (2005) at 21:00 UTC (frame 44373): different types of rainbands and their locations relative to the centre of the storm.

The genesis and development of tropical cyclones require the following conditions to be maintained (e.g. Gray, 1968, 1979): (1) warm ocean waters (surface temperature $T > 27^{\circ}\text{C}$); (2) a conditionally unstable atmosphere where the air temperature decreases fast with height; (3) a relatively moist mid-troposphere to allow the development of widespread thunderstorm activity; (4) a minimum distance of about 500 km from the equator in order for the Coriolis force to be sufficiently large to maintain cyclonic circulation; (5) a near-surface disturbance with sufficient vorticity and low-level convergence to trigger and maintain the cyclonic motion; and (6) low magnitude of vertical wind shear (less than 10 m/s), defined as the difference between the 200- and 850-hPa horizontal wind velocities in the annular region between 200 and 800 km from the TC centre (Chen *et al.*, 2006). The latter condition is important for the maintenance of the deep convection around the centre of the cyclone.

At a first approximation, a tropical cyclone can be seen as a heat engine fuelled by the buoyant motion of warm and saturated (hence convectively highly unstable) air masses that lie directly above the warm tropical and subtropical ocean waters (e.g. Emanuel, 1986, 1989; Renno and Ingersoll, 1996; Marks, 2003). By contrast, extratropical cyclones obtain their energy from the horizontal temperature gradients in the atmosphere (section 2.5.3).

During its mature stage, a tropical cyclone includes four distinct flow regions (Yanai 1964; Smith, 1968; Frank, 1977; Willoughby, 1990; Smith, 2000), as depicted in Fig. 19:

1. Away from the surface boundary (in the altitude range from 2-3 km to about 10 km), frictional stresses are negligible and the horizontal winds are in approximate gradient balance (e.g. La Seur and Hawkins, 1963; Hawkins and Rubsam, 1968; Holland, 1980; Willoughby, 1990, 1991; Vickery *et al.*, 2000). In this region, usually referred to as the *main vortex*, the radial inflow is negligible, whereas the tangential flow is maintained by the balance between the inward-directed pressure gradient force and the sum of the outward-directed centrifugal and Coriolis forces.
2. Within the boundary layer (in the altitude range below 1-2 km), frictional stresses decelerate the tangential flow, reduce the magnitude of the Coriolis and centrifugal forces and result in an inward net force that drives low level convergence. Calculations performed by Smith (1968, 2003), Kepert (2001), Kepert and Wang (2001) and Langousis *et al.* (2008) show that the radial inflow in the boundary layer turns upward before it reaches the TC centre causing vertical fluxes of moisture. Langousis and Veneziano (2009a) showed that these fluxes can be used to obtain accurate estimates for the large-scale mean rainfall intensity field in TCs as a function of the TC characteristics.
3. At altitudes in excess of about 10 km the curved isobars, which are responsible for the TC formation and maintenance, start to flatten. As a consequence, the inward directed pressure gradient force that maintains the cyclonic circulation decreases with increasing height leading to an outward directed net force that drives high-level divergence.
4. Finally, there is a core flow region, called the *eye* of the TC, with diameters of the order of 15-40 km. This region is free of cloud with light tangential winds and a downflow close to the axis.

The condensation of water vapour caused by the ascending motion of humid near-surface air leads to the formation of cloud systems. These systems, which are usually precipitating, are organized around the cyclone centre into long quasi-circular formations usually referred to as *rainbands*. Despite variations of rainband characteristics from one storm to another and during the evolution of a single storm (e.g. Miller, 1958; Barnes *et al.*, 1983; Marks, 1985; Molinari *et al.*, 1999), a number of studies (Willoughby *et al.*, 1984; Powell, 1990; Molinari *et al.*, 1994, 1999, among others) have shown that rainbands, depending on their location relative to the storm centre, share similar structural characteristics and can be organized in three distinct classes: i.e. eyewall, inner-rainbands and outer-rainbands (Fig. 20):

1. The *eyewall* is a well developed convective band that surrounds the eye of the TC. This band has width of approximately 10-15 km with upward directed quasi-steady velocities in the range of 0.5-3 m/s or more, with the larger values being associated with more intense

systems. The quasi-steady updrafts mostly reflect the radial convergence of horizontal fluxes, which become maximum close to the eye of the TC (Smith, 1968; Shapiro, 1983; Kepert, 2001). The eyewall almost always has the highest cloud tops (Jorgensen, 1984a), contains the largest annular mean rainfall intensity (Marks, 1985; Houze *et al.*, 1992) and exhibits weak cellular structure as evidenced by radar observations (e.g. Jorgensen, 1984b; Marks, 1985).

2. The *inner-rainbands* (Molinari *et al.*, 1994; 1999) are a group of spiral bands located outside the eyewall at radial distances smaller than approximately 120 km, and are also referred to as *stationary band complex* (Willoughby *et al.*, 1984). This group moves slowly, if at all, and maintains a rather fixed position relative to the vortex. Rainfall inside the inner-rainband region is mostly stratiform, with active convection covering 5-10% of the total rainfall area and contributing 40-50% of the total rainfall volume (e.g. Marks, 1985; Marks and Houze, 1987; Marks *et al.*, 1992).
3. *Outer-rainbands* typically occur at radial distances larger than approximately 150 km from the TC centre (e.g. Powell, 1990; Molinari *et al.*, 1994). They develop by the increased convergence at the boundary of the vortex envelope, where the convectively unstable environmental air flows around the storm and gives rise to formation of convective cells (e.g. Beer and Giannini, 1980; Ooyama, 1982; Molinari *et al.*, 1994). Consequently, outer-rainbands have more cellular structure than inner ones, which develop in a less unstable atmosphere.

3 Precipitation observation and measurement

3.1 Point measurement of precipitation

3.1.1 Measuring devices

The measurement of precipitation at a point is as easy as placing a bucket at the point of observation and periodically measuring the quantity of water it collects. The collected volume divided by the area of the opening is the precipitation depth. Due to this simplicity, such gauges have been used systematically since many centuries, and must have been discovered independently in different times, perhaps even in the antiquity, and in different places in the world, such as in ancient Greece and ancient India (Kosambi, 2005). However, their records have not survived, so that the oldest available records now are those in Seoul, Korea, already presented in section 1.3 and Fig. 7 (upper), which go back to 1770, even though measurements must have been taken in much earlier periods since 1441 (Arakawa, 1956).

The traditional device for rainfall measurement, known as *rain gauge* or *pluviometer*, is still in use today and, in fact, remains the most accurate device also providing the calibration basis for new measurement devices and techniques. It is a simple cylinder whose opening has an area (e.g. 200-500 cm² according to World Meteorological Organization, 1983) larger than (e.g. tenfold) the cross section of the cylinder, which allows a greater sensitivity of the reading of the rainfall depth in a millimetric ruler attached to the cylinder. In another type of instrument, known as cumulative gauge, which is placed in inaccessible areas, the diameter of the cylinder may be larger than that of the opening so as to be able to store a large volume of precipitation between the times of two visits at the place.

In an autographic (or recording) rain gauge, also known as a *pluviograph* the water depth in the cylinder is recorded with the help of a mechanism involving a floating device. Another type of recording gauge, known as a tipping-bucket gauge, introduces the rainwater to one of a pair of vessels with a known small capacity (typically equivalent to 0.2 mm of rainfall) that is balanced on a fulcrum; when one vessel is filled, it tips and empties and the time of this event is recorded, while the other vessel is brought into position for filling. In traditional autographic devices these recordings are done on a paper tape attached to a revolving cylinder driven by a clockwork motor that is manually wound. In modern instruments this device is often replaced by electronic systems, which provide digital recordings on a data logger and/or a computer connected by a cable or radio link.

A raingauge does not include all precipitation forms, snow in particular, except in light snowfalls when the temperature is not very low and the snow melts quickly. Generally, accurate measurement of snow precipitation (the water equivalent) needs specific instruments, equipped with a heating device to cause melting of snow. If such an instrument is not available, the snow precipitation is estimated as 1/10 of the snowfall depth (see justification in section 1.3).

3.1.2 Typical processing of raingauge data

Measurement of precipitation in raingauges is followed by several consistency checks to locate measurement errors and inconsistencies. Errors are caused by numerous reasons, including human lapses and instrument faults, which may be systematic in case of inappropriate maintenance. Inconsistencies are caused by changes of installed instruments, changes in the environmental conditions (e.g. growing of a tree or building of a house in the vicinity of the raingauge), or movement of the gauge to a new location. When errors are detected, corrections of the measurements are attempted.

The standard meteorological practices include checks of outliers (a measured value is rejected if it is out of preset limits), internal consistency (checks are made whether different variables, e.g. precipitation and incoming solar radiation are compatible with each other), temporal consistency (the consistency of consecutive measurements is checked) and spatial consistency (the consistency of simultaneous measurements in neighbouring stations is checked). Such checks are done in the time scale of measurement (e.g. daily for pluviometers or hourly for pluviographs) but systematic errors can only be located at aggregated (e.g. annual) time scales.

The most popular method applied at an aggregated time scale for consistency check and correcting of inconsistent precipitation data is that of the *double-mass curve*, which is illustrated in Fig. 21. The method has a rather weak statistical background and is rather empirical and graphical (but there is a more statistically sound version in the method by Worsley, 1983). The double-mass curve is a plot of the successive cumulative annual precipitation Σy_i at the gauge that is checked versus the successive cumulative annual precipitation Σx_i for the same period of a control gauge (or the average of several gauges in the same region). If the stations are close to each other and lie in a climatically homogeneous region, the annual values should correlate to each other. A fortiori, if the two series are consistent with each other, the cumulated values Σy_i and Σx_i are expected to follow a proportionality relationship. A departure from this proportionality can be interpreted as a systematic error or inconsistency, which should be corrected. Such a departure is usually reflected in a change in the slope of the trend of the plotted points.

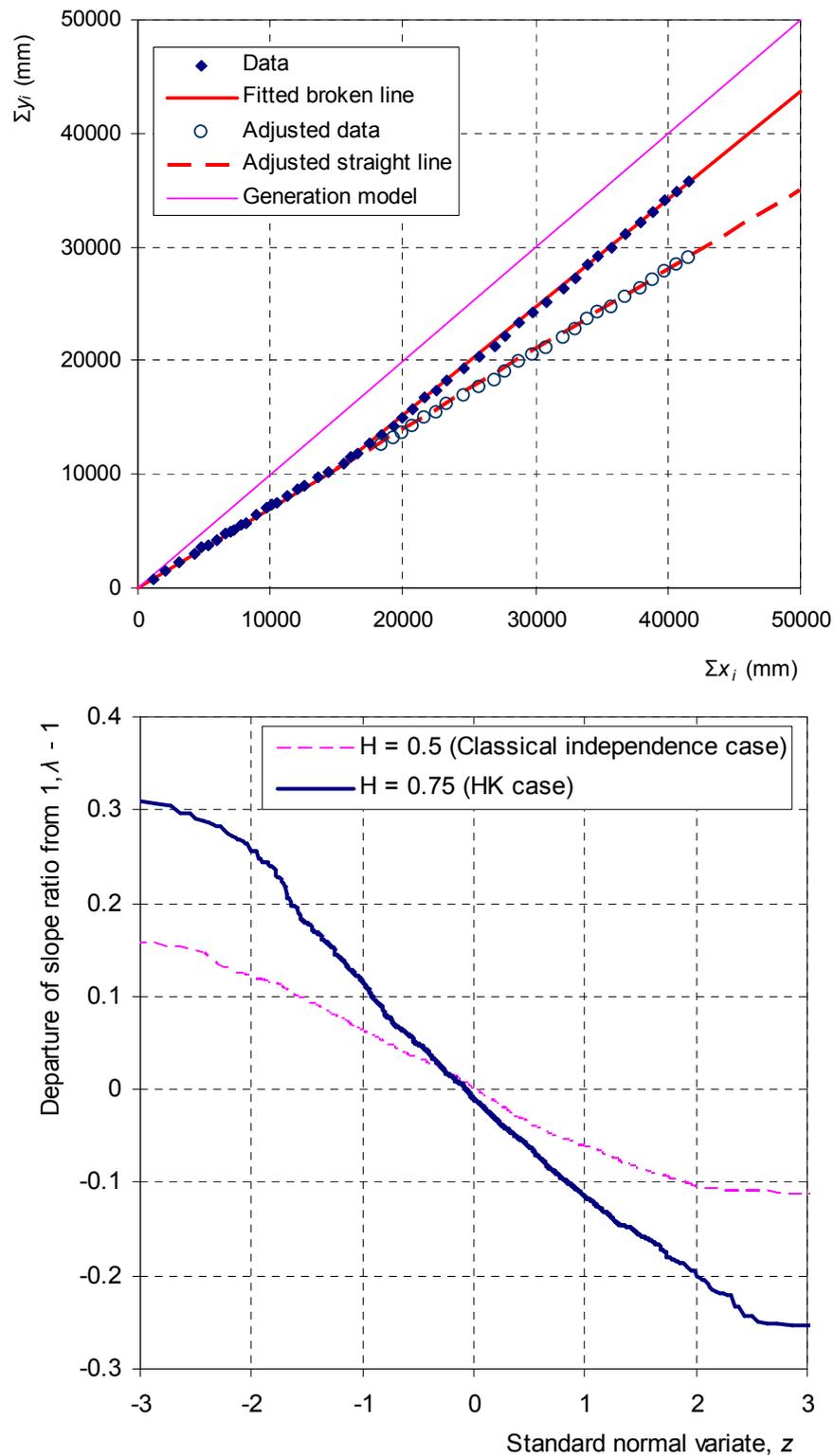


Fig. 21 Illustration of the double-mass curve method and the associated risks in applying it: (upper) Typical double-mass curve for 50 pairs of points, where the first 25 (newest) and the last 25 (oldest) form slopes $m = 0.7$ and $m' = 0.95$, respectively; the adjusted points with $\lambda = m/m' = 0.737$ are also shown. (lower) Comparison of probability distributions of the departure of the ratio λ from unity for series independent in time or with HK behaviour with $H = 0.75$; the distributions were calculated using a Monte Carlo method based on synthetic series with a total size of 1000.

The aggregation of annual values x_i and y_i to calculate Σx_i and Σy_i is typically done from the latest to the oldest year. Fig. 21 (upper) shows the double-mass curve for 50 pairs of values representing annual precipitation at two points, whose cross-correlation (between x_i and y_i) is 0.82. The newest 25 points form a slope of $m = 0.70$, whereas the oldest 25 form a much greater slope, $m' = 0.95$. Assuming that the newest points are the correct ones (with the optimistic outlook that things are better now than they were some years before) we can correct the older 25 annual y_i by multiplying them with the ratio of slopes, $\lambda := m / m' = 0.737$. A second double-mass curve, constructed from the corrected measurements, i.e., from $y_i' := y_i$ for $i \leq 25$ (the newest years) and $y_i' := \lambda y_i$ for $i > 25$ (the oldest years) is also shown in Fig. 21 (upper).

In fact, the data values used in Fig. 21 are not real rainfall data but rather are generated from a stochastic model (Koutsoyiannis 2000b, 2002) so that both stations have equal mean and standard deviation (1000 and 250 mm, respectively), be correlated to each other (with correlation coefficient 0.71) and, most importantly, exhibit HK behaviour (with $H = 0.75$, compatible with the values found in the real world examples of section 1.5). Hence, evidently, all values are “correct”, “consistent”, and “homogeneous”, because they were produced by the same model assuming no change in its parameters. Thus, the example illustrates that the method can be dangerous, as it can modify measurements, seemingly inconsistent, which however are correct. While this risk inheres even in time independent series, it is largely magnified in the presence of HK behaviour. Fig. 21 (lower) provides a normal probability plot of the departure of the ratio λ from unity (where the horizontal axis z is the standard normal distribution quantile and the distributions were calculated by the Monte Carlo method) for two cases: assuming independence in time and assuming HK behaviour with $H = 0.75$ as in the above example. The plots clearly show that, for the same probability, the departure of λ from unity in the HK case is twice as high as in the classical independence case. For the HK case, departures of ± 0.25 from unity appear to be quite normal for 25-year trends and even more so for finer time scales, i.e., ± 0.35 to ± 0.40 for 10-year to 5-year consecutive trends (not shown in figure). Note that the method is typically applied even for correction of as short as 5-year trends (Dingman, 1994), and so its application most probably results in distortion rather than correction of rainfall records.

Apparently, the “correction” of the series using the double-mass curve method removes these “trends” that appear in one of the two time series. Removal of trends results in reduction of the estimated Hurst coefficient or even elimination of the exhibited HK behaviour (Koutsoyiannis, 2003a, 2006b). Thus, if we hypothesize that the HK behaviour is common in precipitation, application of methods such as the double-mass curves may have a net effect of distortion of correct data, based on a vicious circle logic: (a) we assume time independence of the rainfall process; (b) we interpret manifestation of dependence (the HK behaviour in particular) as incorrectness of data; (c) we modify the data so as to remove the influence of dependence; (d) we obtain a series that is much closer to our faulty assumption of independence. The widespread use of the double-mass curve method in routine processing of precipitation time series may thus have caused enormous distortion of real history of precipitation at numerous stations worldwide, also hiding the HK behaviour.

The above discourse aims to issue a warning against unjustified use of consistency checks and correction methods that could eliminate the extreme values (see e.g. the note about the Seoul station in section 1.3) and the long-term variability implied by the HK behaviour; the effect of both these mistreatments of data causes serious underestimation of the design precipitation and flow in

engineering constructions and management decisions. As a general advice for their correct application, we can stress that all methods of this type should never be applied blindly. An inspection of local conditions (environment of the rain gauge station, practices followed by the observer) as well as of the station's archive history is necessary before any action is taken toward altering the data. Unless information on local conditions and archive history justify that inconsistencies or errors exist, corrections of data should be avoided.

3.1.3 Interpolation and integration of rainfall fields

The interpolation problem, i.e., the estimation of an unmeasured precipitation amount y from related precipitation quantities x_i ($i = 1, \dots, n$) is encountered very often in routine hydrologic tasks, such as the infilling of missing values of recorded precipitation at a station or the estimation of precipitation at an ungauged location. The integration problem refers to the estimation of an average quantity y over a specified area (or time period) based on measurements x_i ($i = 1, \dots, n$) of the same quantity and the same time period at different points (or respectively, at different time periods at the same point). The literature provides a huge diversity of methods, most of which, however, could be reduced to a linear statistical relationship applicable to both the interpolation and the integration problem:

$$y = w_1 x_1 + \dots + w_n x_n + \underline{e} \quad (28)$$

where w_i denotes a numerical coefficient (weighting factor) and \underline{e} denotes the estimation error. The same could be written in vector form:

$$y = \mathbf{w}^T \mathbf{x} + \underline{e} \quad (29)$$

with $\mathbf{w} := [w_1, \dots, w_n]^T$ and $\mathbf{x} := [x_1, \dots, x_n]^T$, and the superscript T denotes the transpose of a vector (or a matrix). The notation in equations (28) and (29) suggests that \mathbf{x} , y and \underline{e} are treated as random variables, even though this may not be necessary in some of the existing methods. All interpolation techniques provide a means for estimating the numerical coefficients w_i , either conceptually or statistically, whereas the statistical methods provide, in addition, information about the error. Most commonly, the latter information includes the expected value $\mu_e := E[\underline{e}]$ and its standard deviation $\sigma_e := (\text{Var}[\underline{e}])^{1/2}$. A statistical estimation in which $E[\underline{e}] = 0$ is called unbiased and one in which the mean square error $\text{MSE} := E[\underline{e}^2] = \sigma_e^2 + \mu_e^2$, is the smallest possible is called best; if both these happen, the estimation is called *best linear unbiased estimation* (BLUE). While the BLUE solution is in principle quite simple (see below), the estimation of its weighting factors is not always straightforward. Hence, several simplified statistical methods as well as empirical conceptual methods are in common use. Another reason that explains why such a diversity of methods has emerged is the different type of objects that each of the elements of the vector \mathbf{x} represents. For instance, in temporal interpolation these elements can be observed values at times before and after the time of interpolation. In spatial interpolation, these can be simultaneously observed values for stations lying in the neighbourhood of the point of interpolation. Simultaneous temporal and spatial interpolation, although unusual, may be very useful. For example, an optimal way to infill a missing value in a time series at a specific time would be to include in \mathbf{x} measurements taken in neighbouring gauges at this specific time, as well as measurements taken at the point of interest at preceding and subsequent times.

Let us first examine the different methods in which the estimation of y is based on a single observation $x \equiv x_i$ at one neighbouring (in space or time) point only. Here is a list of options, in which the following notation has been used: $\mu_x := E[\underline{x}]$ and $\mu_y := E[\underline{y}]$ are the expected values of \underline{x} and \underline{y} , respectively; $\sigma_x^2 := E[(\underline{x} - \mu_x)^2]$ and $\sigma_y^2 := E[(\underline{y} - \mu_y)^2]$ are the variances of \underline{x} and \underline{y} , respectively; $\sigma_{xy} := E[(\underline{x} - \mu_x)(\underline{y} - \mu_y)]$ is the covariance of \underline{x} and \underline{y} ; and $\rho_{xy} := \sigma_{xy}/(\sigma_x \sigma_y)$ is the correlation of \underline{x} and \underline{y} .

1. Equality: $y = x$. The single point of observation considered in this naïve type of interpolation is the station i nearest to the interpolation point, with $x \equiv x_i$. As discussed below, this simple interpolation forms the background of the Thiessen method of spatial integration. It is generally biased, with bias $\mu_e = \mu_y - \mu_x$ and its MSE is $\sigma_y^2 + \sigma_x^2 - 2 \sigma_{xy} + \mu_e^2$. However, in the case that the precipitation field is stationary (so that the means and variances at all points are equal to global parameters μ and σ^2 , respectively), it becomes unbiased, with $\text{MSE} = 2\sigma^2(1 - \rho_{xy})$. Evidently, for $\rho_{xy} < 0.5$, the method results in $\text{MSE} > \sigma^2$ and therefore there is no meaning in adopting it for low correlation coefficients (an estimate $x = \mu$ would be more skilful).
2. Normal ratio: $y = w x$ with $w = \mu_y / \mu_x$. This is a better alternative to the equality case, but it requires a sample of measurements to be available for \underline{y} in order to estimate the average μ_y . This estimation is unbiased ($\mu_e = 0$) but not best ($\text{MSE} = \sigma_y^2 + \sigma_x^2 \mu_y^2 / \mu_x^2 - 2 \sigma_{xy} \mu_y / \mu_x$).
3. Homogenous linear regression: $y = w x$ with $w = E[\underline{y}\underline{x}]/E[\underline{x}^2] = (\sigma_{xy} + \mu_x \mu_y) / (\sigma_x^2 + \mu_x^2)$. This is a biased estimation ($\mu_e = \mu_y - w \mu_x$) albeit best ($\text{MSE} = \sigma_y^2 + (\mu_y^2 \sigma_x^2 - 2 \mu_x \mu_y \sigma_{xy} - \sigma_{xy}^2) / (\mu_x^2 + \sigma_x^2)$).
4. Linear regression: $y = w x + b$ with $w = \text{Cov}[\underline{y}\underline{x}]/\text{Var}[\underline{x}] = \sigma_{xy} / \sigma_x^2$ and $b = \mu_y - w \mu_x$. This can be derived from equation (28) by adding an auxiliary variable whose values are always 1 (i.e. $y = w x + b \cdot 1$). It has the properties of being both unbiased and best, with $\text{MSE} = \sigma_y^2(1 - \rho_{xy}^2)$. However, it has the deficiency of potentially resulting in negative values, if $b < 0$, or of excluding values between 0 and b if $b > 0$. Another drawback emerges when many values of y are estimated in an attempt to extend a record of y based on a longer record of x . In this case, the resulting extended record has negatively biased variance, because the method does not preserve variance. To remedy this, a random error e should be added (using the probability distribution of e), which however is not determined in a unique manner and makes the method no longer best.
5. Organic correlation: $y = w x + b$ with $w = \text{sign}[\rho_{xy}] \sigma_y / \sigma_x$ and $b = \mu_y - w \mu_x$. This preserves both mean (i.e. it is unbiased) and variance, but it is not best ($\text{MSE} = 2\sigma_y^2(1 - |\rho_{xy}|)$). Evidently, for $|\rho_{xy}| < 0.5$, the method results in $\text{MSE} > \sigma_y^2$ and therefore adopting it is pointless for low correlation coefficients. Similar to the standard linear regression, the organic correlation retains the deficiency of producing negative values or excluding some positive values.

Coming to the interpolation based on multiple x_i , in the simplest case all weights w_i are assumed equal for all i , i.e., $w_i = 1/n$ so that y is none other than the average of x_i (the *arithmetic mean* method). This simple version is used very often to fill in sparse missing values of raingauge records. The quantities x_i could be simultaneous measurements at neighbouring points (say, within a radius of hundred kilometres), or at neighbouring times, or both. Here, neighbouring times should not necessarily be interpreted in the literal meaning, but with an emphasis on similarity of states. For example, a missing value of monthly precipitation in April 2000 could be estimated by, say, the average of the precipitation of the April months of 1998, 1999, 2001 and 2002. In another version,

the average of all April months with available data are used, but a local average (as we have already discussed in section 1.4) is preferable over an overall average, assuming that precipitation behaves like an HK process rather than a purely random one; this is similar to taking the average of points within a certain distance rather than a global average in spatial interpolation. This is not only intuitive but it can have a theoretical justification (D. Koutsoyiannis, personal notes), according to which for an HK process with $H = 0.7$, a local average based on 3 time steps before and 3 after the interpolation time is optimal (produces lowest MSE); the optimal number of points becomes 2+2 and 1+1 for $H = 0.75$ and $H \geq 0.8$, respectively.

This simple method does not impose any requirement for calculation of statistical quantities for its application. Another method of this type, which takes account of the geographical locations and, in particular, the distances d_i between the interpolated stations, is the method of *inverse distance weighting* (IDW). In each of the basis stations it assigns weights as

$$w_i = \frac{d_i^{-b}}{\sum_{j=1}^k d_j^{-b}} \quad (30)$$

where the constant b is typically assumed to be 2.

Among methods whose application requires statistical quantities to be known, the simplest is a direct extension of the normal ratio method, in which $w_i = (1/n) (\mu_y/\mu_x)$. The BLUE method itself belongs to this type. Initially, we can observe that a simple but biased solution for \mathbf{w} in equation (29) can be easily obtained as

$$\mathbf{w} = \mathbf{C}^{-1} \boldsymbol{\eta}, \quad \mu_e = \mu_y - \mathbf{w}^T \boldsymbol{\mu}_x, \quad \sigma_e^2 = \sigma_y^2 - \boldsymbol{\eta}^T \mathbf{C}^{-1} \boldsymbol{\eta} = \sigma_y^2 - \mathbf{w}^T \boldsymbol{\eta} \quad (31)$$

where $\boldsymbol{\eta} := \text{Cov}[\underline{y}, \underline{x}]$ is the vector whose elements are the covariances of \underline{y} with \underline{x} (see section 1.5) and $\mathbf{C} := \text{Cov}[\underline{x}, \underline{x}]$ is the positive definite symmetric matrix whose elements are the covariances of the vector \underline{x} with itself. One way to make it unbiased is to add an auxiliary variable x_{n+1} whose values are always 1. This is the multivariate extension of the typical linear regression described in point 4 above, and thus it retains the deficiency of potentially producing negative values or excluding some positive values. A better way to make it unbiased is to add a constraint $\mu_y = \mathbf{w}^T \boldsymbol{\mu}_x$ (the bivariate analogue of this is the equality case, described in point 1 above). In the latter case, the MSE becomes

$$\text{MSE} = \sigma_e^2 = \sigma_y^2 + \mu_y^2 + \mathbf{w}^T (\mathbf{C} + \boldsymbol{\mu}_x \boldsymbol{\mu}_x^T) \mathbf{w} - 2 \mathbf{w}^T (\boldsymbol{\eta} + \mu_y \boldsymbol{\mu}_x) \quad (32)$$

Minimization of the MSE with the above constraint using a Lagrange multiplier -2λ , results in the system of equations

$$\mathbf{C} \mathbf{w} + \boldsymbol{\mu}_x \lambda = \boldsymbol{\eta}, \quad \boldsymbol{\mu}_x^T \mathbf{w} = \mu_y \quad (33)$$

whose solution for the $n + 1$ unknowns w_1, \dots, w_n, λ is

$$\mathbf{w}' = \mathbf{C}'^{-1} \boldsymbol{\eta}' \quad (34)$$

where

$$\mathbf{w}' := \begin{bmatrix} \mathbf{w} \\ \lambda \end{bmatrix}, \quad \mathbf{C}' := \begin{bmatrix} \mathbf{C} & \boldsymbol{\mu}_x \\ \boldsymbol{\mu}_x^T & 0 \end{bmatrix}, \quad \boldsymbol{\eta}' := \begin{bmatrix} \boldsymbol{\eta} \\ \mu_y \end{bmatrix} \quad (35)$$

The value of the error is then calculated as

$$\text{MSE} = \sigma_e^2 = \sigma_y^2 + \mathbf{w}'^T \mathbf{C}' \mathbf{w}' - 2 \mathbf{w}'^T \boldsymbol{\eta}' \quad (36)$$

As seen in equations (31) and (34), the application of the method requires a number of covariances to be estimated (specifically, this number is $(n^2 + 3n)/2$, given that \mathbf{C} is symmetric). Not only does this restrict the method's application to points where measurements exist, in order to estimate the covariances, but, when n is large, it is infeasible to reliably estimate so many parameters from data and to derive a positive definite \mathbf{C} . The viable alternative is to assume a parametric stochastic model for the precipitation field. In the simplest case, the field could be assumed stationary and isotropic, where $\mu_{x_i} = \mu_y = \mu$, $\sigma_{x_i} = \sigma_y = \sigma$ and the covariance among any two points i, j is a function f of the geographical distance d_{ij} between these points, i.e., $\sigma_{ij} := \text{Cov}[x_i, x_j] = f(d_{ij})$. In this case equation (35) simplifies to

$$\mathbf{C}' := \begin{bmatrix} \mathbf{C} & \mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix}, \quad \mathbf{w}' := \begin{bmatrix} \mathbf{w} \\ \lambda' \end{bmatrix}, \quad \boldsymbol{\eta}' := \begin{bmatrix} \boldsymbol{\eta} \\ 1 \end{bmatrix} \quad (37)$$

where $\lambda' = \lambda \mu$ and $\mathbf{1}$ is a vector with all its elements equal to 1. The last solution is widely known as *kriging* (although kriging is sometimes formulated not in terms of covariance as here, but in terms of the so-called semivariogram, a notion that is not appropriate for processes with HK behaviour). We can observe from equation (37) that the solution is now independent of μ , as is also the error, which is still calculated from equation (36). It only depends on the covariance function $f(d)$. A function $f(d)$ compatible with the HK behaviour of precipitation, as discussed in section 1.5, is of the form

$$f(d) = \min(c, \alpha d^{4H-4}) \quad (38)$$

where H is the Hurst coefficient and $c \gg 0$ and α are parameters; in particular, c violates theoretical consistency but has been introduced to avoid problems related to the infinite covariance for distance tending to zero.

It can be observed that if the point of interpolation coincides with any one of the basis points i , then $\boldsymbol{\eta}$ is identical to one of the columns of \mathbf{C} and $\boldsymbol{\eta}'$ is identical to one of the columns of \mathbf{C}' . Thus, given the symmetry of \mathbf{C} and \mathbf{C}' , from equation (31) or (34) we obtain that \mathbf{w} is a unit vector, i.e., all elements are zero except one, which will be equal to 1. This shows the consistency of the method, i.e. its property to reproduce the measurements at gauged points with zero error.

All of the above methods that can interpolate at an arbitrary point (rather than only at a gauged one) provide a basis for numerical integration to find the average precipitation over a specific area A . Eventually, these methods result again in equations (28) or (29), where now y is the areal average precipitation. In particular, in the arithmetic mean and the normal ratio methods, because they do not make any assumption about the position of the point to which interpolation refers, the estimate y is an interpolation at any point and a spatial average, as well. The equality method works as follows: The geographical area of interest is divided into polygons, the so called Thiessen polygons,

each of which contains the points nearest to each of the stations. All points belonging to a specified polygon are regarded to have received a precipitation amount equal to that of the station corresponding to this polygon. Thus, in the integration we use either equation (28) or (29), where all gauged x_i in the area are considered with weights $w_i = A_i/A$, whereas A_i and A are the areas of the polygon corresponding to x_i . The remaining methods (IDW, BLUE), can be explicitly put in the form of equation (29), but this is rather tedious if done analytically. A simpler alternative is to make interpolations to many points, for example on a dense square grid. In turn, the gridded interpolations could be used for integration using equal weights for all grid points (i.e. arithmetic mean).

3.2 Radar estimates of precipitation

Radio detection and ranging (RADAR) was developed at the beginning of World War II as a remote sensing technique to measure the range and bearing of distant objects (such as ships and airplanes) by means of radio echoes (e.g. Battan, 1973). Since the early 1970s, radar techniques have also been used for the identification (i.e. shape, size, motion and thermodynamic properties) of precipitation particles. The latter are weather related distributed targets, which contrary to ships and airplanes, have characteristics that evolve in time and depend on the atmospheric conditions.

Because of their ability to provide estimates of areal precipitation quickly (i.e. at time intervals of about 5-15 min), at high resolutions (i.e. down to spatial scales of about 1 km) and over wide areas (i.e. with an effective range of about 200-400 km), radars have found wide application in atmospheric research, weather observation and forecasting (e.g. Atlas *et al.*, 1984; Doviak and Zrnica, 1993; Uijlenhoet, 1999, 2008; Bringi and Chandrasekar, 2001; Krajewski and Smith, 2002; Testik and Barros, 2007). An example is the NEXRAD (Next Generation Weather Radar) network with 159 operational WSR-88D (Weather Surveillance Radar 88 Doppler) units (as of February 2009), deployed throughout the continental United States and at selected locations overseas. According to NOAA's weather service (US National Oceanic and Atmospheric Administration, 2009), since its establishment in 1988, the NEXRAD project has provided significant improvements in severe weather and flash flood warnings, air traffic control and management of natural resources.

3.2.1 Basics of radar observation and measurement

A typical weather radar has 3 main components (Battan, 1973): (1) the *transmitter*, which generates short pulses of energy in the microwave frequency portion of the electromagnetic spectrum, (2) the *antenna*, which focuses the transmitted energy into a narrow beam and (3) the *receiver*, which receives the back-scattered radiation from distant targets that intercept the transmitted pulses.

Some important parameters, and their range of values, that characterize the radar equipment are (Rogers and Yau, 1996): (1) the instantaneous power of the pulse $P_t \approx 10\text{-}10^3$ kW (also referred to as *peak power*), (2) the duration of the pulse $\tau \approx 0.1\text{-}5$ μs , (3) the frequency of the signal $\nu \approx 3\text{-}30$ GHz, (4) the pulse repetition frequency (PRF) $f_r \approx 200\text{-}2000$ Hz, defined as the reciprocal of the time interval t_{max} that separates two distinct pulses (i.e. $t_{\text{max}} = f_r^{-1} \approx 0.5\text{-}5$ ms), and (5) the beamwidth of the antenna $\vartheta \approx 1^\circ$, defined as the angular separation between points where the power of the transmitted signal is reduced to half of its maximum value (or equivalently 3 dB below the maximum). The latter is attained at the beam axis.

The wavelength λ of the signal is defined as the distance between two sequential crests (or troughs) of the electromagnetic wave and it is related to its frequency as:

$$\lambda \nu = c \quad (39)$$

where $c = 3 \times 10^8$ m/s is the velocity of light in a vacuum. It follows from equation (39) that typical frequencies $\nu = 3\text{--}30$ GHz correspond to wavelengths λ between 10 and 1 cm, but most weather radars operate at wavelengths $\lambda = 3 - 10$ cm (X-, C-, and S-band; see e.g. Uijlenhoet and Berne, 2008). Shorter wavelengths are more effectively attenuated by atmospheric hydrometeors and precipitation particles (hence the transmitted signal has small effective range), whereas for longer wavelengths the back-scattered radiation from the precipitation particles does not have sufficient power to be detected by the receiver without noise induced by ground targets (e.g. Uijlenhoet, 2008).

When conducting radar observations and measurements, the direction of the target is obtained from the azimuth and elevation of the antenna when the returning echo is received. The range r of the target is calculated from the relation

$$r = c t / 2 \quad (40)$$

where t is the time interval between the transmission of the pulse and the reception of the echo. If the target is moving, the radial velocity u_r of the target (i.e. in the radar-pointing direction) can be obtained from the frequency shift $\Delta\nu$ of the received relative to the transmitted signal. The frequency shift is caused by the Doppler effect and it is related to u_r as:

$$\Delta\nu = -2u_r/\lambda \quad (41)$$

with positive $\Delta\nu$ being associated with targets that move toward the radar.

If t (the time interval between transmission and reception) is larger than t_{\max} (the reciprocal of the pulse repetition frequency, f_r) the echo from the target will reach the receiver after a new pulse has been transmitted. Hence, targets that return enough energy to be detected by the receiver (see below) and are located at distances $r > r_{\max} = c/(2f_r)$, will appear unrealistically close to the antenna. Thus, r_{\max} is the maximum range within which targets are indicated correctly on the radar screen and it is usually referred to as the *unambiguous range* (Battan, 1973; Rogers and Yau, 1996).

The visibility of a target by the radar depends on whether the returning signal has sufficient power P_r to be detected by the receiver. As an example, we consider a point target (i.e. a target with linear dimension smaller than about 10% of λ) with cross-section A_t located at distance r from the radar. We suppose that the radar transmits pulses with peak power P_t that propagate isotropically in space (i.e. in a 3-dimensional sphere). It follows from simple geometric considerations that the power P_i intercepted by the target is:

$$P_i = \frac{P_t A_t}{4\pi r^2} \quad (42)$$

where $4\pi r^2$ is the surface area of a sphere with radius r . If the transmitted signal is focused in a narrow beam by the antenna (as is commonly the case), equation (42) becomes

$$P_i = G \frac{P_t A_t}{4\pi r^2} \quad (43)$$

where $G = (4\pi A_e)/\lambda^2$ is a dimensionless constant called the *antenna axial gain* that depends on the characteristics (i.e. the wavelength λ) of the signal and the aperture A_e of the antenna.

Assuming that the target scatters the intercepted signal isotropically in space, the power P_r that reaches the radar is:

$$P_r = \frac{P_i A_e}{4\pi r^2} = G \frac{P_t A_t A_e}{(4\pi r^2)^2} = P_t A_t \frac{\lambda^2 G^2}{(4\pi)^3 r^4} \quad (44)$$

If the power P_r is large enough to be detected by the receiver without unwanted echoes (e.g. noise from ground targets) the target is visible to the radar and it is indicated on the radar screen.

For non-isotropic scatterers, the cross-section of the target A_t should be replaced by the back-scatter cross-section σ of the target. For spherical particles with diameter $D < \lambda/10$, usually referred to as *Rayleigh scatterers*, σ can be calculated from the relation (Battan, 1973):

$$\sigma = \frac{\pi^5 |K|^2 D^6}{\lambda^4} \quad (45)$$

where $|K|$ is the amplitude of the complex refractive index ($|K|^2 \approx 0.93$ for liquid water and 0.21 for ice), which characterizes the absorptive and refractive properties of the spherical scatterer. Due to the much higher value of $|K|^2$ for liquid water relative to ice (about 4.5 times higher), the melting layer of ice particles in precipitation generating weather systems appears on the radar screen as a bright band of high reflectivity.

3.2.2 Radar observation of distributed targets and the estimation of precipitation

For a typical weather radar that operates in the C-band portion of the electromagnetic spectrum ($\lambda = 3.75\text{-}7.5$ cm), raindrops and snowflakes (i.e. particles with effective diameters $D < 5\text{-}6$ mm) can be approximated as Rayleigh spherical scatterers with back-scatter cross-section σ given by equation (45). However, there are reasons why atmospheric hydrometeors should not be treated as isolated point targets. One reason is that the pulse transmitted by the radar illuminates simultaneously numerous precipitation particles that are included in a certain volume of air V , referred to as the *resolution volume* of the radar. Hence, the returned signal contains spatially averaged information from the whole population of raindrops and snowflakes in V .

For parabolic antennas, where the beam pattern is approximately the same in all directions, an accurate estimate of V can be obtained by assuming that the resolution volume is a cylinder with effective height equal to half of the pulse length $l = c\tau$ and diameter $d_v = r\vartheta$, i.e. the separation distance between points where the power of the transmitted signal is reduced to half of its maximum value. This gives:

$$V = \pi \left(\frac{r\vartheta}{2}\right)^2 \frac{c\tau}{2} \quad (46)$$

where ϑ is in radians. Equation (46) assumes that all energy in the radar transmitted pulse is contained within the half-power beamwidth; assuming a Gaussian shape of the beam pattern, the denominator of (46) (and, likewise, that of (49) below) should be multiplied by a factor $2 \ln 2$ (Probert-Jones, 1962).

Another reason why raindrops and snowflakes cannot be treated as isolated point targets is their turbulent motion that causes the power P_r of the returned signal to fluctuate in time. To this extent, an accurate approximation of the time averaged power \bar{P}_r (over a sufficiently long interval of about 10^{-2} s), which accounts also for multiple back-scattering cross-sections, is given by (Rogers and Yau, 1996):

$$\bar{P}_r = P_t \frac{\lambda^2 G^2}{(4\pi)^3 r^4} \sum_V \sigma \quad (47)$$

where r is the time-averaged range of the resolution volume V , and the summation is taken over all σ in V . For Rayleigh scatterers, equations (45) and (47) are combined to give

$$\bar{P}_r = P_t \frac{\pi^2 G^2 |K|^2}{64 r^4 \lambda^2} \sum_V D^6 \quad (48)$$

Assuming homogeneity of the population of hydrometeors in V , equation (48) can be written as

$$\bar{P}_r = P_t \frac{\pi^2 G^2 |K|^2}{64 r^4 \lambda^2} V \int_0^\infty n(D) D^6 dD = P_t \frac{\pi^3 G^2 |K|^2 \vartheta^2 c\tau}{512 r^2 \lambda^2} Z = C \frac{|K|^2 Z}{r^2} \quad (49)$$

where $n(D)$ is the size distribution of precipitation particles in V (i.e. number of particles per unit diameter and per unit volume of air), C is the so-called *radar constant* that depends solely on the characteristics of the system under consideration, and

$$Z := \int_0^\infty n(D) D^6 dD \quad (50)$$

is the *reflectivity factor* with units $[\text{Length}^3]$ that depends solely on the size distribution of the precipitation particles. For the Marshall and Palmer (1948) parameterization described by equation (24), equation (50) receives the form

$$Z = 720 n_0 b^{-7} \quad (51)$$

where n_0 and b are the intercept and scale parameters of the exponential size distribution. For the expressions given in section 2.3.2 for rain and snow we obtain:

$$\begin{aligned} \text{(a) } Z &= 296 i^{1.47} \quad (\text{rain}) \\ \text{(b) } Z &= 3902 i^{2.49} \quad (\text{snow}) \end{aligned} \quad (52)$$

where Z has units of mm^6/m^3 and i is the rainfall intensity (or the water equivalent of the accumulated snow at ground level) in mm/h .

For rain, equation (52.a) is very close to the empirical Z - i relationships (usually referred to as Z - R relationships, where $R \equiv i$ denotes the rainfall intensity) found in the literature (e.g. Marshall *et al.*, 1955; Battan, 1973; Uijlenhoet 1999, 2001, 2008), whereas for snow there is more variability and equation (52.b) should be seen only as an approximation. When combined, equations (40), (49) and (52) allow conversion of radar measurements (i.e. \bar{P}_r , t and r) to precipitation intensity i .

3.3 Spaceborne estimates of precipitation

The history of observation of Earth from space started on 4 October 1957, when the Soviet Union successfully launched Sputnik-I, the first artificial satellite. Sputnik-I provided information on the density of the highest layers of the atmosphere and on the radio-signal distribution in the ionosphere. The first launch was immediately followed by the launch of Sputnik-II by the Soviet Union on 3 November 1957 and the launches of Explorer-I (1 February 1958), Vanguard-I (17 March 1958), Vanguard-II (17 February 1959) and TIROS-I (1 April 1960) by the United States of America. The success of TIROS-I in surveying atmospheric conditions (in particular the cloud coverage of earth) opened a new era for meteorological research and development using spaceborne observations.

Since the 1970s, meteorological satellites have become essential in studying the development and evolution of weather related phenomena over the 71% of the Earth's surface covered by sea, where other types of measurements are unavailable. For example, the Tropical Rainfall Measuring Mission (TRMM; Simpson *et al.*, 1988; Kummerow *et al.*, 1998), which started on November, 1997 by the National Aeronautics and Space Administration (NASA) of the United States and the National Space Development Agency (NASDA) of Japan, has provided vast amounts of rainfall and energy estimates in tropical and subtropical regions and advanced the understanding and modelling of extreme rainfall events caused by tropical cyclones (e.g. Lonfat *et al.*, 2004, 2007; Chen *et al.*, 2006, 2007; Langousis and Veneziano, 2009a,b). TRMM data have also been used to improve the accuracy of high resolution weather forecasts produced by limited area models (e.g. Lagouvardos and Kotroni, 2005) and to investigate the relationship between lightning activity, microwave brightness temperatures (see below) and spaceborne radar reflectivity profiles (Katsanos *et al.*, 2007).

We can distinguish two types of sensing by satellites, *passive* and *active*. Passive sensing is based on measuring the radiative intensity emitted or reflected by particles in the atmosphere, such as cloud droplets and hydrometeors of precipitable size. Active sensing is conducted using radar equipment carried by the satellite. Next we discuss some basic principles of passive remote sensing in the visible (V , $\lambda \approx 0.39\text{-}0.77 \mu\text{m}$), infrared (IR; wavelengths $\lambda \approx 0.77 \mu\text{m} - 0.1 \text{ mm}$), and microwave (MW, $\lambda \approx 0.1 \text{ mm} - 10 \text{ cm}$) portions (channels) of the electromagnetic spectrum. The basic principles of operation of active sensors are similar to those of radars and were reviewed in section 3.2. For a more detailed review on the principles and techniques of remote sensing, the reader is referred to Barrett and Martin (1981), Elachi (1987), Stephens (1994) and Kidder and Vonder Haar (1995).

3.3.1 The infrared signature of cloud tops

The high absorptivity of cloud droplets in the infrared spectral range causes clouds to appear opaque in the IR channel. Hence, the infrared radiation received by the satellite's radiometer originates mostly from the cloud tops, which can be accurately approximated as *black bodies*, that is, as objects that absorb all incident radiation and emit it at a rate that depends solely on their temperature. In this case, we can use Stefan-Boltzman's law of radiation (e.g. Barrett and Martin, 1981) to calculate the temperature T_b of the cloud tops from the intensity J of the received IR radiation:

$$T_b = (J/\sigma_{SB})^{1/4} \quad (53)$$

where $\sigma_{SB} = 5.7 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$ is the Stefan-Boltzman constant and T_b is in kelvins. T_b is usually referred to as *brightness temperature* (e.g. Smith, 1993) and, for a given atmospheric lapse rate γ (see section 2.1), it can be used to calculate cloud top heights.

Evidently, lower brightness temperatures T_b correspond to clouds with higher tops and larger probabilities of rain. Hence, we can develop regression equations to relate brightness temperatures to observed surface rainfall rates (e.g. Griffith *et al.*, 1978; Stout *et al.*, 1979; Arkin, 1979; Richards and Arkin, 1981; Arkin and Meisner, 1987; Adler and Negri, 1988). Two important limitations apply (Richards and Arkin, 1981; Liu, 2003): (1) due to the statistical character of the regressed quantities, the accuracy of the rainfall retrieval algorithm increases with increasing scale of spatial or temporal averaging, and (2) the parameters of the regression depend on the climatology of the region and, therefore, cannot be used at regions with different climatic characteristics.

An example of surface rainfall estimation from IR images is the *temperature threshold method* developed by Arkin (1979), Richards and Arkin (1981) and Arkin and Meisner (1987). Arkin (1979) used IR imagery from the Synchronous Meteorological Satellite-1 (SMS-1) and radar data from Global Atmosphere Research Program (GARP) Atlantic Tropical Experiment (GATE) to investigate the correlation between radar-estimated precipitation rates and the fraction of areas with brightness temperature T_b below a certain threshold T_{min} . The study found a maximum correlation (around 0.85) for a brightness temperature threshold $T_{min} \approx 235 \text{ K}$ (-38°C). Richards and Arkin (1981) showed that a linear relationship suffices to describe the dependence between spatially averaged surface rainfall and the fraction of areas with $T_b < 235 \text{ K}$, with error variance that increases with decreasing scale of spatial averaging. Based on these results, Arkin and Meisner (1987) suggested the use of the GOES (Geostationary Operational Environmental Satellite) Precipitation Index (GPI) to calculate spatial rainfall averages in the tropics:

$$\text{GPI} = 3 \text{ (mm/h)} F_c H \quad (54)$$

where GPI is the spatially averaged rainfall accumulation in a grid box of 2.5° latitude \times 2.5° longitude, F_c is the mean fraction (a dimensionless quantity between 0 and 1) of the grid box covered by brightness temperatures $T_b < 235 \text{ K}$ and H is the length of the observation period in hours.

The temperature threshold method of Arkin (1979), Richards and Arkin (1981) and Arkin and Meisner (1987) produces accurate estimates of the spatially averaged rainfall in the tropical belt (30°S to 30°N), at grid scales larger than 2.5° ($\approx 275 \text{ km}$) (Arkin and Meisner, 1987) and for averaging durations greater than about a month (Ba and Nicholson, 1998). The error increases significantly as

we move to mid-latitudes, especially during cold seasons (e.g. Liu, 2003). Extensions of the method include the use of the upper tropospheric humidity (UTH) in the vicinity of convective clouds as an additional predictive variable (Turpeinen *et al.*, 1987), and the combination of IR and visible imagery (i.e. bi-spectral methods; see below) to exclude non-precipitating clouds with high tops.

3.3.2 The visible reflectivity of clouds

The signature of Earth in the visible (V) channel is due to the reflection of the sunlight by clouds and, when the sky is clear, surface features. Consequently, visible imagery is available only during daylight hours. Due to its shorter wavelength, visible radiation can penetrate deeper into clouds than the infrared portion of the electromagnetic spectrum, but similar to the infrared channel, it still represents the upper portion of clouds and serves as an indirect signature of surface rainfall. However, visible reflectivity can complement the infrared brightness temperatures to allow better classification of clouds and qualitative assessment of the probability of precipitation. This is the basis of the well known *bi-spectral* methods (e.g. Lovejoy and Austin, 1979; Bellon *et al.*, 1980; Tsonis and Isaac, 1985; Tsonis, 1987; O' Sullivan *et al.*, 1990; Cheng *et al.*, 1993; Cheng and Brown, 1995; King *et al.*, 1995; Liu, 2003).

The visible reflectivity of clouds increases fast with increasing liquid water path, i.e., the vertically integrated liquid water in the atmospheric column. Hence, we can use infrared brightness temperatures to calculate the altitude of the cloud tops and visible reflectivities to obtain a qualitative estimate of the vertically averaged liquid water of the cloud, which is indicative of the rainfall potential. For example, low brightness temperatures (i.e. cold cloud tops) and high visible reflectivities (i.e. thick clouds) indicate cumulonimbus formations with high probability of precipitation (see section 2.4.1), warm cloud tops and high visible reflectivities indicate stratiform rainfall (see section 2.4.2), whereas cold cloud tops and low visible reflectivities indicate cirrus clouds, which are usually non precipitating.

An example of bi-spectral methods is the RAINSAT technique developed by Lovejoy and Austin (1979) and Bellon *et al.* (1980). This technique uses visible reflectivities to reduce the number of false alarms obtained from the IR channel and more accurately estimate surface rainfall rates. The RAINSAT method was developed using GOES infrared and visible imagery and radar data from tropical (i.e. GATE) and mid-latitude (i.e. McGill weather radar, Quebec, Canada) locations as ground truth. The method was optimized by Cheng *et al.* (1993), Cheng and Brown (1995) for the area of the UK, using IR and visible imagery from the European geostationary satellites Meteosat-2, Meteosat-3 and Meteosat-4 and rainfall retrievals from 9 weather radars located in the United Kingdom and Ireland. A similar cloud classification technique has been proposed by Tsonis and Isaac (1985) and Tsonis (1987). This technique is based on cluster analysis of pixels with different brightness temperatures and visible reflectivities and has been developed using GOES satellite data and rainfall retrievals from the Woodbridge weather radar in Ontario, Canada.

3.3.3 The microwave signature of precipitation

Contrary to the infrared and visible spectral ranges, microwave radiation can effectively penetrate through cloud and rain layers and provide the signature of the integrated contribution of precipitation particles in the atmospheric column. Hence, brightness temperatures obtained from

the MW channel are better linked to surface rainfall rates than the visible reflectivities and IR brightness temperatures.

The type and size of the precipitation particles detected by the microwave radiometer depends on the frequency of the upwelling radiation. Above 80 GHz (i.e. wavelengths $\lambda < 3.75$ mm) ice crystals scatter the upwelling MW radiation and fade the signature of raindrops. Hence, above 80 GHz the radiometer senses only ice, where lower brightness temperatures are associated with more scattering, larger ice particles and higher precipitation intensities at ground level.

Below about 20 GHz (i.e. $\lambda > 1.5$ cm) the radiative intensity of raindrops dominates the microwave signature of hydrometeors in the atmospheric column, whereas ice particles are virtually transparent. Thus, below 20 GHz the microwave radiometer detects the vertically integrated signature of rain water, where higher brightness temperatures are associated with more intense rainfall at ground level. Low-frequency microwave imagery is especially useful when calculating surface rainfall rates over oceans, where the almost constant sea surface temperature and emissivity allow translation of the spatial and temporal variations of brightness temperatures to variations of sea-level rainfall rates (e.g. Liu, 2003). The same is not true over land, where the surface features cause the ground temperature and emissivity to vary significantly in space and time. Another limitation of low frequency microwave images is the saturation of the microwave channel at high rainfall rates, which causes negative biases of the obtained rainfall intensity (e.g. Liu, 2003; Viltard *et al.*, 2006).

Between 20 GHz and 80 GHz scattering and emission by raindrops and ice particles occur simultaneously and the microwave radiation undergoes multiple transformations. Hence, the microwave radiometer detects different rain paths at different microwave frequency ranges.

Combining brightness temperatures from different MW channels to more accurately assess surface rainfall rates is an open research problem and it has driven the development of many rainfall estimation algorithms (Grody, 1991; Spencer *et al.*, 1989; Alishouse *et al.*, 1990; Berg and Chase, 1992; Hinton *et al.*, 1992; Liu and Curry, 1992, 1993; Ferriday and Avery, 1994; Petty, 1994a,b, 2001a,b; Kummerow and Giglio, 1994a,b; Ferraro and Marks, 1995; Kummerow *et al.*, 1996, 2001; Berg *et al.*, 1998; Aonashi and Liu, 2000; Levizzani *et al.*, 2002). For a review of microwave methods of estimation over ocean and land and their advantages and limitations, the reader is referred to Wilheit *et al.*, (1994) and Petty (1995), and Kidd *et al.* (1998) respectively.

4 Precipitation modelling

As already clarified in section 1.5, modelling of precipitation is not possible without using any type of a stochastic approach. Even the deterministic numerical weather forecast models, which determine the state and motion in the atmosphere by solving differential equations, to model precipitation they use *parameterization schemes*. These schemes, instead of describing the detailed dynamics of the precipitation process, establish and use equations of statistical type to quantify the output of the dynamical system. In addition, as mentioned in section 1.5, the modern framework for predicting precipitation particularly as input to hydrological models (the ensemble forecasting) is of the Monte Carlo or stochastic type. The description of these stochastic techniques belongs to the sphere of weather forecasting and is out of the scope of this chapter. In more engineering-oriented applications, precipitation is typically modelled as an autonomous process, without particular

reference to the atmospheric dynamics. Next we outline some of the most widespread modelling practices for precipitation but without details and mathematical formulations, which the interested reader can find in the listed references.

4.1 Rainfall occurrence

From the early stages of the analysis of precipitation intermittency, it was recognized that the rainfall occurrences are not purely random. In other words, rainfall occurrence cannot be modeled (effectively) as a Bernoulli process in discrete time or, equivalently, as a Poisson process in continuous time. It should be recalled that in a Bernoulli process an event (rainfall/wet state) occurs with a probability p (and does not occur with probability $1 - p$) constant in time, and each event is independent of all preceding and subsequent events. In a Poisson process the times of occurrence of events (i.e. the starting times of rainfalls) are random points in time. In this process the time differences between consecutive occurrences are independent identically distributed (IID) with exponential distribution.

Both discrete time and continuous time representations of the rainfall occurrence process, which in fact are closely related (e.g. Foufoula-Georgiou and Lettenmaier, 1986; Small and Morgan, 1986), have been investigated. The most typical tool of the category of discrete time representations is the Markov chain model (Gabriel and Neumann, 1962; Feyerherm and Bark, 1964; Hershfield, 1970; Todorovic and Woolhiser, 1975; Haan *et al.*, 1976; Chin, 1977; Katz, 1977a, b; Kottegoda and Horder, 1980; Roldan and Woolhiser, 1982). In this model, any time interval (e.g. day) can be in one of two states, dry or wet, and it is assumed that the state in a time interval depends on the state in the previous interval.

It was observed, however, that Markov chain models yield unsatisfactory results for rainfall occurrences, especially for dry intervals (De Bruin, 1980). Moreover, the interannual variance of monthly (or seasonal) total precipitation is greater than predicted by Markov chain models, an effect usually referred to as “overdispersion” (Katz and Parlange, 1998). Extended versions of the binary state Markov chains using a higher number of past states may improve performance. Additional states in such model versions have been defined based on combination of states of two consecutive periods (Hutchinson, 1990) or on accounting for the rainfall depth of each interval (Haan *et al.*, 1976). A more effective enhancement is to use transition probabilities taking into account more than one previous interval, which leads to stochastic binary chains of order higher than one (Pegram, 1980; Katz and Parlange, 1998; Clarke, 1998). In more recent developments, to account for a long number of previous time intervals and simultaneously avoid an extremely high number of transition probabilities, it was proposed that, instead of the sequence of individual states of these intervals, one could use conditional probabilities based on aggregation of states of previous intervals (Sharma and O’Neill, 2002). Similarly, one could use a discrete wetness index based on the number of previous wet intervals (Harrold *et al.*, 2003). An extension of the Markov chain approach to multiple sites has been studied by Pegram and Seed (1998).

In a more recent study, Koutsoyiannis (2006a) used the principle of maximum entropy, interpreted as maximum uncertainty, to explain the observed dependence properties of the rainfall occurrence process, including the overdispersion or clustering behavior and persistence. He quantified intermittency by the probability $p^{(1)}$ that a time interval of length 1 h is dry, and dependence by the probability that two consecutive intervals are dry, i.e., by $p^{(2)}$, where in general $p^{(k)}$ denotes the

probability that an interval of length k is dry. Using these two probabilities and a multi-scale entropy maximization framework, he was able to determine any conditional or unconditional probability of any sequence of dry and wet intervals at any time scale. Thus, he described the rainfall occurrence process including its dependence structure at all scales using only two parameters. The dependence structure appeared to be non-Markovian yet not over-exponential. Application of this theoretical framework to the rainfall data set of Athens indicated good agreement of theoretical predictions and empirical data at the entire range of scales for which probabilities dry and wet can be estimated (from one hour to several months). An illustration is given in Fig. 22.

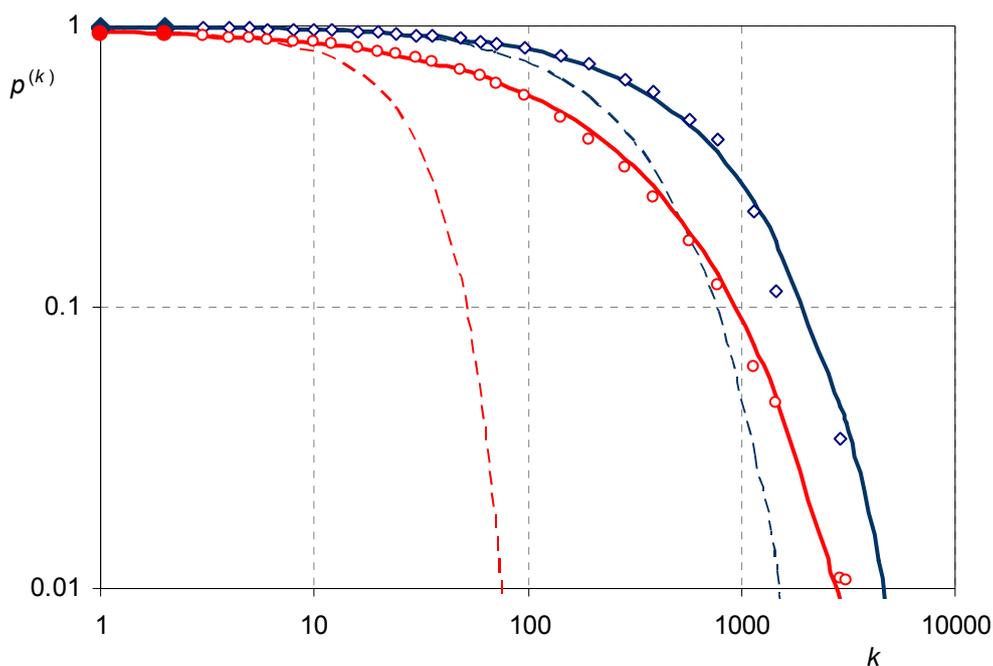


Fig. 22 Probability dry $p^{(k)}$ vs. scale k (in h), as estimated from a hourly rainfall data set in Athens, Greece, and predicted by the maximum entropy model in Koutsoyiannis (2006a) for the entire year (circles and red full line) and the dry season (June-September; diamonds and blue full line). The model was fitted using two data points in each case (marked in full in the plot), i.e. the probability dry for 1 h, $p \equiv p^{(1)}$, and 2 h, $p^{(2)}$, which are respectively 0.9440 and 0.9335 for the entire year and 0.9888 and 0.9860 for the dry season. The final model is expressed as $p^{(k)} = p [1 + (\xi^{-1/\eta} - 1)(k - 1)]^\eta$, where the parameters are respectively $\eta = 0.63$ and $\xi = 0.816$ for the entire year and $\eta = 0.83$ and $\xi = 0.801$ for the dry season. For comparison lines resulting from the Markov chain model are also plotted (dashed lines).

In the continuous time representation of the rainfall occurrence process, the dominant tools are the cluster-based point processes (Waymire and Gupta, 1981a,b,c). These are essentially based on the prototype of the spatial distribution of galaxies devised by Neyman and Scott (1952) to describe their property of “clustering” relative to the Poisson process. With reference to storms, if they were regarded as instantaneous pulses positioned at random points in time, the logarithm of probability that the interarrival time exceeds a value x , or the log survival function, would be proportional to x . However, empirical evidence suggests that the log survival function is a nonlinear concave function of x , which indicates a tendency for clustering of rainfall events relative to the Poisson model (Foufoula Georgiou and Lettenmaier, 1986). This clustering has been modeled by a cascade of two Poisson processes, corresponding to two characteristic time scales of arrivals of storms and storm cells.

The Neyman-Scott process with instantaneous pulses was the first applied to rainfall occurrence (Kavvas and Delleur, 1981; Rodriguez-Iturbe *et al.*, 1984), later succeeded by the Neyman-Scott rectangular pulses and the very similar Bartlett-Lewis rectangular pulse models (Rodriguez-Iturbe *et al.*, 1987). The Bartlett-Lewis rectangular pulse model, which is the most typical and successfully applied model of this type, assumes that rainfall occurs in the form of storms of certain durations and that each storm is a cluster of random cells. The general assumptions of the rainfall occurrence process are:

1. Storm origins \underline{t}_i occur according to a Poisson process with rate λ .
2. Origins \underline{t}_{ij} of cells of each storm i arrive according to a Poisson process with rate β .
3. Arrivals of each storm i terminate after a time \underline{v}_i exponentially distributed with parameter γ .
4. Each cell has a duration \underline{w}_{ij} exponentially distributed with parameter η .

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

A major problem of these models was their inability to reproduce the probability of zero rainfall at multiple time scales (Velghe *et al.*, 1994). In this respect, Foufoula-Georgiou and Guttorp (1986) noted that the Neyman-Scott model parameters are scale dependent and thus cannot be attributed a physical meaning. To ameliorate this, modifications of both the Neyman-Scott model (Entekhabi *et al.*, 1989) and the Bartlett-Lewis model (Rodriguez-Iturbe *et al.*, 1988; Onof and Wheater, 1993, 1994) were proposed. These are in fact based on the randomization of the mean interarrival time of one of the two Poisson processes. Evaluation and comparison of several cluster-based rectangular pulse models for rainfall were done by Velghe *et al.* (1994) and Verhoest *et al.* (1997), whereas a comprehensive review of Poisson-cluster models has been provided by Onof *et al.* (2000). An extension of the concept introducing a third Poisson process was proposed by Cowpertwait *et al.* (2007).

4.2 Rainfall quantity

In the discrete time representations of rainfall occurrence, the rainfall quantity in each wet interval is modelled separately from the occurrence process, usually based on statistical analysis of the observed record. In the point process representations, the storms and cells are abstract quantities that do not fully correspond to real-world objects. Therefore, they cannot be identified in the recorded time series. An assumption is typically made that each cell has a uniform intensity \underline{x}_{ij} with a specified distribution, and based on all assumptions, the statistical characteristics of the rainfall process at one or more time scales are derived analytically (Rodriguez-Iturbe *et al.*, 1987, 1988). These statistical characteristics are compared to the empirically derived statistics and, by minimizing the departures of the two, the model parameters are estimated. The distribution of the uniform intensity \underline{x}_{ij} is typically assumed to be exponential with parameter $1 / \mu_x$. Alternatively, one can choose a two-parameter gamma distribution with mean μ_x and standard deviation σ_x . In this manner, the point process models describe the entire rainfall process, including occurrence and

quantity. A demonstration of the model is shown in Fig. 23. However, in some cases (e.g. Gyasi-Agyei and Willgoose, 1997) point processes have been used to simulate merely rainfall occurrences and then have been combined with other models that simulate rainfall depths. Other modelling approaches for the rainfall process (including its intermittency) are reviewed in Srikanthan and McMahon (2001).

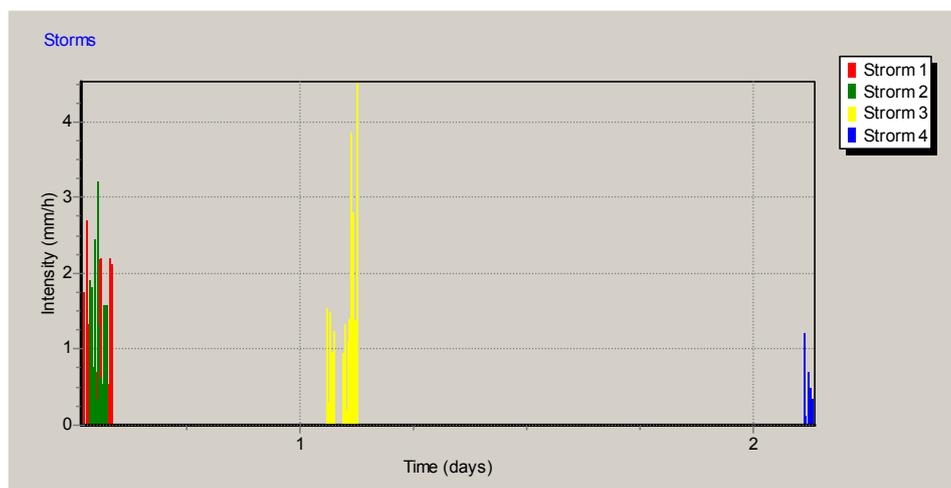


Fig. 23 Simulated realization of a series of four storms from the Bartlett-Lewis rectangular pulse model (modified version with randomly varying η) occurring within three days (notice the overlap of storms 1 and 2, which is allowed by the model), implemented by the Hyetos software (see section 4.4). The model parameters are $\lambda = 0.94 \text{ d}^{-1}$, $\kappa = \theta/\eta = 1.06$, $\varphi = \gamma/\eta = 0.059$, $\alpha = 2.70$, $\nu = 0.0068 \text{ d}^{-1}$, and $\mu_X = \sigma_X = 24.3 \text{ mm/d}$.

With their typical assumptions, including those of the exponential or gamma distribution for rain cell amount, the point process models, despite providing satisfactory representation of the process at a specific time scale or a small range of time scales, cannot really perform satisfactorily over a wide range of scales and also lead to exponential distribution tails, whereas it has been recently recognized that the tails must be of power type (see sections 1.5 and 5.2). Generally, the distribution function of rainfall varies among different time scales. At very fine scales, the density is J-shaped, i.e., with a mode at zero, and perhaps with density tending to infinity as the rainfall depth or the intensity tends to zero. At coarse time scales such as monthly (for wet months) and annual, the distribution becomes bell-shaped and tends to normal as the scale increases. However, its tail always departs from the exponential tail of the normal distribution. In fact, for theoretical reasons, if at the right tail the survival function is a power function of the rainfall depth or intensity x , with exponent $1/\kappa$, i.e., $F^*(x) \propto x^{-1/\kappa}$ (see equation (18)), then it will be of the same type and will have precisely the same exponent $1/\kappa$ at any time scale (the proof is omitted). This behaviour of the tail is perhaps the only invariant distributional property across all scales, whereas the shape of the body of the distribution varies significantly across different scales. However, even this variation must have a simple and unique explanation, which is the principle of maximum entropy. Specifically, Koutsoyiannis (2005a) has shown that all diverse shapes of the distribution across different scales can be derived from the principle of maximum entropy constrained on known mean and variance.

Papalexiou and Koutsoyiannis (2008) proposed a single distribution (a power-transformed beta prime distribution, also known as generalized beta of 2nd kind; see also Koutsoyiannis, 2005a) with four parameters, which provides good fits for rainfall intensity at time scales from hourly to annual. Only one of the four parameters (corresponding to the exponent of the tail) is invariant across

scales. If the range of scales of interest is smaller, then specific special cases of this distribution can be used as good approximations. For example the three-parameter Burr type VII distribution, which has the advantage of providing a closed form of the quantile function, can be used effectively for time scales from a few minutes to a couple of months (Papalexiou and Koutsoyiannis, 2009).

4.3 Space-time models

Space-time modelling of precipitation is one of the most demanding tasks of stochastic modelling in hydrology and geophysics. Rainfall intermittency should be modelled in both space and time, along with the motion of rainfall fields, the rainfall quantity and its temporal and spatial structure. One of the relatively simple solutions has been provided by the extension of point process models used for the rainfall process at a single site. This extension introduces a description of rainfall cells in space, in addition to that in time, and a motion of the cells. As an example, we summarize here the Gaussian displacement spatial-temporal rainfall model (GDSTM; Northrop, 1996, 1998). This model, is a spatial analogue of a point process model having a temporal structure similar to that of the Bartlett-Lewis rectangular pulse model described above and a spatial structure known as the Gaussian displacement structure, introduced by Cox and Isham (1988).

Similar to its single-site analogue, GDSTM assumes that rainfall is realized as a sequence of storms, each consisting of a number of cells. Both storms and cells are characterized by their centers, durations and areal extents (see sketch in Fig. 24) and, in addition, cells have certain uniform rainfall intensity. Specifically, the following assumptions characterize storms and cells.

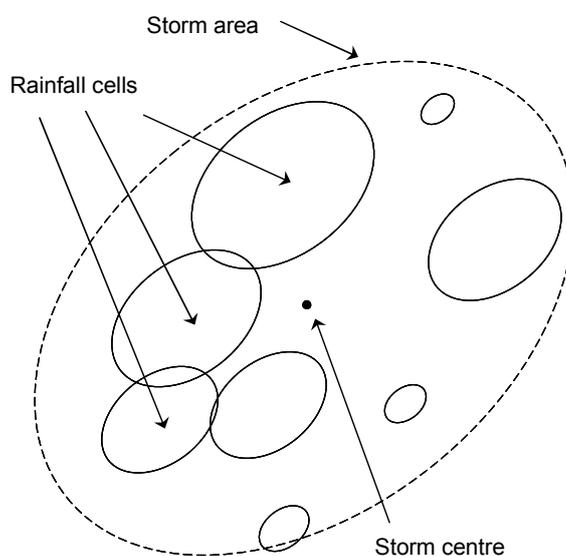


Fig. 24 Sketch of the spatial structure of the Gaussian displacement spatial-temporal rainfall model.

Storm centers arrive according to a homogeneous Poisson process of rate λ in two-dimensional space (denoted by x, y) and time (denoted by t) and move with a uniform velocity (V_x, V_y) . Each storm has a finite duration \underline{L} (assumed exponentially distributed with parameter $\beta = 1 / \mu_L$) and an infinite areal extent, represented by an elliptical geometry with eccentricity ε and orientation ϑ , and incorporates a certain number of rainfall cells. However, a storm can be assigned a finite “storm area”, the area that contains a certain percentage of rainfall cells. The storm area varies randomly and in each storm is determined in terms of the realization of a random variable \underline{w} , which

determines uniquely (for the specific storm) a set of parameters σ_x^2 , σ_y^2 and ρ that determine the displacement of cell centers from the storm center. Specifically, \underline{w} is Gamma-distributed with shape and scale parameters determined in terms of the eccentricity ε and the mean storm area μ_s . At the same time, the parameter ρ is determined in terms of the eccentricity ε and the storm orientation, ϑ . Following the generation of \underline{w} , the parameters σ_x^2 and σ_y^2 are determined in terms of the eccentricity ε , the storm orientation ϑ , and the value of \underline{w} .

Each rainfall cell is assigned a center $(\underline{x}_c, \underline{y}_c, \underline{t}_c)$. The time origin \underline{t}_c follows a Poisson process starting at the time ordinate of the storm origin \underline{t}_0 (with the first cell being located at this point) and ending at $\underline{t}_0 + \underline{L}$. The expected number of cells within that time interval is $\mu_c = 1 + \beta / \gamma$, where γ is the cell generation Poisson process parameter. The spatial displacements from the storm center are random variables jointly normally distributed with zero means, variances σ_x^2 and σ_y^2 , and correlation ρ . Given these parameters, the displacement $\Delta \underline{x}$ of each cell is generated as a normal variate $(0, \sigma_x)$ and the displacement $\Delta \underline{y}$ as a normal variate $(\mu_{y|x}, \sigma_{y|x})$. Furthermore, each cell has a finite duration \underline{D} (assumed exponentially distributed with parameter $1 / \mu_D$) and an elliptical area with major axis \underline{a} , forming an angle ϑ with the x axis (west-east), and minor axis $\underline{b} = \sqrt{1 - \varepsilon^2} \underline{a}$. It is assumed that \underline{a} is a random variable gamma distributed with shape and scale parameters depending on the mean storm area μ_A and the eccentricity ε , respectively. Finally, each cell has an intensity \underline{x} independent of any other variable, exponentially distributed with parameter $1 / \mu_x$.

The model is defined in terms of 11 independent parameters, namely: (1) the rate of storm arrival (number of storms per area per time), λ ; (2) the mean cell duration, μ_D ; (3) the mean storm duration, μ_L ; (4) the mean cell area, μ_A ; (5) the mean storm area, μ_s ; (6) the mean number of cells per storm, μ_c ; (7) the mean cell intensity, μ_x ; (8 and 9) the components of the cell and storm velocity in the x direction (east), V_x , and in the y direction (north), V_y ; (10) the cell and storm eccentricity, ε ; and (11) the cell and storm orientation, ϑ .

Similar to its single-point analogue, the entities of the spatial point process model are abstract. To make the model outputs comparable to reality, integration from continuous time over a specific time scale and/or spatial scale is needed, from which the first and second order rainfall statistics are calculated. The latter serve as the basis for parameter estimation using either rain gauge or radar data. Due to model complexity the calculation of the statistics can be done only numerically; hence the entire model application (and the parameter estimation in particular, which needs numerical optimization, e.g. using the generalized reduced gradient method) is laborious.

4.4 Rainfall disaggregation and downscaling

Both disaggregation and downscaling refer to the generation of a precipitation field at a specific temporal and/or spatial scale given a known precipitation field (measured or simulated) at a certain larger temporal and/or spatial scale (lower resolution). Disaggregation and downscaling are very useful procedures and have several applications, such as in the following cases.

1. Global-scale weather prediction models provide rainfall forecasts at a low resolution, e.g. grid size of 50 km. Hydrologic models require the description of the precipitation field at a much higher resolution, with grid size of the order of 1 km..

2. Satellite precipitation estimates are available at a spatial scale greater than or equal to 0.25° (latitude and longitude), or about 28 km at the equator, and a temporal scale of 3 h. Again hydrologic applications require higher resolutions.
3. The majority of historical point rainfall records come from daily raingauges which have often been operational for several decades. The number of raingauges providing hourly or sub-hourly resolution data is smaller by about an order of magnitude. However, hydrologic applications, especially flood studies, usually need hourly or even sub-hourly data.
4. In complex problems of stochastic generation of precipitation time series or precipitation fields, it is difficult to reproduce simultaneously the long-term and the short-term stochastic structure of precipitation using a single model. A better approach is to couple several models, starting from a large-scale model to represent the long-term behaviour. The outputs of the latter are then disaggregated into finer scales. Note, however, that in a recent study Langousis and Koutsoyiannis (2006) developed a stochastic framework capable of reproducing simultaneously the long-term and the short-term stochastic structure of hydrological processes, avoiding the use of disaggregation.

While disaggregation and downscaling are similar in nature, they also have a difference that distinguishes them. Downscaling aims at solely producing a precipitation field \underline{y} with the required statistics at the scale of interest, being statistically consistent with the given field \underline{x} at the finer scale. Disaggregation demands full and precise consistency, which introduces an equality constraint in the problem of the form

$$\underline{C} \underline{y} = \underline{x} \quad (55)$$

where \underline{C} is a matrix of coefficients. For example, assuming that \underline{x} is an annual amount of precipitation at a station and \underline{y} is the vector consisting of the twelve monthly precipitation values at the same station, \underline{C} will be a row vector with all its elements equal to 1, so that equation (55) represents the requirement that the sum of all monthly precipitation amounts must equal the annual amount.

Task 1 could be accomplished by running a second meteorological model at the limited area of interest. Such models, known as *limited-area models*, can have much higher resolution than global models. The description of this type of downscaling, known as *dynamical downscaling* because it is based on the atmospheric dynamics, is out of the scope of this chapter. In contrast, a stochastic procedure need not refer to the dynamics, and is generic and appropriate for both downscaling and disaggregation and for all above tasks 1-4. This generic procedure resembles the interpolation procedure described in section 3.1.3, but there are two important differences. First, it is necessary to include the error terms in the generation procedure (recall that in interpolation, which is a point estimation, knowing only the mean and variance of the error was sufficient). Second, the generated values \underline{y} at the different points should be statistically consistent to each other. This precludes the separate application of an algorithm at each point of interest and demands simultaneous generation at all points. In turn, this demands that the error terms in different points should be correlated to each other. All these requirements could be summarized in the linear generation scheme,

$$\underline{y} = \underline{A} \underline{x} + \underline{B} \underline{v} \quad (56)$$

where \mathbf{A} and \mathbf{B} are matrixes of coefficients and \mathbf{v} is a vector of independent random variables, so that the term $\mathbf{B} \mathbf{v} =: \mathbf{e}$ corresponds to the error term in interpolation (cf. (29)). In disaggregation, equation (56) should be considered simultaneously with equation (55).

For Gaussian random fields without intermittency, the application of equations (55) and (56) is rather trivial. However, the intermittency of the rainfall processes and the much skewed distributions at fine time scales are severe obstacles for rainfall disaggregation. To overcome such obstacles, several researchers have developed a plethora of rather *ad hoc* disaggregation models (see review by Koutsoyiannis, 2003b). However, the application of the above theoretically consistent scheme is still possible, if combined with a stochastic model accounting for intermittency (e.g. a Bartlett-Lewis model) and if an appropriate strategy is used to implement equation (55). Such a strategy includes recursive application of equation (56) until the error in equation (55) becomes relatively low, and is followed by correction of the error of the accepted final iteration by appropriate adjusting procedures, which should not alter the covariance structure of the precipitation field. The general strategy of stochastic disaggregation is described in Koutsoyiannis (2001) and two implementations for temporal rainfall disaggregation at a fine (hourly) scale at a single site and at multiple sites are described in Koutsoyiannis and Onof (2001) and Koutsoyiannis *et al.* (2003), respectively. The models described in the latter two papers, named Hyetos and MuDRain, respectively, are available on line (www.itia.ntua.gr/en/software/) and have been used in several applications worldwide. Typical results of the two models are shown in Fig. 25 and Fig. 26, respectively.

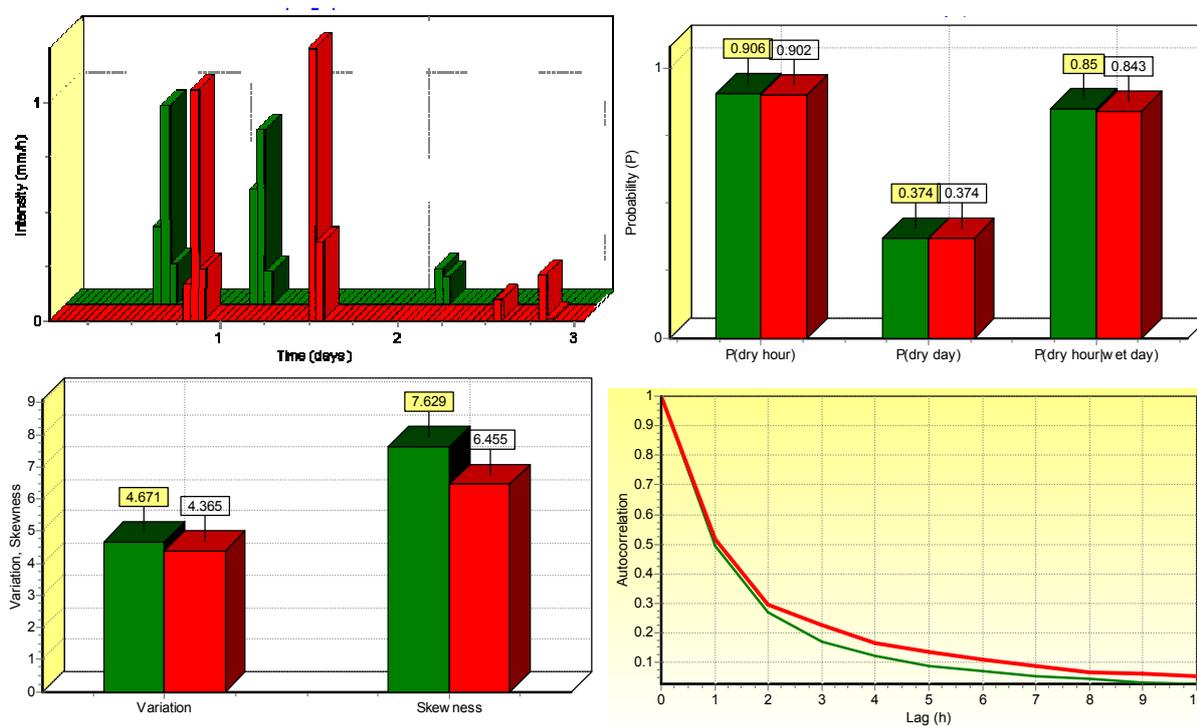


Fig. 25 Typical screens produced by the Hyetos software during disaggregation of daily to hourly rainfall data, where plots in green and red refer to disaggregated and original data respectively. Upper left panel shows typical hyetographs, where the green (disaggregated) plot is the result of the storms shown in Fig. 23 converted to a hyetograph at an hourly scale. Notice that while daily totals match, the temporal distribution of rainfall differs in the disaggregated and original hyetographs. However, in statistical sense the disaggregated series resemble the original, as shown in the other panels compare statistics of disaggregated and original series.

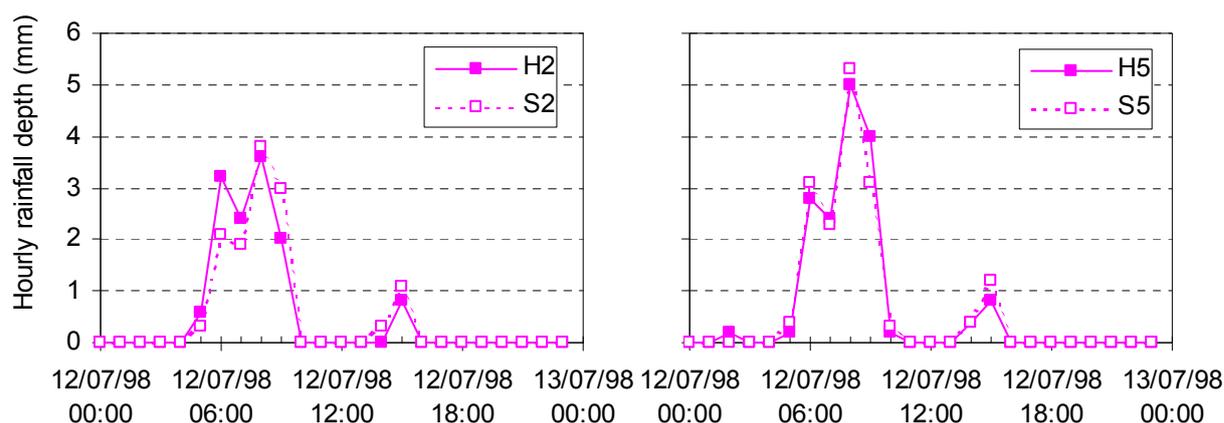


Fig. 26 While, as shown in Fig. 25 (upper left panel), in single-variate disaggregation, the produced hyetographs resemble the actual ones only in a statistical sense, multivariate disaggregation reproduces the actual shapes of hyetographs provided that fine scale (e.g. hourly) data exist in at least one of the stations. The two panels show a comparison of historical (marked H) and simulated (by the MuDRain disaggregation model; marked S) hyetographs at a day with relatively high rainfall (~16mm) at two raingauges (2 and 5) in the Brue catchment located in South-Western England (redrawn from Koutsoyiannis *et al.*, 2003).

4.5 Multifractal models

Rainfall models of multifractal type have for a long time been known to accurately reproduce several statistical properties of actual rainfall fields in finite but practically important ranges of scales: typically from below 1 hour to several days in time and from below 10 km to more than 100 km in space (Schertzer and Lovejoy, 1987; 1989; Tessier *et al.*, 1993; Fraedrich and Larnder, 1993; Olsson, 1995; Lovejoy and Schertzer, 1995; Over and Gupta, 1996; Carvalho *et al.*, 2002; Nykanen and Harris, 2003; Kundu and Bell, 2003; Deidda *et al.* 2004, 2006; Gebremichael and Krajewski, 2004; Calenda *et al.*, 2005; Gebremichael *et al.*, 2006; Veneziano and Langousis, 2005; García-Marín *et al.*, 2007; Langousis and Veneziano, 2007). These properties include the scaling of the moments of different orders (Schertzer and Lovejoy, 1987; Menabde *et al.*, 1997; Deidda *et al.*, 1999, Deidda, 2000), the power law behavior of spatial and temporal spectral densities (Olsson, 1995; Tessier *et al.*, 1996; Deidda *et al.*, 2004, 2006), the alteration of wet and dry intervals (Over and Gupta, 1996; Schmitt *et al.*, 1998; Olsson, 1998; Güntner *et al.*, 2001; Langousis and Veneziano, 2007) and the distribution of extremes (Hubert *et al.*, 1998; Veneziano and Furcolo, 2002; Veneziano and Langousis, 2005; Langousis and Veneziano, 2007; Langousis *et al.*, 2007, Veneziano *et al.*, 2009). Significant deviations of rainfall from multifractal scale invariance have also been pointed out. These deviations include breaks in the power-law behavior of the spectral density (Fraedrich and Larnder, 1993; Olsson, 1995; Menabde *et al.*, 1997), lack of scaling of the non-rainy intervals in time series (Schmitt *et al.*, 1998), differences in scaling during the intense and moderate phases of rainstorms (Venugopal *et al.* 2006), the power deficit at high-frequencies relative to multifractal models (Perica and Foufoula-Georgiou, 1996a,b; Menabde *et al.*, 1997; Menabde and Sivapalan, 2000), and more complex deviations as described in Veneziano *et al.* (2006a).

Next we review some basic properties of stationary multifractal processes and discuss a simple procedure to construct discrete multifractal fields based on the concept of multiplicative cascades.

For a detailed review on the generation of multifractal processes and their applications in hydrological modelling and forecasting, the reader is referred to Veneziano and Langousis (2010).

Let $\underline{i}^{(d)}(t)$ be the average rainfall intensity averaged over time scale d at time t . The stochastic process $\underline{i}^{(d)}(t)$ is said to be stationary multifractal if, for any time scale d , its statistics remain unchanged when the time axis is contracted by a factor $r > 1$ and the intensity is multiplied by a random variable \underline{g}_r , that is

$$\underline{i}^{(d/r)}(t) \stackrel{d}{=} \underline{g}_r \underline{i}^{(d)}(t) \quad (57)$$

where $\stackrel{d}{=}$ denotes equality in (any finite-dimensional) distribution. The notation implies that the distribution of \underline{g}_r depends only on r and not on time t or the intensity $\underline{i}^{(d)}$. Obviously, the mean of \underline{g}_r is 1 and furthermore \underline{g}_r is assumed to be stochastically independent from $\underline{i}^{(d)}$ at the higher scale d . The distribution of \underline{g}_r characterizes the scaling properties as well as many other characteristics of the rainfall process including the marginal distribution, intermittency, distribution of extremes etc. Equation (57) need not apply for arbitrarily large time scales but rather applies up to a maximum scale $d = d_{\max}$. In rainfall, d_{\max} seems to be of the order of several days and it is representative of the mean interarrival time of rainfall events (Langousis and Veneziano, 2007; Langousis *et al.*, 2007; Veneziano *et al.*, 2007). We note for comparison that the related equation in the simple scaling (HK) representation of section 1.5 is $(\underline{i}^{(d/r)} - \mu) \stackrel{d}{=} r^{1-H} (\underline{i}^{(d)} - \mu)$ or $\underline{i}^{(d/r)} \stackrel{d}{=} \mu(1 - r^{1-H}) + r^{1-H} \underline{i}^{(d)}$, so that, when the HK process has zero mean, it can be viewed as a special case of the multifractal process in which the random variable \underline{g}_r is replaced by a deterministic power function of the resolution r .

A property of stationary multifractal processes, which has been used to verify multifractality, is that the spectral density $s(\omega)$ behaves like ω^{-b} where ω is the frequency, and $b < 1$ is a constant (e.g. Fraedrich and Larnder, 1993; Olsson, 1995, Deidda *et al.*, 2004; Hsu *et al.*, 2006). More comprehensive checks of multifractality involve the dependence of statistical moments of different orders on scale. In particular, under perfect multifractality $E[(\underline{i}^{(d)})^q] \propto E[(\underline{g}_r)^q] \propto d^{-K(q)} \propto r^{K(q)}$, where $K(q)$ is a convex function, usually referred to as “moment scaling function” (Gupta and Waymire, 1990; Veneziano, 1999). All concepts and methods are readily extended to space-time rainfall (Veneziano *et al.*, 2006b).

A simple procedure to construct discrete stationary multifractal fields is based on iterative application of equation (57) starting from a large time scale $d \leq d_{\max}$ and gradually decreasing the time scale (i.e. at resolutions $r \propto m^n$, where $m > 1$ and $n \geq 1$ are integers). The contraction by the same factor $r = m$ at each step simplifies generation, since only the distribution of $\underline{g}_r \equiv \underline{g}_m$ is needed. This forms the concept of so-called isotropic discrete multiplicative cascade. Its construction in the D -dimensional cube S_D starts at level 0 with a single “tile” $\Omega_{0_1} \equiv S_D$ with constant unit intensity inside Ω_{0_1} . At level $n = 1, 2, \dots$ (or equivalently at resolutions $r = m^D, m^{2D}, \dots$) each “tile” at the previous level $n - 1$ is partitioned into m^D “tiles” where $m > 1$ is the integer multiplicity of the cascade. The intensity inside each cascade “tile” Ω_{n_i} ($i = 1, \dots, m^{nD}$) is obtained by multiplying that of the parent “tile” at level $n - 1$ by an independent copy \underline{y}_i of a unit-mean random variable \underline{y} , called the generator of the cascade. Clearly, for $r = m^{nD}$, $\underline{g}_r = \underline{y}_1 \underline{y}_2 \dots \underline{y}_n$. For illustration, Fig. 27 shows a simulated realization of a 2-dimensional binary (i.e. $m = 2$) discrete multiplicative cascade developed to level $n = 8$.

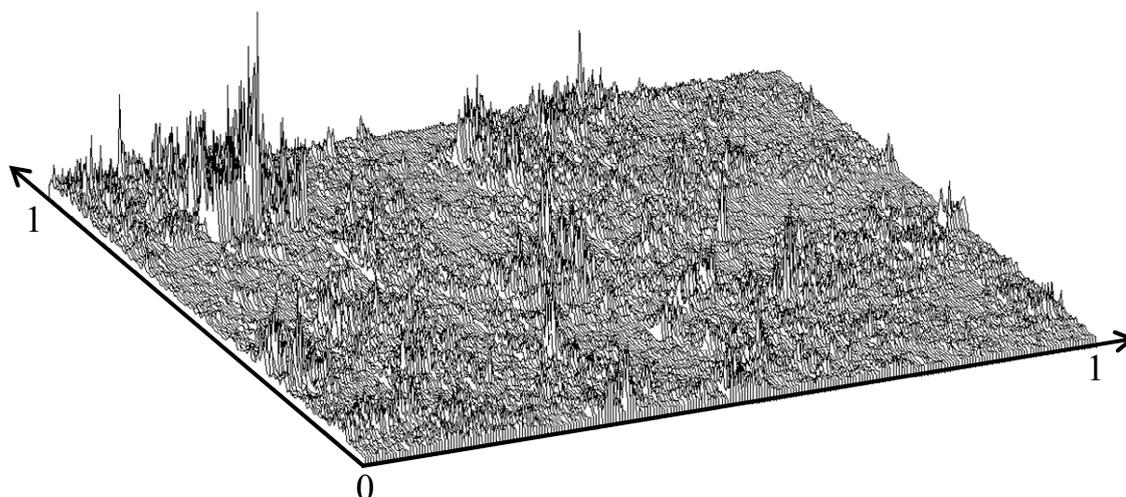


Fig. 27 Simulated realization of a 2-dimensional stationary multifractal field. The random variable y is taken to be lognormal with unit mean value and log-variance $(\sigma_{\ln y})^2 = 0.2 \ln(2)$.

5 Precipitation and engineering design

5.1 Probabilistic vs. deterministic design tools

The design and management of flood protection works and measures requires reliable estimation of flood probability and risk. A solid empirical basis for this estimation can be offered by flow observation records with an appropriate length, sufficient to include a sample of representative floods. In practice, however, flow measurements are never enough to support flood modelling. The obvious alternative is the use of hydrologic models with rainfall input data to generate streamflow. Notably, even when flow records exist, rainfall probability still has a major role in hydrological practice; for instance in major hydraulic structures the design floods are estimated from appropriately synthesised design storms (e.g. US Department of the Interior, Bureau of Reclamation, 1977, 1987; Sutcliffe, 1978). The need to use rainfall data as the basis of hydrologic design becomes even more evident in the study of engineering structures and urban water management systems that modify the natural environment, so that past flood records, even if they exist, are no longer representative of the future modified system.

Hydrologic design does not necessarily require full modelling of the rainfall process, of the type discussed in section 4. Usually, in design studies, the focus is on extreme rainfall, which, notably, may not be represented well in such models, which are better for the “average” behaviour of rainfall. However, historically, the perception of intense rainfall and the methodologies devised to model it have suffered from several fallacies spanning from philosophical to practical issues, which we describe next to cast a warning against their acceptance and use.

The first fallacy is of a rather philosophical type. As discussed in section 1.5, the modelling of the rainfall process in pure deterministic terms has been proven to be problematic. However, deterministic thinking in science is strong enough, so that after the failure in providing full descriptions, it was headed to determining physical bounds to precipitation in an attempt to design risk-free constructions or practices. The resulting concept of probable maximum precipitation (PMP),

that is, an upper bound of precipitation that is physically feasible (World Meteorological Organization, 1986), is perhaps one of the biggest failures in hydrology. Using elementary logic we easily understand that even the terminology is self-contradictory, and thus not scientific. Namely, the word “probable” contradicts the existence of a deterministic limit.

Several methods to determine PMP exist in literature and are described in World Meteorological Organization (1986). However, examination in depth of each of the specific methods separately will reveal that they are all affected by logical inconsistencies. While they are all based on the assumption of the existence of a deterministic upper limit, they determine this limit statistically. This is obvious in the so called “statistical approach” by Hershfield (1961, 1965), who used 95 000 station-years of annual maximum daily rainfall belonging to 2645 stations, standardized each record and found the maximum over the 95 000 standardized values. Naturally, one of the 95 000 standardized values would be the greatest of all others, but this is not a deterministic limit to call PMP (Koutsoyiannis, 1999). If one examined 95 000 additional measurements one might have found an even higher value. Thus the logical problem here is the incorrect interpretation that an observed maximum in precipitation is a physical upper limit.

The situation is perhaps even worse with the so-called moisture maximization approach of PMP estimation (World Meteorological Organization, 1986), which seemingly is more physically (hydrometeorologically) based than the statistical approach of Hershfield. In fact, however, it suffers twice by the incorrect interpretation that an observed maximum is a physical upper limit. It uses a record of observed dew point temperatures to determine an upper limit, which is the maximum observed value. Then it uses this “limit” for the so called “maximization” of an observed sample of storms, and asserts the largest value among them as PMP. Clearly, this is a questionable statistical approach, because (a) it does not assign any probability to the value determined and (b) it is based only on one observed value (known in statistics as the highest order statistic), rather than on the whole sample, and thus it is enormously sensitive to one particular observation of the entire sample (Papalexiou and Koutsoyiannis, 2006; Koutsoyiannis, 2007). Thus, not only does the determination of PMP use a statistical approach (rather than deterministic physics), but it uses bad statistics. The arbitrary assumptions of the approach extend beyond the confusion of maximum observed quantities with physical limits. For example, the logic of moisture maximization at a particular location is unsupported given that a large storm at this location depends on the convergence of atmospheric moisture from much greater areas.

Rational thinking and fundamental philosophical and scientific principles can help identify and dispel such fallacies. In particular, the Aristotelian notions of *potentia* (Greek ‘*dynamis*’) and of potential infinite (Greek ‘*apeiron*’; Aristotle, *Physics*, 3.7, 206b16) that “exists in no other way, but ... potentially or by reduction” (and is different from mathematical complete infinite) would help us to avoid the PMP concept. In fact, this does not need a great deal of philosophical penetration. The same thing is more practically expressed as “conceptually, we can always imagine that a few more molecules of water could fall beyond any specified limit” (Dingman, 1994). Yet the linkage to the Aristotelian notions of *potentia* and potential infinity may make us more sensitive in seeing the logical inconsistencies (see also Koutsoyiannis, 2007).

According to Popper (1982) the extension of the Aristotelian idea of *potentia* in modern terms is the notion of probability. Indeed, probability provides a different way to perceive the intense rainfall

and flood and to assign to each value a certain probability of exceedence (see next session) avoiding the delusion of an upper bound of precipitation and the fooling of decision makers that they can build risk free constructions. In this respect the criticism of the PMP and the probable maximum flood (PMF) involves logical, technical, philosophical and ethical issues (e.g. Benson, 1973).

One typical argument against the use of probabilistic approaches, in favour of PMP, which is very old yet popular even today, has been stated by Horton (1931; from Klemes, 2000), "It is, however, important to recognize the nature of the physical processes involved and their limitations in connection with the use of statistical methods. ... Rock Creek cannot produce a Mississippi River flood any more than a barnyard fowl can lay an ostrich egg". However, this argument reveals an incorrect perception of probability and statistics. In a probability theoretic context there is not a logical inconsistency. Assuming, for example that the annual peak flood of the Mississippi river (x_M) is on the average (μ_M), a million times larger than the average (μ_C) flood of a certain small creek (x_C), and assuming that both x_M and x_C have a lognormal distribution with standard deviation $\sigma_{\ln x}$ of logarithms of about 0.3 (which is roughly equal to the coefficient of variation of the annual flood peaks, assumed equal in the two streams), one can readily find that the probability that the flood in the creek x_C in some year exceeds the mean annual flood μ_M of Mississippi is $\Phi^*(\zeta) := 1 - \Phi_G(\zeta)$ where Φ_G is the standard normal distribution function and $\zeta = \ln(\mu_M/\mu_C)/\sigma_{\ln x}$ or $\zeta = \ln(10^6)/0.3 = 46$. For large ζ , the approximation $\ln \Phi^*(\zeta) = -(1/2)[\ln(2\pi\zeta^2) + \zeta^2]$ holds (e.g. Abramowitz and Stegun, 1965), hence $\ln \Phi^*(\zeta) = -1062.75$, so that the probability of exceedence is $\Phi^*(\zeta) = 10^{-462}$. That is, according to the probabilistic approach, the return period of the event that the small creek flood matches or exceeds the mean annual flood of the Mississippi is 10^{462} years. Assuming that the age of the universe is of the order of 10^{10} years, one would wait, on the average, 10^{452} times the age of the universe to see this event happen—if one foolishly hoped that the creek, the Mississippi and the Earth would exist for such a long time. Evidently, such a low probability could be regarded as synonymous to impossibility, which shows that the probabilistic approach does not regard the floods of Mississippi equivalent to those of a small creek (see also an example about the age of a person by Feller, 1950).

5.2 Extreme rainfall distribution

Having being exempted from the concept of an upper limit to precipitation and having adopted a probabilistic approach, the real problem is how the rainfall intensity grows as the probability of exceedence decreases. Clearly, as the probability of exceedence tends to zero, the intensity tends to infinity. There exists a mathematically proven lower limit to the rate of this growth, which is represented by an exponential decay of the probability of exceedence with intensity. The alternative is a power law decay and, as already mentioned in section 1.5, the two options may lead to substantial differences in design quantities for high return periods. In this respect, the most important questions, which have not received definite answers yet, are again related to the notion of infinity.

Accordingly, the distribution tails are important to know in engineering design. However, the study of the tails is difficult and uncertain because the tails refer to infrequent events that require very long records to appear. Traditionally, rainfall records are analyzed in two ways. The most frequent is to choose the highest of all recorded precipitation intensities (for a given averaging time scale) at each year and form a statistical sample with size equal to the number of years of the record. The

other is to form a sample with all recorded intensities over a certain threshold irrespectively of the year they occurred. Usually the threshold is chosen high enough, so that the sample size is again equal to the number of years of the record. This however is not necessary: it can well be set equal to zero, so that all recorded intensities are included in the sample. However, the threshold simplifies the study and helps focus the attention on the distribution tail.

If $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ are random variables representing the recorded average intensities within a year at nonoverlapping time periods equal to a chosen time scale d , then the maximum among them $\underline{y} := \max(\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n)$ has a distribution function $H_n(y)$ fully dependent on the joint distribution function of \underline{x}_j . Assuming that \underline{x}_j are IID with common distribution function $F(x)$, then $H_n(x) = [F(x)]^n$. If n is not constant but rather can be regarded as a realisation of a random variable (corresponding to the fact that the number of rainfall events is not the same in each year) with Poisson distribution with mean ν , then the distribution function H becomes (e.g. Todorovic and Zelenhasic, 1970; Rossi et al., 1984),

$$H(x) = \exp\{-\nu[1 - F(x)]\} \quad (58)$$

In particular, if the threshold has been chosen with the above rule (to make the sample size equal to the number of years of the record) then obviously $\nu = 1$. Equation (58) expresses in a satisfactory approximation the relationship between the above two methodologies and the respective distributions F and H . The two options discussed above are then represented as follows:

1. Exponential tail

$$F(x) = 1 - \exp(-x/\lambda + \psi), \quad H(x) = \exp[-\exp(-x/\lambda + \psi)], \quad x \geq \lambda\psi \quad (59)$$

where $\lambda > 0$ and $\psi > 0$ are parameters, so that $\lambda\psi$ represents the specified threshold. Here F is the exponential distribution and H is the Gumbel distribution, also known as extreme value type I (EV1) distribution.

2. Power tail

$$F(x) = 1 - \left[1 + \kappa \left(\frac{x}{\lambda} - \psi\right)\right]^{-1/\kappa}, \quad H(x) = \exp\left\{-\left[1 + \kappa \left(\frac{x}{\lambda} - \psi\right)\right]^{-1/\kappa}\right\}, \quad x \geq \lambda\psi \quad (60)$$

where $\lambda > 0$, $\psi > 0$ and $\kappa > 0$ are parameters and $\lambda\psi$ represents the specified threshold. Here F is the generalized Pareto distribution (a generalized form of equation (18)) and H is the generalized extreme value (GEV) distribution. In the case $\kappa > 0$ considered here, GEV is also called the extreme value type II (EV2) distribution. The case $\kappa < 0$ is mathematically possible and is called the extreme value type III (EV3) distribution. However, this is inappropriate for rainfall as it puts an upper bound ($\lambda\psi$) for x , which is inconsistent. The case $\kappa = 0$, corresponds precisely to the exponential tail (exponential and Gumbel distributions).

For years the exponential tail and the Gumbel distribution have been the prevailing models for rainfall extremes, despite the fact that they yield unsafe (the smallest possible) design rainfall values. Recently, however, their appropriateness for rainfall has been questioned. Koutsoyiannis (2004a, 2005a, 2007) discussed several theoretical reasons that favour the power/EV2 over the exponential/EV1 case. As already mentioned (section 1.5.5), Koutsoyiannis (2004b, 2005a) compiled an ensemble of annual maximum daily rainfall series from 169 stations in the Northern Hemisphere

(28 from Europe and 141 from the USA) roughly belonging to six major climatic zones and all having lengths from 100 to 154 years. The analysis provides sufficient support for the general applicability of the EV2 distribution model worldwide. Furthermore, the ensemble of all samples was analysed in combination and it was found that several dimensionless statistics are virtually constant worldwide, except for an error that can be attributed to a pure statistical sampling effect. This enabled the formation of a compound series of annual maxima, after standardization by the mean, for all stations (see Fig. 13, which shows the distribution of a compound sample over threshold of all stations, except one in which only annual maxima existed). The findings support the estimation of a unique κ for all stations, which was found to be 0.15.

Additional empirical evidence with the same conclusions is provided by the Hershfield's (1961) data set, which was the basis of the formulation of Hershfield's PMP method. Koutsoyiannis (1999) showed that this data set does not support the hypothesis of an upper bound in precipitation, that is PMP. Rather it is consistent with the EV2 distribution with $\kappa = 0.13$, while the value $\kappa = 0.15$ can be acceptable for that data set too (Koutsoyiannis, 2004b). This enhances the trust that an EV2 distribution with $\kappa = 0.15$ can be regarded as a generalized model appropriate for mid latitude areas of the Northern Hemisphere.

In a recent study, Veneziano *et al.* (2009) used multifractal analysis to show that the annual rainfall maximum for time scale d can be approximated by a GEV distribution and that typical values of κ lie in the range 0.09 to 0.15 with the larger values being associated with more arid climates. This range of values agrees well with the findings of Koutsoyiannis (1999, 2004b, 2005a). Similar results were provided by Chaouche (2001) and Chaouche *et al.* (2002). Chaouche (2001) exploited a data base of 200 rainfall series of various time steps (month, day, hour, minute) from the five continents, each including more than 100 years of data. Using multifractal analyses it was found that (a) an EV2/Pareto type law describes the rainfall amounts for large return periods; (b) the exponent of this law is scale invariant over scales greater than an hour (as we stated in section 4.2, it cannot be otherwise because this is dictated by theoretical reasons); and (c) this exponent is almost space invariant. Other studies have also expressed scepticism for the appropriateness of the Gumbel distribution for the case of rainfall extremes and suggested hyper-exponential tail behaviour. Coles *et al.* (2003) and Coles and Pericchi (2003) concluded that inference based on a Gumbel distribution model fitted to the annual maxima may result in unrealistically high return periods for certain observed events and suggested a number of modifications to standard methods, among which is the replacement of the Gumbel model with the GEV model. Mora *et al.* (2005) confirmed that rainfall in Marseille (a raingauge included in the study by Koutsoyiannis, 2004b) shows hyper-exponential tail behaviour. They also provided two regional studies in the Languedoc-Roussillon region (south of France) with 15 and 23 gauges, for which they found that a similar distribution with hyper-exponential tail could be fitted. This finding, when compared to previous estimations, leads to a significant increase in the depth of rare rainfall. On the same lines, Bacro and Chaouche (2006) showed that the distribution of extreme daily rainfall at Marseille is not in the Gumbel law domain. Sisson *et al.* (2006) highlighted the fact that standard Gumbel analyses routinely assign near-zero probability to subsequently observed disasters, and that for San Juan, Puerto Rico, standard 100-year predicted rainfall estimates may be routinely underestimated by a factor of two. Schaefer *et al.* (2006) using the methodology by Hosking and Wallis (1997) for regional precipitation-frequency

analysis and spatial mapping for 24-hour and 2-hour durations for the Washington State, USA, found that the distribution of rainfall maxima in this State generally follows the EV2 distribution type.

5.3 Ombrian relationships

One of the major tools in hydrologic design is the ombrian relationship, more widely known by the misnomer rainfall intensity-duration-frequency (IDF) curve. An ombrian relationship (from the Greek ‘ombros’, rainfall) is a mathematical relationship estimating the average rainfall intensity i over a given time scale d (sometimes incorrectly referred to as duration) for a given return period T (also commonly referred to as frequency, although “frequency” is generally understood as reciprocal to “period”). Several forms of ombrian relationships are found in the literature, most of which have been empirically derived and validated by the long use in hydrologic practice. Attempts to give them a theoretical basis have often used inappropriate assumptions and resulted in oversimplified relationships that are not good for engineering studies.

In fact, an ombrian relationship is none other than a family of distribution functions of rainfall intensity for multiple time scales. This is because, the return period is tied to the distribution function, i.e., $T = \delta / [1 - F(x)]$, where δ is the mean interarrival time of an event that is represented by the variable x , typically 1 year. Thus, a distribution function such as one of those described in section 4.2, is at the same time an ombrian relationship. This has been made clear in Koutsoyiannis *et al.* (1998) who showed that the empirical considerations usually involved in the construction of ombrian curves are not necessary at all and create difficulties and confusion.

However, the direct use in engineering design of a fully consistent multiscale distribution function may be too complicated. Simplifications are possible to provide satisfactory approximations, given that only the distribution tail is of interest and that the range of scales of interest in engineering studies is relatively narrow. Such simplifications, which were tested recently and were found to be reasonable (Papalexioiu and Koutsoyiannis, 2009) are:

1. The separability assumption, according to which the influences of return period and time scale are separable (Koutsoyiannis *et al.*, 1998), i.e.,

$$i(d, T) = a(T) / b(d) \quad (61)$$

where $a(T)$ and $b(d)$ are mathematical expressions to be determined.

2. The use of the Pareto distribution for the rainfall intensity over some threshold at any time scale, as discussed in section 5.2; this readily provides a simple expression for $a(T)$.
3. The expression of $b(d)$ in the simple form

$$b(d) = (1 + d/\vartheta)^\eta \quad (62)$$

where $\vartheta > 0$ and $\eta > 0$ are parameters. A justification of this relationship, which is a satisfactory approximation for time scales up to a few days, can be found in Koutsoyiannis (2006a).

Based on assumptions 1-3, we easily deduce that the final form of the ombrian relationship is

$$i(d, T) = \lambda' \frac{(T/\delta)^\kappa - \psi'}{(1 + d/\vartheta)^\eta} \quad (63)$$

where $\psi' > 0$, $\lambda' > 0$ and $\kappa > 0$ are parameters. In particular, as discussed in section 5.2, κ is the tail-determining parameter and unless a long record exists, which could support a different value, it should be assumed $\kappa = 0.15$. Equation (63) is dimensionally consistent, provided that ϑ has units of time (as well as δ), λ' has units of intensity, and κ and ψ are dimensionless. The numerator of equation (63) differs from a pure power law that has been commonly used in engineering practice, as well as in some multifractal analyses. Consistent parameter estimation techniques for ombrian relationships have been discussed in Koutsoyiannis *et al.* (1998) as well as in Chapter 46, “Statistical Hydrology”, of this volume.

Acknowledgments

We thank Vasiliki Kotroni and Remko Uijlenhoet for their insightful and constructive comments and suggestions on an earlier version of this chapter. Their detailed reviews greatly improved the presentation of our work. We also thank Federico Lombardo for his comments on the final draft.

Cross References

Chapter 46, “Statistical Hydrology”

Chapter 79, “Abstraction of atmospheric humidity”

References

- Abramowitz M and Stegun IA (1965) *Handbook of Mathematical Functions*, Dover, New York.
- Adler RF and Negri AJ (1988) A Satellite infrared technique to estimate tropical convective and stratiform rainfall, *J. Appl. Meteor.*, 27, 30–51.
- Alishouse JC, Snyder SA, Vongsathorn J and Ferraro RR (1990) Determination of oceanic total precipitable water from the SSM/I, *Geoscience and Remote Sensing, IEEE Transactions*, 28(5), 811-816.
- Anthes RA (1982) *Tropical Cyclones: Their Evolution, Structure, and Effects*, Meteorological Monographs Number 41, American Meteorological Society, Boston, MA.
- Aonashi K and Liu G (2000) Passive microwave precipitation retrievals using TMI during the Baiu period of 1998. Part I: Algorithm Description and Validation, *J. Appl. Meteor.*, 39, 2024–2037.
- Arakawa H (1956) On the secular variation of annual totals of rainfall at Seoul from 1770 to 1944, *Theoretical and Applied Climatology*, 7(2), 205–211.
- Arkin PA (1979) The Relationship between fractional coverage of high cloud and rainfall accumulations during GATE over the B-scale array, *Mon. Wea. Rev.*, 107, 1382–1387.
- Arkin PA and Meisner BN (1987) The relationship between large-scale convective rainfall and cold cloud over the Western Hemisphere during 1982-84, *Mon. Wea. Rev.*, 115, 51–74.

- Atlas D and Ulbrich CW (1977) Path and area integrated rainfall measurement by microwave attenuation in the 1-3 cm band, *J. Appl. Meteor.*, 16, 1322-1331.
- Atlas D, Ulbrich CW and Meneghini R (1984) The multiparameter remote measurement of rainfall, *Radio Sci.*, 19, 3-21.
- Austin PM and Houze RA (1972) Analysis of the Structure of Precipitation Patterns in New England, *J. Appl. Meteor.*, 11, 926-935.
- Ba MB and Nicholson SE (1998) Analysis of convective activity and its relationship to the rainfall over the rift valley lakes of East Africa during 1983-90 using the Meteosat infrared channel, *J. Appl. Meteor.*, 37, 1250-1264.
- Bacro J-N and Chaouche A (2006) Incertitude d'estimation des pluies extrêmes du pourtour méditerranéen: illustration par les données de Marseille, *Hydrol. Sci. J.*, 51(3), 389-405.
- Baloutsos G, Bourletsikas A and Kaoukis K (2005), Study and investigation of the fog precipitation characteristics in the fir forest of Agios Nikolaos, Evritania, Greece, in Greek, *Geotechnica Epistemonica Themata*, 2(16), 34-45.
- Barnes GM, Zipser EJ, Jorgensen D and Marks FD (1983) Mesoscale and convective structure of hurricane rainband, *J. Atmos. Sci.*, 40, 2125-2137.
- Barrett EC and Martin DW (1981) *The Use of Satellite Data in Rainfall Monitoring*, Academic Press, London, U.K.
- Battan LJ (1973) *Radar Observations of the Atmosphere*, The University of Chicago Press, Chicago, U.S.A.
- Beard KV (1976) Terminal velocity and shape of cloud and precipitation drops aloft, *J. Atmos. Sci.*, 33, 851-864.
- Beer T and Giannini L (1980) Tropical cyclone cloud bands, *J. Atmos. Sci.*, 37, 1511-1520.
- Bellon A, Lovejoy S and Austin G (1980) Combining satellite and radar data for the short-range forecasting of precipitation, *Mon. Wea. Rev.*, 108, 1554-1566.
- Bennetts DA and Sharp JC (1982) The relevance of conditional symmetric instability to the prediction of mesoscale frontal rainbands, *Quart. J. Roy. Meteor. Soc.*, 108, 595-602.
- Benson MA (1973), Thoughts on the design of design floods, in *Floods and Droughts, Proc. 2nd Intern. Symp. in Hydrology*, pp. 27-33, Water Resources Publications, Fort Collins, Colorado.
- Berg W and Chase R (1992) Determination of mean rainfall from the special sensor microwave/imager (SSM/I) using a mixed lognormal distribution, *J. Atmos. Oceanic Technol.*, 9, 129-141.
- Berg W, Olson W, Ferraro R, Goodman SJ and LaFontaine FJ (1998) An assessment of the first- and second-generation navy operational precipitation retrieval algorithms, *J. Atmos. Sci.*, 55, 1558-1575.

- Bergeron T (1935) On the physics of cloud and precipitation, *Proc. Assem. Int. Union Geodesy Geophys.*, 5th, Lisbon.
- Berry EX and Reinhardt RL (1974a) An Analysis of cloud drop growth by collection: Part I. Double distributions, *J. Atmos. Sci.*, 31, 1814-1824.
- Berry EX and Reinhardt RL (1974b) An analysis of cloud drop growth by collection: Part II. Single initial distributions, *J. Atmos. Sci.*, 31, 1825-1831.
- Bluestein HB (2003) Tornadoes, In: *Encyclopedia of Atmospheric Sciences*, Eds. Holton JR, Curry JA and Pyle JA, pp. 2290-2297, Academic Press, London, UK.
- Bowen EG (1950) The formation of rain by coalescence, *Australian J. Sci. Res.*, A3, 192-213.
- Bringi VN and Chandrasekar V (2001) *Polarimetric Doppler Weather Radar*, Cambridge Univ. Press, New York.
- Brock JR (1972) Condensational growth of atmospheric aerosols, In: *Aerosols and Atmospheric Chemistry*, Ed. Hidy GM, Academic Press, New York, 149-153.
- Browning KA and Ludlam FH (1962) Airflow in convective storms. *Quart. J. Roy. Meteor. Soc.*, 88, 117-135.
- Byers HR and Braham RR (1949) *The Thunderstorm*, U.S. Dept. of Commerce, Washington.
- Calenda G, Gorgucci E, Napolitano F, Novella A and Volpi E (2005) Multifractal analysis of radar rainfall fields over the area of Rome, *Advances in Geosc.*, 2, 293-299.
- Carvalho LMV, Lavallée D and Jones C (2002) Multifractal properties of evolving convective systems over tropical South America, *Geophys. Res. Lett.*, 29, doi: 10.1029/2001GL014276.
- Chaouche K (2001) Approche multifractale de la modélisation stochastique en hydrologie. Thèse, Ecole Nationale du Génie Rural, des Eaux et des Forêts, Centre de Paris, France (www.engref.fr/thesechaouche.htm).
- Chaouche K, Hubert P and Lang G (2002) Graphical characterisation of probability distribution tails, *Stoch. Environ. Res. Risk Assess.*, 16(5), 342-357.
- Chen SS, Lonfat M, Knaff JA and Marks FD Jr (2006) Effects of vertical wind shear and storm motion on tropical cyclone rainfall asymmetries deduced from TRMM, *Mon. Wea. Rev.*, 134, 3190-3208.
- Chen SS, Price JF, Zhao W, Donelan M and Walsh EJ (2007) The CBLAST-hurricane program and the next-generation fully coupled atmosphere-wave-ocean models for hurricane research and prediction, *Bull. Amer. Meteor. Soc.*, 88, 311-317.
- Cheng M and Brown R (1995) Delineation of precipitation areas by correlation of Meteosat visible and infrared data with radar data, *Mon. Wea. Rev.*, 123, 2743-2757.

- Cheng M, Brown R and Collier C (1993) Delineation of precipitation areas using Meteosat infrared and visible data in the region of the United Kingdom, *J. Appl. Meteor.*, 32, 884–898.
- Chin EH (1977) Modelling daily precipitation occurrence process with Markov chains, *Water Resour. Res.*, 13(6), 959-956.
- Clarke RT (1998) *Stochastic Processes for Water Scientists*, Wiley, New York.
- Coles S and Pericchi L (2003) Anticipating catastrophes through extreme value modelling, *Appl. Statist.*, 52, 405–416.
- Coles S, Pericchi LR and Sisson S (2003) A fully probabilistic approach to extreme rainfall modeling, *J. Hydrol.*, 273(1–4), 35–50.
- Cotton WR and Anthes RA (1989) *Storm and Cloud Dynamics*, International Geophysics Series Vol. 44, Academic Press Inc., USA
- Cotton WR, Thompson G and PW Mielke (1994) Realtime mesoscale prediction on workstations, *Bull. Amer. Meteor. Soc.*, 75, 349-362.
- Cowpertwait P, Isham V and Onof C (2007) Point process models of rainfall: developments for fine-scale structure, *Proceedings of the Royal Society A*, 463 (2086), 2569-2587.
- Cox D and Isham V (1988), A simple spatial-temporal model of rainfall, *Proc. R. Soc. London, Ser. A*, 415, 317–328.
- De Bruin, HAR (1980) A stochastic description of wet and dry spells in terms of an effective number of days, *J. of Hydrol.*, 45, 91-99.
- Deidda R (2000) Rainfall downscaling in a space-time multifractal framework, *Wat. Resour. Res.*, 36(7), 1779-1794.
- Deidda R, Benzi R and Siccardi F (1999) Multifractal modeling of anomalous scaling laws in rainfall, *Wat. Resour. Res.*, 35(6), 1853-1867.
- Deidda R, Badas MG and Piga E (2004) Space-time scaling in high-intensity tropical ocean global atmosphere coupled ocean-atmosphere response experiment (TOGA-COARE) storms, *Water Resour. Res.*, 40, doi:10.1029/2003WR002574.
- Deidda R, Grazia-Badas M and Piga E (2006) Space–time multifractality of remotely sensed rainfall fields, *J. Hydrol.*, 322, 2-13, doi:10.1016/j.jhydrol.2005.02.036.
- Dingman SL (1994), *Physical Hydrology*, Prentice Hall, Englewood Cliffs, New Jersey.
- Doviak R and Zrnich D (1993) *Doppler Radar and Weather Observations*, 562 pp., Elsevier, New York.
- Drake RL (1972a) A general mathematical survey of the coagulation equation, In: *Topics in Current Aerosol Research* (Part 2), Eds. Hidy GM and Brock JR, Pergamon Press, Oxford, 201-384.
- Drake RL (1972b) The scalar transport equation of coalescence theory: moments and kernels, *J. Atmos. Sci.*, 29, 537-547.

- Elachi C (1987) *Introduction to the Physics and Techniques of Remote Sensing*, Wiley Series in Remote Sensing, John Wiley and Sons, USA.
- Emanuel KA (1986) An air-sea interaction theory for tropical cyclones. Part I: Steady state maintenance, *J. Atmos. Sci.*, 43, 585-604.
- Emanuel KA (1989) Polar lows as arctic hurricanes, *Tellus*, 41A, 1-17.
- Entekhabi D, Rodriguez-Iturbe I and Eagleson PS (1989) Probabilistic representation of the temporal rainfall process by the modified Neyman-Scott rectangular pulses model: Parameter estimation and validation, *Water Resour. Res.*, 25(2), 295-302.
- Epstein ES (1969) Stochastic dynamic prediction, *Tellus*, 21, 739-757.
- Federer B and Waldvogel A (1975) Hail and raindrop size distributions from a swiss multicell storm, *J. Appl. Meteor.*, 14, 91-97.
- Feingold G and Levin Z (1986) The lognormal fit to raindrop spectra from frontal convective clouds in Israel, *J. Clim. Appl. Meteor.*, 25, 1346-1363.
- Feller W (1950) *An introduction to Probability Theory and its Applications*, Wiley, New York.
- Ferraro RR and Marks GF (1995) The development of SSM/I rain-rate retrieval algorithms using ground-based radar measurements, *J. Atmos. Oceanic Technol.*, 12, 755-770.
- Ferriday JG and Avery SK (1994) Passive microwave remote sensing of rainfall with SSM/I: Algorithm development and implementation, *J. Appl. Meteor.*, 33, 1587-1596.
- Feyerherm AM and Bark LD (1964) Statistical methods for persistent precipitation patterns, *J. Appl. Meteorol.*, 4, 320-328.
- Findeisen W (1938) Die Kolloid Meteorologischen Vorgang der Niederschlags Bildung, *Meteorol. Z.*, 55, 121-133.
- Fletcher NH (1962) *The Physics of Rainclouds*, Cambridge University Press.
- Foufoula-Georgiou E and Guttorp P (1986) Compatibility of continuous rainfall occurrence models with discrete rainfall observations, *Water Resour. Res.*, 22(8), 1316-1322.
- Foufoula-Georgiou E and Lettenmaier DP (1986) Continuous-time versus discrete-time point process models for rainfall occurrence series, *Water Resour. Res.*, 22(4), 531-542.
- Fraedrich K and Larnder C (1993) Scaling regimes of composite rainfall time series, *Tellus*, 45A, 289-298.
- Frank WM (1977) The Structure and energetics of the tropical cyclone: I. Storm structure, *Mon. Wea. Rev.*, 105, 1119-1135.
- Gabriel KR and Neumann J (1962) A Markov chain model for daily rainfall occurrences at Tel Aviv, *Royal Meteorol. Soc. Quart. J.*, 88, 90-95.

- García-Marín AP, Jiménez-Hornero FJ and Ayuso JL (2007) applying multifractality and the self-organized criticality theory to describe the temporal rainfall regimes in Andalusia (southern Spain), *Hydrol. Process.*, doi:10.1002/hyp.6603.
- Gebremichael M and Krajewski WF (2004) Assessment of the statistical characterization of small-scale rainfall variability from radar: Analysis of TRMM ground validation datasets, *J. Appl. Meteor.*, 43(8), 1180-1199.
- Gebremichael M, Over TM and Krajewski WF (2006) Comparison of the scaling properties of rainfall derived from space- and surface-based radars, *J. of Hydrometeor.*, 7, 1277-1294.
- Georgakakos KP, Carsteanu AA, Sturdevant PL and Cramer JA (1994) Observation and analysis of Midwestern rain rates, *J. Appl. Met.*, 33, 1433–1444.
- Gillespie DT (1972) The stochastic coalescence model for cloud droplet growth, *J. Atmos. Sci.*, 29, 1496-1510.
- Gillespie DT (1975) Three models for the coalescence growth of cloud drops, *J. Atmos. Sci.*, 32, 600-607.
- Gray WM (1968) Global view of the origins of tropical disturbances and storms, *Mon. Wea. Rev.*, 96, 669-700.
- Gray WM (1979) Hurricanes: Their formation, structure and likely role in the tropical circulation, In: *Meteorology Over the Tropical Oceans*, Ed. Shaw DB, *Roy. Meteor. Soc.*
- Grell GA, Dudhia J and Stauffer DR (1995) A description of the fifth-generation Pennsylvania State/NCAR mesoscale model (MM5), NCAR Tech. Note TN-398+STR122 pp [Available from UCAR Communications, P.O. Box 3000, Boulder, CO 80307].
- Griffith CG, Woodley WL, Grube PG, Martin DW, Stout J and Sikdar DN (1978) rain estimation from geosynchronous satellite imagery—Visible and infrared studies, *Mon. Wea. Rev.*, 106, 1153–1171.
- Grody NC (1991) Classification of snow cover and precipitation using the special sensor microwave imager, *J. Geophys. Res.*, 96(D4), 7423–7435.
- Gunn KLS and Marsall JS (1958) The distribution with size of aggregate snowflakes, *J. Meteor.*, 15, 452-461.
- Gunn R and Kinzer GD (1949) The terminal velocity of fall for water drops in stagnant air, *J. Meteor.*, 6, 243-248.
- Güntner A, Olsson J, Calver A and Gannon B (2001) Cascade-based disaggregation of continuous rainfall time series: The influence of climate, *Hydrol. Earth Syst. Sci.*, 5, 145-164.
- Gupta VK and Waymire E (1990) Multiscaling properties of spatial rainfall and river flow distributions, *J. Geophys. Res.*, 95, 1999-2009.

- Gyasi-Agyei Y and Willgoose GR (1997) A hybrid model for point rainfall modeling, *Water Resour. Res.*, 33(7) 1699-1706.
- Haan CT, Allen DM and Street JO (1976) A Markov chain model of daily rainfall, *Water Resour. Res.*, 12(3), 443-449.
- Hakim GJ (2003) Cyclogenesis, In: *Encyclopedia of Atmospheric Sciences*, Eds. Holton JR, Curry JA and Pyle JA, pp. 589-594, Academic Press, London, UK.
- Hane CE (1986) Extratropical squall lines and rainbands, Ch. 16, pp. 359-389, In: *Mesoscale Meteorology and Forecasting*, Ed. Ray PS, *Amer. Meteor. Soc.*, Boston, USA.
- Harr RD (1982) Fog drip in the Bull Run Municipal watershed, Oregon, *Water Resources Bulletin* 18, 785-789.
- Harrold TI, Sharma A and Sheather SJ (2003) A nonparametric model for stochastic generation of daily rainfall occurrence, *Water Resour. Res.*, 39(10), 1300, doi: 10.1029/2003WR002182.
- Harrold TW and Austin PM (1974) The Structure of precipitation systems—A review, *J. Rech. Atmos.*, 8, 41-57.
- Hawkins HF and Rubsam DT (1968) Hurricane Hilda, 1964. II. Structure and budgets of the hurricane on October 1, 1964, *Mon. Wea. Rev.*, 96, 617-636.
- Hemond HF and Fechner-Levy EJ (2000) *Chemical Fate and Transport in the Environment*, Second Edition, Academic Press, London, UK.
- Hershfield DM (1961) Estimating the probable maximum precipitation, *Proc. ASCE, J. Hydraul. Div.*, 87(HY5), 99-106.
- Hershfield DM (1965) Method for estimating probable maximum precipitation, *J. American Waterworks Assoc.*, 57, 965-972.
- Hershfield DM (1970) A comparison of conditional and unconditional probabilities for wet- and dry-day sequences, *J. Appl. Meteorol.*, 9(5), 825-827.
- Hinton BB, Olson WS, Martin DW and Auvine B (1992) A passive microwave algorithm for tropical oceanic rainfall, *J. Appl. Meteorol.*, 31, 1379-1395.
- Hobbs PV, Bowdle DA and Radke LF (1985) Particles in the lower troposphere over the high plains of the United States. Part I: Size Distributions, Elemental Compositions and Morphologies, *J. Clim. Appl. Meteorol.*, 24, 1344-1356.
- Holland GJ (1980) An analytic model of the wind and pressure profiles in hurricanes, *Mon. Wea. Rev.*, 108, 1212-1218.
- Hong S-Y, Dudhia J and Chen S-H (2004) A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation, *Mon. Wea. Rev.*, 132, 103-120.

- Horton RE (1931) The field, scope and status of the science of hydrology, *Trans. Am. Geophys. Union, 12th Annu. Mtg.*, National Research Council, Washington D.C., 189-202.
- Hosking JRM and Wallis JR (1997) *Regional Frequency Analysis—An Approach Based on L-Moments*, Cambridge Univ. Press, New York.
- Houghton HG (1985) *Physical Meteorology*, MIT Press, Cambridge, Mass.
- Houze RA (1993) *Cloud Dynamics*, Academic Press, California, U.S.A.
- Houze RA, Hobbs PV, Herzegh PH and Parsons DB (1979) Size distributions of precipitation particles in frontal clouds, *J. Atmos. Sci.*, 36, 156-162.
- Houze RA, Marks FD Jr and Black RA (1992) Dual-aircraft Investigation of the inner core of hurricane Norbert. Part II: Mesoscale Distribution of Ice Particles, *J. Atmos. Sci.*, 49, 943-962.
- Hsu HM, Moncrieff MW, Tung WW and Liu C (2006) Multiscale temporal variability of warm-season precipitation over North America: Statistical analysis of radar measurements, *J. Atmos. Sci.*, 63, 2355-2368.
- Hubert P, Bendjoudi H, Schertzer D and Lovejoy S (1998) A multifractal explanation for rainfall intensity-duration-frequency curves, In: *Heavy Rains and Flash Floods*, edited by C. Llasat, P. Versace, and E. Ferrari, pp. 21 – 28, Natl. Res. Council., Group for Prev. from Hydrol. Disasters, Cosenza, Italy.
- Hurst HE (1951) Long term storage capacities of reservoirs, *Trans. Am. Soc. Civil Engrs.*, 116, 776–808.
- Hutchinson MF (1990) A point rainfall model based on a three-state continuous Markov occurrence process, *J. Hydrol.*, 114, 125-148.
- Jorgensen DP (1984a) Mesoscale and convective-scale characteristics of mature hurricanes. Part I: General observations by research aircraft, *J. Atmo. Sci.*, 41(8), 1268-1286.
- Jorgensen DP (1984b) Mesoscale and convective-scale characteristics of mature hurricanes. Part II: Inner core structure of hurricane Allen (1980), *J. Atmo. Sci.*, 41(8), 1287-1311.
- Joss J and Gori E (1978) Shapes of raindrop size distributions, *J. Appl. Meteor.*, 17, 1054-1061.
- Joss J and Waldvogel A (1990) Precipitation measurement and hydrology, In: *Radar in Meteorology*, Ed. Atlas D, Battan Memorial and 40th Anniversary Radar Meteorology Conference, American Meteorological Society, 577-597.
- Katsanos D, Lagouvardos K, Kotroni V and Argiriou A (2007) Relationship of lightning activity with microwave brightness temperatures and spaceborne radar reflectivity profiles in the Central and Eastern Mediterranean, *J. Appl. Meteor.*, 46, 1901-1912.
- Katz RW (1977a) An Application of chain-dependent processes to meteorology, *J. Appl. Prob.*, 14, 598-603.

- Katz RW (1977b) Precipitation as a chain-dependent process. *J. Appl. Meteorol.*, 16, 671-676.
- Katz RW and Parlange MB (1998) Overdispersion phenomenon in stochastic modeling of precipitation, *J. Climate*, 11, 591-601.
- Kavvas ML and Delleur JW (1981) A stochastic cluster model of daily rainfall sequences, *Water Resour. Res.*, 17(4), 1151-1160.
- Kepernt J (2001) The dynamics of boundary layer jets within the tropical cyclone core. Part I: Linear theory. *J. Atmo. Sci*, 58, 2469-2484.
- Kepernt J and Wang Y (2001) The dynamics of boundary layer jets within the tropical cyclone core. Part II: Nonlinear enhancement. *J. Atmos. Sci.*, 58, 2485-2501.
- Kessler E (1969) On the distribution and continuity of water substance in atmospheric circulation, *Meteor. Monogr.*, 10(32).
- Kidd C, Kniveton D and Barrett EC (1998) The advantages and disadvantages of statistically derived—empirically calibrated passive microwave algorithms for rainfall estimation, *J. Atmos. Sci.*, 55, 1576–1582.
- Kidder SQ and Vonder Haar TH (1995) *Satellite Meteorology*, Academic Press, London, UK.
- King PW, Hogg WD and Arkin PA (1995) The role of visible data in improving satellite rain-rate estimates, *J. Appl. Meteor.*, 34, 1608–1621.
- Klemes V (2000) Tall tales about tails of hydrological distributions, *J. Hydrol. Engineering*, 5(3), 227-231 and 232-239.
- Kolmogorov AN (1933) Grundbegriffe der Wahrscheinlichkeitsrechnung, *Ergebnisse der Math.* (2), Berlin; 2nd English Edition (1956): Foundations of the Theory of Probability, 84 pp. New York:Chelsea Publishing Company.
- Kolmogorov AN (1940) Wiener'sche Spiralen und einige andere interessante Kurven in Hilbert'schen Raum, *Dokl. Akad. Nauk URSS*, 26, 115–118.
- Kosambi DD (2005) *The Culture and Civilization of Ancient India in Historical Outline*, Vikas Publishing House Pvt Ltd (www.inae.org/metallurgy/archives_pdf/kosambi.pdf).
- Kotroni V and Lagouvardos K (2004) Evaluation of MM5 high-resolution real-time forecasts over the urban area of Athens, Greece, *J. Appl. Meteor.*, 43, 1666-1678.
- Kottegoda NT and Horder MA (1980) Daily flow model based on rainfall occurrences using pulses and a transfer function, *J. of Hydrol.*, 47, 215-234.
- Koutsoyiannis D (1999) A probabilistic view of Hershfield's method for estimating probable maximum precipitation, *Water Resour. Res.*, 35(4), 1313-1322.
- Koutsoyiannis D (2000a) Lecture notes on Hydrometeorology - Part 1, Edition 2, 157 pages, National Technical University of Athens, Athens (<http://www.itia.ntua.gr/en/docinfo/116/>).

- Koutsoyiannis D (2000b) A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series, *Water Resources Research*, 36 (6), 1519–1533.
- Koutsoyiannis D (2001) Coupling stochastic models of different time scales, *Water Resources Research*, 37 (2), 379–392.
- Koutsoyiannis D (2002) The Hurst phenomenon and fractional Gaussian noise made easy, *Hydrological Sciences Journal*, 47 (4), 573–595.
- Koutsoyiannis D (2003a) Climate change, the Hurst phenomenon, and hydrological statistics, *Hydrological Sciences Journal*, 48 (1), 3–24.
- Koutsoyiannis D (2003b) Rainfall disaggregation methods: Theory and applications, *Proceedings, Workshop on Statistical and Mathematical Methods for Hydrological Analysis*, edited by D. Piccolo and L. Ubertini, Rome, 1–23, Università di Roma "La Sapienza".
- Koutsoyiannis D (2004a) Statistics of extremes and estimation of extreme rainfall, 1, Theoretical investigation, *Hydrological Sciences Journal*, 49 (4), 575–590.
- Koutsoyiannis D (2004b) Statistics of extremes and estimation of extreme rainfall, 2, Empirical investigation of long rainfall records, *Hydrological Sciences Journal*, 49 (4), 591–610.
- Koutsoyiannis D (2005a) Uncertainty, entropy, scaling and hydrological stochastics, 1, Marginal distributional properties of hydrological processes and state scaling, *Hydrological Sciences Journal*, 50 (3), 381–404.
- Koutsoyiannis D (2005b) Uncertainty, entropy, scaling and hydrological stochastics, 2, Time dependence of hydrological processes and time scaling, *Hydrological Sciences Journal*, 50 (3), 405–426.
- Koutsoyiannis D (2006a) An entropic-stochastic representation of rainfall intermittency: The origin of clustering and persistence, *Water Resources Research*, 42 (1), W01401, doi:10.1029/2005WR004175.
- Koutsoyiannis D (2006b) Nonstationarity versus scaling in hydrology, *Journal of Hydrology*, 324, 239–254.
- Koutsoyiannis D (2007) A critical review of probability of extreme rainfall: principles and models, in *Advances in Urban Flood Management*, ed. Ashley R, Garvin S, Pasche E, Vassilopoulos A and Zevenbergen C, 139–166, Taylor and Francis, London.
- Koutsoyiannis D (2010) A random walk on water, *Hydrology and Earth System Sciences*, 14, 585–601.
- Koutsoyiannis D and Xanthopoulos T (1999) *Engineering Hydrology*, Edition 3, 418 pp., National Technical University of Athens, Athens (in Greek).
- Koutsoyiannis D and Onof C (2001) Rainfall disaggregation using adjusting procedures on a Poisson cluster model, *Journal of Hydrology*, 246, 109–122.

- Koutsoyiannis D and Montanari A (2007), Statistical analysis of hydroclimatic time series: Uncertainty and insights, *Water Resources Research*, 43 (5), W05429, doi:10.1029/2006WR005592.
- Koutsoyiannis D, Kozonis D and Manetas A (1998) A mathematical framework for studying rain–fall intensity-duration-frequency relationships, *J. Hydrol.*, 206(1-2), 118-135.
- Koutsoyiannis D, Onof C and Wheater HS (2003) Multivariate rainfall disaggregation at a fine timescale, *Water Resources Research*, 39 (7), 1173, doi:10.1029/2002WR001600.
- Koutsoyiannis D, Mamassis N and Tegos A (2007), Logical and illogical exegeses of hydrometeorological phenomena in ancient Greece, *Water Science and Technology: Water Supply*, 7 (1): 13–22.
- Koutsoyiannis D, Efstratiadis A, Mamassis N and Christofides A (2008) On the credibility of climate predictions, *Hydrological Sciences Journal*, 53 (4), 671–684.
- Krajewski WF and Smith JA (2002) Radar hydrology: Rainfall estimation, *Adv. Wat. Resour.*, 25(8-12), 1387-1394, doi: 10.1016/S0309-1708(02)00062-3.
- Kummerow C and Giglio L (1994a) A passive microwave technique for estimating rainfall and vertical structure information from space. Part I: Algorithm description, *J. Appl. Meteor.*, 33, 3–18.
- Kummerow C and Giglio L (1994b) A passive microwave technique for estimating rainfall and vertical structure information from space. Part II: Applications to SSM/I data, *J. Appl. Meteor.*, 33, 19–34.
- Kummerow C, Barnes W, Kozu T, Shiue J and Simpson J (1998) The Tropical Rainfall Measuring Mission (TRMM) sensor package, *J. Atmos. Oceanic Technol.*, 15, 809–817.
- Kummerow CD, Olson WS and Giglio L (1996) A simplified scheme for obtaining precipitation and vertical hydrometeor profiles from passive microwave sensors, *IEEE Trans. Geosci. and Remot. Sens.*, 34, 1213-1232.
- Kummerow C, Hong Y, Olson WS, Yang S, Adler RF, McCollum J, Ferraro R, Petty G, Shin DB and Wilheit TT (2001) The evolution of the Goddard profiling algorithm (GPROF) for rainfall estimation from passive microwave sensors, *J. Appl. Meteor.*, 40, 1801–1820.
- Kundu PK and Bell TL (2003) A Stochastic model of space-time variability of mesoscale rainfall: statistics of spatial averages, *Water Resour. Res.*, 39(12), doi:10.1029/2002WR001802.
- La Seur NE and Hawkins HF (1963) An analysis of hurricane Cleo (1958) based on data from research reconnaissance aircraft, *Mon. Wea. Rev.*, 91, 694-709.
- Lagouvardos K and Kotroni V (2005) Improvement of high resolution weather forecasts through humidity adjustment, based on satellite data. *Quart. J. Roy. Meteor. Soc.*, 131, 2695-2712.
- Lagouvardos K, Kotroni V, Koussis A, Feidas C, Buzzi A and Malguzzi P (2003) The meteorological model BOLAM at the National Observatory of Athens: assessment of two-year operational use, *J. Appl. Meteor.*, 42, 1667-1678.

- Landsea CW (2000) Climate variability of tropical cyclones, In: *Storms*, Vol. 1, Eds. Pielke R Jr and Pielke R Sr, Routledge, NY.
- Langmuir I (1948) The production of rain by a chain reaction in cumulus clouds at temperatures above freezing, *J. Meteor.*, 5, 175-192.
- Langousis A and Koutsoyiannis D (2006) A stochastic methodology for generation of seasonal time series reproducing over-year scaling behavior, *J. Hydrol.*, 322 (1-4), 138-154.
- Langousis A and Veneziano D (2007) Intensity-duration-frequency curves from scaling representations of rainfall, *Wat. Resour. Res.*, 43, doi: 10.1029/2006WR005245.
- Langousis A and Veneziano D (2009a) Theoretical model of rainfall in tropical cyclones for the assessment of long-term risk, *J. Geophys. Res.*, doi:10.1029/2008JD010080.
- Langousis A and Veneziano D (2009b) Long term rainfall risk from tropical cyclones in coastal areas, *Wat. Resour. Res.*, 45, W11430, doi:10.1029/2008WR007624.
- Langousis A, Veneziano D, Furcolo P, and Lepore C (2007) Multifractal rainfall extremes: theoretical analysis and practical estimation, *Chaos Solitons and Fractals*, doi:10.1016/j.chaos.2007.06.004.
- Langousis A, Veneziano D and Chen S (2008) Boundary layer model for moving tropical cyclones, In: *Hurricanes and Climate Change*, Eds. Elsner J and Jagger TH, Springer.
- Leith CE (1974) Theoretical skill of Monte Carlo forecasts, *Mon. Wea. Rev.*, 102, 409-418.
- Levizzani V, Amorati R and Meneguzzo F (2002) A review of satellite-based rainfall estimation methods, MUSIC – Multiple Sensor Precipitation Measurements, Integration, Calibration and Flood Forecasting, Deliverable 6.1, Italy.
- Lin YL, Farley RD and Orville HD (1983) Bulk parameterization of the snow field in a cloud model, *J. Climate Appl. Meteor.*, 22, 1065-1092.
- Liu G (2003) Precipitation, In: *Encyclopedia of Atmospheric Sciences*, Eds. Holton JR, Curry JA and Pyle JA, pp. 1972-1979, Academic Press, London, UK.
- Liu G and Curry JA (1992) Retrieval of precipitation from satellite microwave measurement using both emission and scattering, *J. Geophys. Res.*, 97, 9959–9974.
- Liu G and Curry JA (1993) Determination of characteristic features of cloud liquid water from satellite microwave measurements, *J. Geophys. Res.*, 98(D3), 5069–5092.
- Liu JY and Orville HD (1969) Numerical modeling of precipitation and cloud shadow effects on mountain-induced cumuli, *J. Atmos. Sci.*, 26, 1283-1298.
- Locatelli JD and Hobbs PV (1974) Fall speeds and masses of solid precipitation particles, *J. Geophys. Res.*, 79, 2185-2197.

- Lonfat M, Marks FD Jr and Chen SS (2004) Precipitation distribution in tropical cyclones using the Tropical Rainfall Measuring Mission (TRMM) microwave imager: A Global perspective, *Mon. Wea. Rev.*, 132, 1645-1660.
- Lonfat M, Rogers R, Marchok T and Marks FD Jr (2007) A parametric model for predicting hurricane rainfall, *Mon. Wea. Rev.*, 135, 3086-3097.
- Long AB (1971) Validity of the finite-difference droplet equation, *J. Atmos. Sci.*, 28, 210-218.
- Lorenz EN (1963) Deterministic Nonperiodic Flow, *J. Atmos. Sci.*, 20, 130–141.
- Lovejoy S and Austin GL (1979) The delineation of rain areas from visible and IR satellite data for GATE and mid-latitudes, *Atmosphere-Ocean*, 17(1), 77-92.
- Lovejoy S and Schertzer D (1995) Multifractals and rain, In: *New Uncertainty Concepts in Hydrology and Hydrological Modelling*, Ed. Kundzewicz ZW, Cambridge.
- Lull HW (1964) Ecological and silvicultural aspects. Sec. 6, in *Handbook of Applied Hydrology*, ed. Chow VT, New York, NY, McGraw-Hill.
- Lutgens FK and Tarbuck EJ (1992) *The Atmosphere: An Introduction to Meteorology*, 5th Edition, Prentice-Hall, Englewood Cliffs, NJ.
- Mandelbrot BB and van Ness JW (1968) Fractional Brownian Motion, fractional noises and applications, *SIAM Review*, 10, 422 – 437.
- Marks FD (1985) Evolution of the structure of precipitation in hurricane Allen (1980), *Mon. Wea. Rev.*, 113, 909-930.
- Marks FD (2003) Hurricanes, In: *Encyclopedia of Atmospheric Sciences*, Eds. Holton JR, Curry JA and Pyle JA, pp. 942-966, Academic Press, London, UK.
- Marks FD and Houze RA Jr (1987) Inner-core structure of hurricane Alicia from airborne doppler radar observations, *J. Atmos. Sci.*, 44, 1296-1317.
- Marks FD, RA Houze Jr and Gamache JF (1992) Dual-aircraft investigation of the inner core of hurricane Norbert. Part I: Kinematic Structure, *J. Atmos. Sci.*, 49, 919-942.
- Marshall JS and Palmer WM (1948) The distribution of raindrops with size, *J. Meteor.*, 5, 165-166.
- Marshall JS, Hirschfeld W and Gunn KLS (1955) Advances in radar weather, *Adv. Geophys.*, 2, 1-56.
- Mason BJ (1971) *The Physics of Clouds*, Clarendon Press, Oxford.
- Mass CF, Ovens D, Westrick K and Colle B (2002) Does increasing horizontal resolution produce more skillful forecasts? *Bull. Amer. Meteorol. Soc.*, 83, 407-430.
- Meehl GA, Stocker TF, Collins WD, Friedlingstein P, Gaye AT, Gregory JM, Kitoh A, Knutti R, Murphy JM, Noda A, Raper SCB, Watterson IG, Weaver, AJ and Zhao Z-C (2007) Global climate projections. In: *Climate Change 2007: The Physical Science Basis*. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change

- (ed. by Solomon S., Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M and Miller HL), 747–845, Cambridge University Press, Cambridge, UK.
- Menabde M and Sivapalan M (2000) Modelling of rainfall time series and extremes using bounded random cascades and levy-stable distributions, *Wat. Resour. Res.*, 36(11), 3293-3300.
- Menabde M, Harris D, Seed A, Austin G and Stow D (1997) Multiscaling properties of rainfall and bounded random cascades, *Wat. Resour. Res.*, 33(12), 2823-2830.
- Miller BL (1958) Rainfall rates in Florida hurricanes, *Mon. Wea. Rev.*, 86(7), 258-264.
- Molinari J, Moore PK, Idone VP, Henderson RW and Saljoughy AB (1994) Cloud-to-ground lightning in hurricane Andrew, *J. Geophys. Res.*, 99, 16 665-16 676.
- Molinari J, Moore PK and Idone VP (1999) Convective structure of hurricanes as revealed by lightning locations, *Mon. Wea. Rev.*, 127, 520-534.
- Mora RD, Bouvier C, Neppel L and Niel H (2005) Approche régionale pour l'estimation des distributions ponctuelles des pluies journalières dans le Languedoc-Roussillon (France), *Hydrol. Sci. J.*, 50(1), 17-29.
- Neyman J and Scott EL (1952) A theory of the spatial distribution of galaxies, *Astrophysical Journal*, 116, 144-163.
- Northrop P (1996) Modelling and statistical analysis of spatial-temporal rainfall fields, PhD thesis, Department of Statistical Science, University College London.
- Northrop PJ (1998) A clustered spatial-temporal model of rainfall, *Proc. R. Soc. Lond.*, A454, 1875-1888.
- Nykanen DK and Harris D (2003) Orographic influences on the multiscale statistical properties of precipitation, *J. Geophys. Res.*, 108(D8), doi:10.1029/2001JD001518.
- Olson DA, Junker NW and Korty B (1995) Evaluation of 33 years of quantitative precipitation forecasting at the NMC, *Wea. Forecasting*, 10, 498-511.
- Olsson J (1995) Limits and characteristics of the multifractal behavior of a high-resolution rainfall time series, *Nonlinear processes in Geoph.*, 2, 23-29.
- Olsson, J (1998) Evaluation of a scaling cascade model for temporal rainfall disaggregation, *Hydrol. Earth Syst. Sci.*, 2, 19-30.
- Onof C and Wheeler HS (1993) Modelling of British rainfall using a Random Parameter Bartlett-Lewis Rectangular Pulse Model, *J. Hydrol.*, 149, 67-95.
- Onof C and Wheeler HS (1994) Improvements to the modeling of British rainfall using a Modified Random Parameter Bartlett-Lewis Rectangular Pulses Model, *J. Hydrol.*, 157, 177-195.

- Onof C, Chandler RE, Kakou A, Northrop P, Wheeler HS and Isham V (2000) Rainfall modelling using Poisson-cluster processes: a review of developments, *Stochastic Environmental Research and Risk Assessment*, 14, 384-411.
- Ooyama KV (1982) Conceptual evolution of theory and modeling of the tropical cyclone, *J. Meteor. Soc. Jpn.*, 60, 369-379.
- Orlanski I (1975) A rational subdivision of scales for atmospheric processes, *Bull. Am. Meteor. Soc.*, 56(5), 527-530.
- O'Sullivan F, Wash CH, Stewart M and Motell CE (1990) Rain estimation from infrared and visible GOES satellite data, *J. Appl. Meteor.*, 29, 209-223.
- Over TM and Gupta VK (1996) A space-time theory of mesoscale rainfall using random cascades, *J. Geophys. Res.*, 101(D21), 26 319- 26 331.
- Papalexiou SM and Koutsoyiannis D (2006) A probabilistic approach to the concept of probable maximum precipitation, *Advances in Geosciences*, 7, 51-54.
- Papalexiou SM and Koutsoyiannis D (2008) Probabilistic description of rainfall intensity at multiple time scales, *IHP 2008 Capri Symposium: "The Role of Hydrology in Water Resources Management"*, Capri, Italy, UNESCO, International Association of Hydrological Sciences (www.itia.ntua.gr/en/docinfo/884/).
- Papalexiou SM and Koutsoyiannis D (2009) Ombrian curves: from theoretical consistency to engineering practice, *8th IAHS Scientific Assembly / 37th IAH Congress*, Hyderabad, India, 2009 (www.itia.ntua.gr/en/docinfo/926/).
- Papoulis A (1991) *Probability, Random Variables, and Stochastic Processes*, 3rd ed., New York: McGraw-Hill.
- Pegram GGS (1980) An auto-regressive model for multi-lag Markov chains, *J. Appl. Probab.*, 17, 350-362.
- Pegram GGS and Seed AW (1998) The feasibility of stochastically modelling the spatial and temporal distribution of rainfields, WRC Report No. 550/1/98 to the Water Research Commission, Pretoria, South Africa.
- Perica S and Foufoula-Georgiou E (1996a) Linkage of scaling and thermodynamic parameters of rainfall: Results from midlatitude mesoscale convective systems, *J. Geophys. Res.*, 101(D3), 7431-7448.
- Perica S and Foufoula-Georgiou E (1996b) Model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions, *J. Geophys. Res.*, 101(D21), 26347-26361.
- Petty GW (1994a) Physical retrievals of over-ocean rain rate from multichannel microwave imaging. Part I: Theoretical characteristics of normalized polarization and scattering indices, *Meteor. Atmos. Phys.*, 54, 79-99.

- Petty GW (1994b) Physical retrievals of over-ocean rain rate from multichannel microwave imaging. Part II: Algorithm implementation, *Meteor. Atmos. Phys.*, 54, 101–121.
- Petty GW (1995) The status of satellite-based rainfall estimation over land, *Remote Sens. Environ.* 51, 125-137.
- Petty GW (2001a) Physical and microwave radiative properties of precipitating clouds. Part I: Principal component analysis of observed multichannel microwave radiances in tropical stratiform rainfall, *J. Appl. Meteor.*, 40, 2105–2114.
- Petty GW (2001b) Physical and microwave radiative properties of precipitating clouds. Part II: A parametric 1D rain-cloud model for use in microwave radiative transfer simulations, *J. Appl. Meteor.*, 40, 2115–2129.
- Pinsky M and Khain A (2004) Collisions of small drops in a turbulent flow. Part II: Effects of flow accelerations, *J. Atmo. Sci.*, 61, 1926-1939.
- Pinsky M, Khain A and Shapiro M (1999) Collisions of small drops in a turbulent flow. Part I: Collision efficiency. Problem Formulation and Preliminary Results, *Amer. Meteor. Soc.*, 56, 2585-2600.
- Pinsky M, Khain A and Shapiro M (2000) Stochastic effects of cloud droplet hydrodynamic interaction in a turbulent flow, *Atmos. Res.*, 53, 131-169.
- Poincaré H (1908), *Science et Méthode*; reproduced in Poincaré H (1956), *Chance, The World of Mathematics*, New York: Simon & Schuster.
- Powell MD (1990) Boundary layer structure and dynamics in outer hurricane rainbands. Part I: Mesoscale rainfall and kinematic structure, *Mon. Wea. Rev.*, 118, 891-917.
- Popper K (1982) *Quantum Physics and the Schism in Physics*, Unwin Hyman, London.
- Probert-Jones JR (1962) The radar equation in meteorology. *Quart. J. Roy. Meteor. Soc.*, 88, 485–495.
- Ramage C (1971) *Monsoon Meteorology*, International Geophysics Series, Vol. 15, Academic Press, San Diego, CA.
- Randall DA, Wood RA, Bony S, Colman R, Fichefet T, Fyfe J, Kattsov V, Pitman A, Shukla J, Srinivasan J, Stouffer RJ, Sumi A, and Taylor KE (2007), Climate Models and Their Evaluation, in Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, and Miller HL (ed.): *Climate Change 2007: The Physical Science Basis*, Contribution of Working Group I to the Fourth Assessment Report of the IPCC, Cambridge University Press: Cambridge, UK and New York, NY, USA, 589–662.
- Reisner J, Rasmussen RJ and Bruintjes RT (1998) Explicit forecasting of supercooled liquid water in winter storms using the MM5 mesoscale model, *Quart. J. Roy. Meteor. Soc.*, 124B, 1071-1107.
- Renno N and Ingersoll A (1996) Natural convection as a heat engine: A theory for CAPE, *J. Atmo. Sci.*, 53(4), 572-585.

- Richards F and Arkin P (1981) On the relationship between satellite-observed cloud cover and precipitation, *Mon. Wea. Rev.*, 109, 1081–1093.
- Robert CP (2007) *The Bayesian Choice, From Decision-Theoretic Foundations to Computational Implementation*, Springer.
- Robertson D (1974) Monte Carlo simulations of drop growth by accretion, *J. Atmos. Sci.*, 31, 1344-1350.
- Rodriguez-Iturbe I, Gupta VK and Waymire E (1984) Scale considerations in the modeling of temporal rainfall, *Water Resour. Res.*, 20(11), 1611-1619.
- Rodriguez-Iturbe I, Cox DR, and Isham V (1987) Some models for rainfall based on stochastic point processes, *Proc. R. Soc. Lond.*, A 410, 269-298.
- Rodriguez-Iturbe I, Cox DR, and Isham V (1988) A point process model for rainfall: Further developments, *Proc. R. Soc. Lond.*, A 417, 283-298.
- Rogers RR and Yau MK (1996) *A Short Course in Cloud Physics*, Third Edition, Butterworth Heinemann, USA.
- Roldan J and Woolhiser DA (1982) Stochastic daily precipitation models, 1, A comparison of occurrence processes, *Water Resour. Res.*, 18(5), 1451-1459.
- Rossi F, Fiorentino M and Versace P (1984) Two-component extreme value distribution for flood frequency analysis, *Water Resour. Res.*, 20(7), 847-856.
- Rotunno R (1986) Tornadoes and tornadogenesis, Ch. 18, pp. 414-436, In: *Mesoscale Meteorology and Forecasting*, Ed. Ray PS, *Amer. Meteor. Soc.*, Boston, USA.
- Ryan BF (1996) On the global variation of precipitating layer clouds, *Bull. Amer. Meteor. Soc.*, 77, 53-70.
- Ryan BT, Blau HH, Thuna PC, Cohen ML and Roberts GD (1972) Cloud microstructure as determined by optical cloud particle spectrometer, *J. Appl. Meteor.*, 11, 149-156.
- Schaefer MG, Barker BL, Taylor GH and Wallis JR (2006) Regional precipitation-frequency analysis and spatial mapping for 24-hour and 2-hour durations for Wahington State, *Geophysical Research Abstracts*, Vol. 8, 10899.
- Schertzer D and Lovejoy S (1987) Physical modeling and analysis of rain and clouds by anisotropic scaling of multiplicative processes, *J. Geophys. Res.*, 92, 9693-9714.
- Schertzer D and Lovejoy S (1989) Generalized scale invariance and multiplicative processes in the atmosphere, *Pageoph.*, 130, 57-81.
- Schmitt F, Vannitsem S and Barbosa A (1998) Modeling of rainfall time series using two-state renewal processes and multifractals, *J. Geophys. Res.*, 103(D18)92, 23181-23193.
- Scorer RS and Wexler H (1963) *A colour guide to clouds*, Pergamon Press, Oxford.

- Scott WT (1968) On the connection between the telford and kinetic equation approaches to droplet coalescence theory, *J. Atmos. Sci.*, 25, 871-873.
- Scott WT (1972) Comments on "Validity of the finite-difference droplet collection equation", *J. Atmos. Sci.*, 29, 593-594.
- Seltzer MA, Passarelli RE and Emanuel KA (1985) The possible role of symmetric instability in the formation of precipitation bands, *J. Atmos. Sci.*, 42, 2207-2219.
- Sempere-Torres D, Porrà JM and Creutin JD (1994) A general formulation for raindrop size distribution, *J. Appl. Meteorol.*, 33, 1494-1502.
- Sempere-Torres D, Porrà JM and Creutin JD (1998) Experimental evidence of a general description for raindrop size distribution properties, *J. Geophys. Res.*, 103(D), 1785–1797.
- Sharma A and O'Neill R (2002) A nonparametric approach for representing interannual dependence in monthly streamflow sequences, *Water Resour. Res.*, 38(7), 1100, doi: 10.1029/2001WR000953.
- Shapiro LJ (1983) The asymmetric boundary layer flow under a translating hurricane, *J. Atmos. Sci.*, 40, 1984-1998.
- Simpson J, Adler RF and North GR (1988) Proposed Tropical Rainfall Measuring Mission (TRMM) satellite, *Bull. Amer. Meteor. Soc.*, 69: 278-295.
- Sisson SA, Pericchi LR and Coles SG (2006) A case for a reassessment of the risks of extreme hydrological hazards in the Caribbean, *Stoch. Environ. Res. Risk Assess.*, 20, 296–306.
- Skamarock W, Klemp J, Dudhia J, Gill D, Barker D, Wang W and Powers J (2005) *A description of the advanced Research WRF Version 2*, Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research, Boulder, Colorado, USA.
- Slingo J (2003) Monsoon, In: *Encyclopedia of Atmospheric Sciences*, Eds. Holton JR, Curry JA and Pyle JA, pp. 1365-1370, Academic Press, London, UK.
- Slinn WGN (1975) Atmospheric aerosol particles in surface-level air, *Atmospheric Environment*, 9, 763-764.
- Small MJ and Morgan DJ (1986) The relationship between a continuous-time renewal model and a discrete Markov chain model of precipitation occurrence, *Water Resour. Res.*, 22(10), 1422-1430.
- Smith JA (1993) Precipitation, In: *Handbook of Applied Hydrology*, Ed. Maidment DA, Ch. 3, pp. 3.1-3.47, McGraw-Hill, New York.
- Smith RK (1968) The surface boundary layer of a hurricane, *Tellus*, 20, 473-484.
- Smith RK (2000) The role of cumulus convection in hurricanes and its representation in hurricane models, *Rev. Geophys.*, 38, 465-489.

- Smith RK (2003) A Simple model of hurricane boundary layer, *Quart. J. Roy. Met. Soc.*, 129, 1007-1027.
- Spencer RW, Goodman HM and Hood RE (1989) Precipitation retrieval over land and ocean with the SSM/I: Identification and characteristics of the scattering signal, *J. Atmos. Oceanic Technol.*, 6, 254–273.
- Srikanthan R and McMahon TA (2001) Stochastic generation of annual, monthly and daily climate data: A review, *Hydrology and Earth System Sciences*, 5(4), 653–670.
- Stephens GL (1994) *Remote Sensing of the Lower Atmosphere: An Introduction*, Oxford University Press, New York, USA.
- Stewart SR (2002) Tropical Cyclone Report, Tropical Storm Allison, 5-17 June 2001, US National Oceanic and Atmospheric Administration, <http://www.nhc.noaa.gov/2001allison.html>.
- Stoelinga MT, Woods CP, Locatelli JD and Hobbs PV (2005), On the representation of snow in bulk microphysical parameterization schemes, WRF/MM5 Users' Workshop, June 2005.
- Stout JS, Martin DW and Sikdar DN (1979) Estimating GATE rainfall with geosynchronous satellite images, *Mon. Wea. Rev.*, 107, 585–598.
- Sutcliffe JV (1978) *Methods of Flood Estimation, A Guide to Flood Studies Report*, Report 49, Institute of Hydrology, Wallingford, UK.
- Telford JW (1955) A new aspect of coalescence theory, *J. Meteor.*, 12, 436-444.
- Telford JW and Chai SK (1980) A new aspect of condensation theory, *Pure Appl. Geophys.*, 118, 720-742.
- Telford JW and Wagner PB (1981) Observations of condensation growth determined by entity type mixing, *Pure Appl. Geophys.*, 119, 934-965.
- Tessier Y, Lovejoy S and Schertzer D (1993) Universal multifractals in rain and clouds: Theory and observations, *J. Appl. Meteor.*, 32, 223-250.
- Tessier Y, Lovejoy S, Hubert P, Schertzer D, and Pecknold S (1996) Multifractal analysis and modeling of rainfall and river flows and scaling, causal transfer functions, *J. Geophys. Res.*, 101(D21), 26427–26440.
- Testik FY and Barros AP (2007) Toward elucidating the microstructure of warm rainfall: A survey, *Rev. Geophys.*, 45, RG2003, doi:10.1029/2005RG000182.
- Thompson G, Rasmussen RM and Manning K (2004) Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis, *Mon. Wea. Rev.*, 132, 519–542.
- Todorovic P and Woolhiser DA (1975) A stochastic model of n-day precipitation, *J. Appl. Meteorol.*, 14, 17-24.

- Todorovic P and Zelenhasic E (1970) A stochastic model for flood analysis, *Water Resour. Res.*, 6(6), 1641-1648.
- Tsonis A (1987) Determining rainfall intensity and type from GOES imagery in the midlatitudes, *Remot. Sens. Env.*, 21, 29-36.
- Tsonis A and Isaac G (1985) On a new approach for instantaneous rain area delineation in the midlatitudes using GOES data, *J. Appl. Meteor.*, 24, 1208-1218.
- Turpeinen OM, Abidi A and Belhouane W (1987) Determination of rainfall with the ESOC precipitation index, *Mon. Wea. Rev.*, 115, 2699-2706.
- Twomey S (1964) Statistical effects in the evolution of a distribution of cloud droplets by coalescence, *J. Atmos. Sci.*, 21, 553-557.
- Twomey S (1966) Computation of rain formation by coalescence, *J. Atmos. Sci.*, 23, 405-411.
- Uijlenhoet R (1999) *Parameterization of Rainfall Microstructure for Radar Meteorology and Hydrology*, Doctoral Dissertation, Wageningen University, The Netherlands.
- Uijlenhoet R (2001) Raindrop size distributions and radar reflectivity - rain rate relationships for radar hydrology, *Hydrol. Earth Syst. Sci.*, 5, 615-627.
- Uijlenhoet R (2008) Precipitation physics and rainfall observation, in *Climate and the Hydrological Cycle*, eds. Bierkens MFP, Dolman AJ and Troch PA, IAHS Special Publication 8, 59-97, Wallingford, UK: IAHS Publications.
- Uijlenhoet R and Berne A (2008) Stochastic simulation experiment to assess radar rainfall retrieval uncertainties associated with attenuation and its correction, *Hydrol. Earth Syst. Sci.*, 12, 587-601.
- Uijlenhoet R and Sempere-Torres D (2006) Measurement and parameterization of rainfall microstructure, *J. Hydrol.*, 328(1-2), 1-7, ISSN 0022-1694, doi: 10.1016/j.jhydrol.2005.11.038.
- Ulbrich C (1983) Natural variations in the analytical form of the raindrop size distribution, *J. Clim. Appl. Meteorol.*, 22(10), 1764-1775.
- Ulbrich CW and Atlas D (1998) Rainfall microphysics and radar properties: Analysis methods for drop size spectra, *J. Appl. Meteor.*, 37, 912-923.
- US Department of the Interior, Bureau of Reclamation (1977) *Design of Arch Dams*. US Government Printing Office, Denver, Colorado, USA.
- US Department of the Interior, Bureau of Reclamation (1987) *Design of Small Dams*, third edn. US Government Printing Office, Denver, Colorado, USA.
- National Oceanic and Atmospheric Administration (2009) About the Radar Operations Center, <http://www.roc.noaa.gov/WSR88D/About.aspx>.

- Velghe T, Troch PA, De Troch FP and Van de Velde J (1994) Evaluation of cluster-based rectangular pulses point process models for rainfall, *Water Resour. Res.*, 30(10) 2847-2857.
- Veneziano D (1999) Basic properties and characterization of stochastically self-similar processes in R^D , *Fractals*, 7(1), 59-78.
- Veneziano D and Furcolo P (2002) Multifractality of rainfall and intensity-duration-frequency curves, *Wat. Resour. Res.*, 38(12), 1306-1317.
- Veneziano D and Langousis A (2005) The areal reduction factor a multifractal analysis, *Wat. Resour. Res.*, 41, W07008, doi:10.1029/2004WR003765.
- Veneziano D and Langousis A (2010) Scaling and fractals in hydrology, In: *Advances in Data-based Approaches for Hydrologic Modeling and Forecasting*, Ed. Sivakumar B, World Scientific, Chapter 4, 145p.
- Veneziano D, Furcolo P and Iacobellis V (2006a) Imperfect scaling of time and space-time rainfall, *J. Hydrol.*, 322(1-4), 105-119.
- Veneziano D, Langousis A, and Furcolo P (2006b) Multifractality and rainfall extremes: a review, *Wat. Resour. Res.*, 42, W06D15, doi:10.1029/2005WR004716.
- Veneziano D, Lepore C, Langousis A and Furcolo P (2007) Marginal methods of intensity-duration-frequency estimation in scaling and nonscaling rainfall, *Wat. Resour. Res.*, 43, W10418, doi:10.1029/2007WR006040.
- Veneziano D, Langousis A and Lepore C (2009) New Asymptotic and pre-asymptotic results on rainfall maxima from multifractal theory, *Wat. Resour. Res.*, 45, W11421, doi:10.1029/2009WR008257.
- Venugopal V, Roux SG, Fofoula-Georgiou E and Arneodo A (2006) Revisiting multifractality of high resolution temporal rainfall using a wavelet-based formalism, *Wat. Resour. Res.*, 42, W06D14, doi:10.1029/2005WR004489.
- Verhoest N, Troch PA and De Troch FP (1997) On the applicability of Bertlett-Lewis rectangular pulses models in the modeling of design storms at a point, *J. Hydrol.*, 202, 108-120.
- Vickery PJ, Skerlj PF and Twisdale LA (2000) Simulation of hurricane risk in the US using empirical track model, *J. of Structural Engineering*, 126(10), 1222-1237.
- Viltard N, Burlaud C and Kummerow C (2006) Rain retrieval from TMI Brightness temperature measurements using a TRMM PR – based data base, *J. Applied Meteor. and Climat.*, 45, 455-466.
- Vohl O, Mitra S, Wurzler S and Pruppacher H (1999) A wind tunnel study on the effects of turbulence on the growth of cloud drops by collision and coalescence, *J. Atmo. Sci.*, 56, 4088-4099.
- Wallace JW, and Hobbs PV (1977) *Atmospheric Science, An Introductory Survey*, Academic Press, San Diego, Ca.

- Wang B, Ding Q and Jhun JG (2006) Trends in Seoul (1778-2004) summer precipitation, *Geophysical Research Letters*, 33 (15), L15803.
- Wang B, Jhun JG and Moon BK (2007) Variability and singularity of Seoul, South Korea, rainy season (1778--2004), *J. Clim.*, 20(11), 2572-2580.
- Waymire E and Gupta VK (1981a) The mathematical structure of rainfall representations, 1, A review of the stochastic rainfall models, *Water Resour. Res.*, 17(5), 1261-1272.
- Waymire E and Gupta VK (1981b) The mathematical structure of rainfall representations, 2, A review of the theory of point processes, *Water Resour. Res.*, 17(5), 1273-1285.
- Waymire E and Gupta VK (1981c) The mathematical structure of rainfall representations, 3, Some applications of the point process theory to rainfall processes, *Water Resour. Res.*, 17(5), 1287-1294.
- Weisman ML and Klemp JB (1982) The dependence of numerically simulated convective storms on vertical wind shear and buoyancy, *Mon. Wea. Rev.*, 110, 504-520.
- Weisman ML and Klemp JB (1984) The structure and classification of numerically simulated convective storms in directionally varying wind shears, *Mon. Wea. Rev.*, 112, 2479-2498.
- Weisman ML and Klemp JB (1986) Characteristics of isolated convective storms, Ch. 15, pp. 331-358, In: *Mesoscale Meteorology and Forecasting*, Ed. Ray PS, Amer. Meteor. Soc., Boston, USA.
- Wilheit T, Adler R, Avery S, Barrett E, Bauer P, Berg W, Chang A, Ferriday J, Grody N, Goodman S, Kidd C, Kniveton D, Kummerow C, Mugnai A, Olson W, Petty G, Shibata A and Smith E (1994) Algorithms for the retrieval of rainfall from passive microwave measurements, *Remote Sensing Reviews*, 11, 163-194.
- Willoughby HE (1990) Gradient balance in tropical cyclones, *J. Atmos. Sci.*, 47, 265-274.
- Willoughby HE (1991) Reply, *J. Atmos. Sci.*, 48, 1209-1212.
- Willoughby HE, Marks FD Jr and Feinberg RJ (1984) Stationary and moving convective bands in hurricanes, *J. Atmos. Sci.*, 41(22), 3189-3211.
- Wisner C, Orville HD and Myers C (1972) A numerical model of hail-bearing cloud, *J. Atmos. Sci.*, 29, 1160-1181.
- World Meteorological Organization (1983) *Guide to Meteorological Instruments and Methods of Observation*, Publication 8, Fifth Edition, World Meteorological Organization, Geneva.
- World Meteorological Organization (1986), *Manual for Estimation of Probable Maximum Precipitation*, Operational Hydrology Report 1, 2nd edition, Publication 332, World Meteorological Organization, Geneva.
- Worsley KJ (1983) Testing a two-phase multiple regression, *Technometrics* 25(1), 35-41.
- Yanai M (1964) Formation of tropical cyclones, *Rev. Geophys.*, 2, p. 367.