Monsoons govern the climate of much of the tropics, where billions of people live. After months of winter drought, the long-awaited summer rains restore water to parched lands to support vegetation and crops. This article explains the causes of the seasonal occurrence of monsoons and describes their evolution. Finally, the present and past variations of the monsoons are discussed, as well as the expected future development.

1. The importance of monsoons
The word **monsoon** conjures up images of dark clouds, torrential rains, flooded countryside, cities paralyzed by water. In the words of Pierre Gouru [1], “physical conditions are necessary for a complete understanding of geography. Without a thorough science of the physical environment, the examination of the human elements of the landscape is exposed to an inaccurate interpretation of the relations existing between humans and their natural support”. Thus, **Monsoon Asia** [2] groups together the densely populated countries of the southern and eastern seaboard of Asia, from India to China and Japan, which are subject to this particular meteorological regime that is deeply rooted in their cultures, their societies and their economies. In these regions of fertile plains and deltas where winter drought alternates with abundant summer rainfall, people have founded rich civilizations since ancient times. The cultivation of rice, the staple food of these populations, is totally dependent on the monsoon rains (Figure 1). When the rains are early or late, deficient or excessive, the rhythm of preparation of dykes and fields, sowing, transplanting and harvesting are disrupted, often with serious socio-economic consequences.

Civilizations are deadly and climatic factors have sometimes been held responsible [3]. Several studies have shown the link between **major upheavals** such as the end of the Indus civilization around 1300BC, the decline of the Tang dynasty in China around 900, the fall of the Khmer empire around 1330, and periods of deficient monsoon. In the Americas, similar decreases in summer rains essential to agricultural resources may have caused the fall of the Maya civilization in the tenth century and the mysterious exodus of the Anasazi around 1300. More recently in Africa, the severe drought in the Sahel between 1970 and 1990, with rainfall falling by about half, caused food insecurity, population movements to cities, and economic, social and political instability.

The monsoon is a **meteorological phenomenon of great magnitude**, whose beneficial or harmful consequences affect billions of people. For centuries, many studies have been devoted to understanding the physical mechanisms behind its occurrence and its disturbances, to monitor its evolution in real time, and to try to predict its evolution in the short, medium and long term. We try here to present the main aspects.

### 2. Main characteristics of monsoons

The monsoon [4] is a particular meteorological regime which concerns not only Asia, but also other regions of the **tropical and subtropical domain**. Two criteria must be verified to qualify a monsoon:

(i) the alternation of a **wet season**, during the summer of the hemisphere concerned (June-July-August in the north, December-January-February in the south), and a **dry season** during the winter;

(ii) the reversal of the **direction of the prevailing winds**, blowing in the lower layers from the ocean to the continent in summer and from the continent to the ocean in winter.
Figure 2 shows contrasting seasonal situations. In boreal summer (Figure 2a), the main rainfall areas are located in Central America, sub-Saharan Africa, India and Southeast Asia, with an extension over the equatorial western Pacific. They are fed by trans-equatorial flows that load with moisture as they pass over the ocean. These low-level winds are driven towards thermal lows within the warmest continental regions. During the austral summer (Figure 2b), rainfall areas are found over South America, southern Africa, Indonesia and northern Australia, similarly fed by cross-equatorial winds directed towards continental thermal lows.

The above conditions are verified over southwestern North America, sub-Saharan Africa, India, Southeast Asia and northern Australia. It is more difficult to identify a strict monsoon regime over South America and Southern Africa due to less seasonal changes in prevailing winds. The winter monsoon over the Indian Ocean appears here as the northern part of the trans-equatorial flow directed towards the large area of precipitation extending from Indonesia to northern Australia and the South Pacific. To some extent, oceanic monsoons can also be evoked over the southern Indian Ocean and the northwestern Pacific.

3. Seasonal thermal contrast and inter-hemispheric transfer
To explain these significant seasonal variations in atmospheric circulation, we must go back to the basics of the **Earth's energy balance**. First of all, let us recall that the Earth is an approximately spherical planet whose own daily rotation axis is inclined by 23.44° with respect to the axis of its annual rotation around the Sun. The situation would be much simpler if this inclination were zero. The maximum **solar irradiance** would then occur throughout the year at the equator where the ascending branch of the **Hadley circulation** would generate a perpetual rain belt. On our tilted Earth, the energy balance varies greatly in each hemisphere between summer and winter and as a function of latitude (Figure 3).

Almost all the energy arriving on Earth comes from the Sun in the form of **visible radiation**, with wavelengths between 0.38 and 0.75 µm. The power actually received by the surface (Figure 3a) depends not only on the latitude, with a maximum at the **sub-solar point** [6] at 23.44°N at the June solstice and at 23.44°S at the December solstice, but also on the **albedo** [7] which quantifies the proportion of incident radiation directly reflected back to space. Thus, despite a very long daylight period, the Arctic and Antarctic regions covered by highly reflective snow and ice (with an albedo close to 1) receive relatively little solar radiation. The slightly higher values in December-January-February are due to a closer distance to the Sun than in June-July-August.

The Earth loses energy through **thermal radiation** from its surface and atmosphere to space (Figure 3b). Physical laws indicate that this radiation depends on the temperature (expressed in Kelvin) of the emitting body to the power of 4, and that the wavelength of the maximum is inversely proportional to this temperature. Due to the surface and atmospheric temperatures, terrestrial radiation is emitted in the **infrared** [8] with a maximum around 10 µm wavelength, and is more intense in tropical and equatorial regions, which are warmer than the polar regions. The relative minima around 10°S in December-January-February and 10°N in June-July-August are due to the dense cloud masses of the **Inter-Tropical Convergence Zone** [9], with very high and therefore very cold tops. However, we note that the variation in latitude is much less marked than for incident solar radiation (Figure 3a).

The balance is **positive** when a region receives more energy from solar radiation than it loses from terrestrial radiation, it is negative in the opposite case. Thus, we can see (Figure 3c) that during the boreal summer (June-July-August), latitudes between 10°S and 70°N gain energy, while latitudes between 10°S and 80°S and above 70°N lose energy. The situation is reversed during the austral summer (December-January-February). But temperatures in these regions do not rise or fall as quickly as this simple radiation balance would dictate. In fact, the excess heat received in the summer hemisphere is exported to the winter hemisphere by the movements of the atmosphere and ocean currents, which considerably reduces seasonal temperature contrasts (See: The **slow and powerful ocean circulation** and **Atmospheric circulation: its organization**).
The atmospheric situation at solstice can be represented by the simple model of a tank with two parts separated by a removable rigid wall (Figure 4). The left part representing the atmosphere above the winter hemisphere is filled with a cold fluid, the right part representing the summer hemisphere is filled with the same fluid at higher temperature. Hydrostatic equilibrium implies that the pressure at the bottom of the tank is proportional to the weight of the overlying column of fluid. Since the density of a fluid decreases (becomes less dense) as its temperature increases, the pressure at the bottom of the warmer right side is lower than that of the colder left side (Figure 4a).

When the removable wall is removed, the horizontal pressure force, directed towards the lower pressure, causes the cold fluid to move to the right in the bottom of the tank (Figure 4b). To compensate for this transfer, an equal volume of hot fluid is drawn to the left at the top of the tank. The pressure difference between the two parts of the tank gradually diminishes with this exchange of mass and heat. Horizontal pressure equilibrium and the end of fluid movement are reached when a stable stratification is achieved with the cold fluid at the bottom of the tank and the hot fluid above (Figure 4c).

In the case of the Earth’s atmosphere, the differential of received solar power and emitted terrestrial power (Figure 3c) maintains a warming of the summer hemisphere and a cooling of the winter hemisphere. This perpetuates an inter-hemispheric pressure difference and the induced intermediate circulation (Figure 4b) with a trans-equatorial flow directed from the winter hemisphere to the summer hemisphere in the lower layers of the atmosphere and a compensatory flow of opposite direction in the upper troposphere. Thanks to this circulation, the excess heat received in summer is transferred aloft to the winter hemisphere, which in exchange exports relatively cold air to the summer hemisphere near the surface.

The Earth rotates on itself in 24 hours, and this has consequences for the trans-equatorial circulation resulting from the temperature difference between the summer and winter hemispheres. The Coriolis force caused by this rotation acts on all moving bodies on the surface of the globe. It deflects their trajectories to the right in the Northern Hemisphere and to the left in the Southern Hemisphere. Its intensity is maximum at the poles and zero at the equator. The pressure force carries air from the high (cold) pressure of the winter hemisphere to the low (warm) pressure, but the wind in the Earth's atmosphere does not follow a straight path between these two regions.
During the **boreal summer**, the wind from the southern hemisphere undergoes a decreasing westward deviation as it approaches the equator. As it passes in the Northern Hemisphere, it is increasingly deflected eastward away from the equator (Figure 5a). During the **austral summer**, the wind from the Northern Hemisphere is first directed to the southwest and then to the southeast after crossing the equator (Figure 5b). Thus, the monsoons in the Northern Hemisphere are fed by southwesterly flows, those in the Southern Hemisphere by northwesterly flows (Figure 2).

The circulation reverses around 10-15 km altitude, in the upper part of the tropical troposphere. There is no rigid upper boundary in the atmosphere, as in the reservoir in Figure 4, however the increase in temperature with altitude in the stratosphere inhibits the vertical movement of the air beyond the tropopause. An area of high pressure with a divergent anticyclonic circulation [10] is thus established above the warm low layers. Part of this divergent flow heads towards the equator with an easterly component and forms the **Tropical Easterly Jet**, an area of strong winds that extends from Southeast Asia to West Africa above 10 km altitude during the boreal summer. Similar strong, but less pronounced, high-altitude easterlies occur near the equator in connection with the American, Australian and southern African monsoons (see: *Jet streams*). The flow that crosses the equator gradually acquires a westerly component and contributes to the formation of the **Subtropical Western Jet** in the winter hemisphere. By interacting with the **Polar Westerly Jet** located higher in latitude in the same hemisphere, this circulation can influence the development of weather disturbances in winter mid-latitudes.

The inter-hemispheric transfers associated with the contrasting radiation balance between the summer and winter hemispheres are not unique to the planet Earth. A similar seasonal phenomenon is observed in the Martian atmosphere, which is admittedly much more tenuous than ours, due to an inclination of 25.19° close to the terrestrial value of 23.44°, with moreover a strong difference in topography between the two Martian hemispheres. On Earth, the presence of continents and oceans modulates this circulation and produces the specific characteristics of the monsoon regime.

### 4. The consequences of the ocean-continent thermal contrast

Land and water surfaces respond differently to the energy input from solar radiation (Figure 6). An important property of surfaces is the heat capacity or specific heat, which is the amount of heat (expressed in Joule) that must be supplied to a mass of 1 kg of material to raise its temperature by 1°C. Its value for liquid water is 4185 J kg⁻¹ °C⁻¹ at standard conditions of temperature (0°C) and pressure (1013.25 hPa). The heat capacity of soils varies according to their nature: from 600 to 1000 J kg⁻¹ °C⁻¹ for dry minerals, from 1000 to 2000 J kg⁻¹ °C⁻¹ for organic matter. These values increase with soil moisture. When subjected to the same energy input, a continental soil heats up 2 to 7 times more than the same mass of water.
A difference between continents and oceans concerns the vertical distribution of the solar energy received. Incident radiation that arrives on a continental surface and is not reflected is absorbed in the surface layer, which is only a few centimetres thick (Figure 6a). Thus, the continental surface heats up rapidly. The infrared heat radiation emitted by the surface diffuses vertically and in turn heats the first few hundred metres of the atmosphere above. The heat also spreads deep down, but only for a few metres. When the ground is wet, either by nature or as a result of precipitation, the solar energy received causes the partial evaporation of this liquid water into atmospheric water vapour, at the expense of warming the surface and lower atmospheric layers.

The situation is very different over the ocean (Figure 6b). First of all, surface evaporation is important, which reduces the heat available for warming. In addition, visible solar radiation penetrates to a depth of several tens of metres, spreading the energy input over a much deeper layer. The wind at the surface causes internal circulations that lead to vertical mixing, accentuating the thermal dilution in the mixed oceanic layer, which is several tens of metres deep. For the same amount of incident radiation, the heating of the ocean surface and the atmosphere above it is therefore less than on a continent.

Continental surfaces respond quickly to seasonal variations in solar irradiance. By May-June, the Sahara, the Arabian Peninsula, Pakistan and northern India and, to a lesser extent, the southwestern United States and northern Mexico reach temperatures in excess of 30°C. A similar situation is observed in the Southern Hemisphere in November-December in northern Australia and, with a lesser amplitude, in southern Africa. The slower warming of the tropical oceans is also slowed down by the intensification of the vertical mixing induced by the trans-equatorial winds blowing over their surface during the summer. In fact, the maximum temperatures at the surface and in the mixed layer are not reached until a few months after the solstice: late August – early September in the Northern Hemisphere, late February – early March in the Southern Hemisphere.

Since the surface pressure depends on the weight of the overlying air column, it is lower over warm land (warm air is less dense) than over relatively colder oceans (cold air is denser). Thus, thermal lows are established from June until the end of August around 20-30°N, from December until the end of February around 15-25°S. The size of these lows depends on the size of the continents concerned: larger in India and Southeast Asia than over the Sahara or northern Australia, much smaller in North America and southern Africa. As the air is always heading towards the low pressure areas, while undergoing a deviation due to the Coriolis force, the thermal lows attract the trans-equatorial flow which has become loaded with moisture during its passage over the tropical oceans. The monsoon can begin.

5. The monsoon rains
In spring, the convection zone[12] made up of rainstorm clusters is located near the equator, where the warmest surface waters are found (Figure 7a). In early summer, the thermal anomaly that builds up over the tropical continents draws the moisture-laden southwesterly (in the Northern Hemisphere) or northwesterly (in the Southern Hemisphere) flow into the lower layers during its long journey over the equatorial and tropical oceans. With the presence of drier air aloft, this generates, initially in coastal areas (Figure 7b), a situation of strong convective instability conducive to the development of numerous powerful rainstorm systems. These systems dump large amounts of precipitation at the surface and locally restabilize the atmosphere by heating the middle and upper troposphere through a significant release of latent heat resulting from the condensation of the water vapour associated with the monsoon flow into water droplets and ice particles.

After a few days of such activity, the stabilized coastal regions return to a certain calm. However, the inland areas remain warm and the moist monsoon flow that enters them brings the same type of instability (Figure 7c). Clouds and thunderstorms thus progress further in latitude in successive phases of advance and stabilization that last several days. The continental surfaces become humid and contribute to the recharging of the water vapour flow through evaporation. The same regions spared from rainfall dry out and warm up, becoming unstable again, and return to more or less intense rainstorm activity for several days. In the upper troposphere, the divergent anticyclonic flow at the top of the convective systems feeds the Tropical Easterly Jet and, beyond the equator, the Subtropical Western Jet. These upper tropospheric winds intensify as disturbances develop over the continent between the coast and higher latitudes.

This situation of advancing and retreating monsoon rains persists throughout the summer (July – August – September in the Northern Hemisphere, January – February – March in the Southern Hemisphere). Locally, there are active or inactive phases of the monsoon, with a greater accumulation of rainfall towards the coast directly exposed to the humid oceanic flow, and more sporadic and intense precipitation events moving towards the interior of the continent. Monsoon activity is also modulated by the various weather disturbances that propagate with different timing and speed through the tropical atmosphere. All this makes it difficult to forecast rainfall on a local or regional scale within a few days or weeks, with important consequences for agricultural activities that are highly dependent on this water supply.

At the end of the summer, the continental warm anomaly decreases due to less solar radiation (Figure 7d). The thermal depression fills up and the monsoon flow loses its intensity. The lesser supply of moist air reduces convective instability in the interior of the continent. The rainstorms diminished in intensity and regressed to the coast, before dissipating completely. The monsoon is over and the atmosphere clears. Aloft, the Tropical Easterly Jet and the Subtropical Western Jet disappear. Meanwhile, the tropical continents in the opposite hemisphere start to warm up, the thermal lows gradually deepen and the low-level flow reverses. Soon the monsoon will start on the other side of the equator.

6. Monsoon variability: present, past, future
Monsoon activity varies from year to year, mainly in relation to ENSO (El Niño – Southern Oscillation) via the perturbations induced on the ocean surface temperature and on the large-scale circulation of the tropical atmosphere. The impact is greatest in the Northern Hemisphere, particularly in India and Southeast Asia. El Niño phases with warming in the central and eastern equatorial Pacific cause weaker monsoons with less rainfall. In contrast, La Niña phases with warming of the western equatorial Pacific produce more active monsoons with more intense and longer-lasting rainfall.

Other longer-term modes of natural climate variability also influence monsoons on multi-decadal time scales. For example, the East Asian monsoon is more pronounced during positive phases (e.g. 1975-2005) of the Pacific Decadal Oscillation, when the western Pacific cools while the eastern ocean warms. Similarly, the West African monsoon is more active during the positive phases of the Atlantic Multi-decadal Oscillation (1930-1960, 1995-present), with a warm anomaly over the North Atlantic and a more intense deep circulation, than during the cold phases (1900-1930, 1960-1995).

In older times, climate reconstructions using different tracers (e.g. historical accounts, tree rings, speleothems [13], lake levels, phyto- and zoo-plankton populations [14], ...) show more active monsoons when the hemisphere concerned is abnormally warm in summer, weaker rainfall when it is colder. The monsoons in the Northern Hemisphere were thus more intense during the warm episodes of the Roman period (from -250 to +400) and the Medieval Climatic Optimum (from 950 to 1250), and weaker during the cold episodes of the Dark Ages (from 400 to 950) and the Little Ice Age (from 1250 to 1850).

Further back in time, the Holocene Climatic Optimum (9000 to 5000 BC) resulted from more sunshine than at present during the boreal summer, due to a slightly more pronounced tilt of the Earth's rotation axis (about 24° compared to 23.44° at present) and a slightly smaller distance between the Sun and the Earth at that season. This warm phase coincided with the Subpluvial Period of the Neolithic period, which was marked by a humid and rainy climate due to a very active monsoon over northern Africa and the Arabian Peninsula. Intense and frequent rainfall fed a vast network of rivers and lakes with varied vegetation and numerous fauna, where today the arid and barren deserts of the Sahara and Arabia stretch. Numerous traces of human settlements attest to conditions that were much more favourable than today. The end of this happy wet period led to massive exoduses of populations. Some of these climatic refugees, who settled in the verdant Nile Valley, may have been the precursors of the civilization of Ancient Egypt.

In the distant geological past, a very special situation existed 330 to 180 million years ago. All the continental masses were then united in a single super-continent, Pangaea, roughly C-shaped around the equator, with the Thetys Ocean inside the C (Figure 8). This exceptional configuration was very favourable to the monsoon regime, with strong trans-equatorial flows filled with...
oceanic moisture running from the southern hemisphere northwards in boreal summer, and from north to south in austral summer. The coastal areas of Pangea, along the branches of C, thus received enormous amounts of rain, favouring the development of luxuriant vegetation. But these mega-monsoons could not penetrate to the heart of Pangea, where immense deserts resulted from the near absence of precipitation and very marked daily and seasonal temperature cycles.

If we now look into the future, global warming caused by the massive release of greenhouse gases into the atmosphere by human activities will certainly have an impact on monsoon activity. Firstly, a warmer atmosphere can hold more moisture, which favours more sustained rainfall. Secondly, numerical models of future climate indicate a stronger warming in the Northern Hemisphere than in the South. This implies a strengthening of the summer thermal contrast between the tropical areas of northern Africa and Asia, and the equatorial Atlantic, Indian and western Pacific oceans. As a result, monsoon flows and associated rainfall are expected to increase with longer wet seasons, more frequent and sometimes more intense rainfall events. On the contrary, further warming of the eastern Pacific is expected to weaken the North American monsoon. Little change is expected in the Southern Hemisphere.

This is not expected to result in a complete greening of the Sahara and the Arabian and Asian deserts, as during the Holocene Climatic Optimum mentioned above. The climatic perturbations induced by the increase in greenhouse gases are indeed not as favourable as the increase in summer solar radiation due to orbital variations. However, it may be reassuring to consider that global warming, whose predicted consequences are rather negative for large parts of the planet, could ultimately contribute to a return of vegetation in currently arid areas, thus providing better environmental conditions for Sahelian, Indian and Asian populations who have had to endure severe droughts in the past decades. Nevertheless, the multiplication of episodes of intense rainfall and flooding requires adaptations of agricultural practices and developments, particularly urban ones, which are difficult to manage.

7. Messages to remember

The monsoon regime covers large areas of the tropical regions of America, Africa, Asia and Australia, and the Indian and western Pacific oceans.

The alternation of a wet and a dry season is the result of the difference in solar energy supply between the summer and winter hemispheres, and of the fact that the continents are warming up more rapidly than the oceans.

Monsoons are influenced by climate variations. Past examples show that they are generally more active during warm anomalies. In the Northern Hemisphere, they will probably become stronger with the current global warming.

References

Cover image. Monsoon clouds over Bangladesh as seen from the International Space Station [Source: © NASA / Marshall Space Flight Center, ISS June 3, 2002, Creative Commons BY-NC 2.0]


[4] The term monsoon is used in the singular to designate the meteorological phenomenon in its generality, and monsoons in the plural to refer to specific situations in different tropical regions.


[6] At a given moment, the subsolar point is the point on the planet closest to the Sun. At this point, the solar rays reach the
The planet perpendicular to its surface.

[2] The **albedo** is the ratio of the flux of reflected light energy to the flux of incident light energy.

[3] Infrared is the range of electromagnetic waves with a wavelength between 0.7 µm and 1 mm.

[4] The **Inter-Tropical Convergence Zone** is a region of low pressure near the equator, a few hundred kilometers wide, with numerous rainstorm clusters and weak winds of changing direction.

[5] An **anticyclonic** circulation rotates clockwise in the Northern Hemisphere, counter-clockwise in the Southern Hemisphere. The air rotates in the opposite **cyclonic** direction, anti-clockwise in the Northern Hemisphere and clockwise in the Southern Hemisphere, around the low pressure areas.

[6] A **divergent** circulation moves away from an area of high pressure. A **converging** circulation moves towards an area of low pressure.

[7] **Convection** refers to the overall movement of a fluid (liquid or gas) and the transport of its properties during the movement. In the atmosphere, convection results mainly from vertical variations in temperature and humidity.

[8] **Speleothems** are mineral deposits or concretions caused by the action of water infiltrating underground cavities.

[9] **Phytoplankton** are microscopic algae, often unicellular, that multiply with light, CO2 and nutrient salts. **Zoo-plankton** are microscopic, herbivorous or carnivorous animals that feed on living matter.