

CLOUD: A PARTICLE BEAM FACILITY TO INVESTIGATE THE INFLUENCE OF COSMIC RAYS ON CLOUDS

Jasper Kirkby

CERN, Geneva, Switzerland

Abstract

Palaeoclimatic data provide extensive evidence for solar forcing of the climate during the Holocene¹ and the last ice age, but the underlying mechanism remains a mystery. However recent observations suggest that cosmic rays may play a key role. Satellite data have revealed a surprising correlation between cosmic ray intensity and the fraction of the Earth covered by low clouds [1, 2]. Since the cosmic ray intensity is modulated by the solar wind, this may be an important clue to the long-sought mechanism for solar-climate variability. In order to test whether cosmic rays and clouds are causally linked and, if so, to understand the microphysical mechanisms, a novel experiment known as CLOUD² has been proposed [3]–[5]. CLOUD proposes to investigate ion-aerosol-cloud microphysics under controlled laboratory conditions using a beam from a particle accelerator, which provides a precisely adjustable and measurable artificial source of cosmic rays. The heart of the experiment is a precision cloud chamber that recreates cloud conditions throughout the atmosphere.

1 INTRODUCTION

That there is a causal connection between the observed variations in the forces of the Sun, the terrestrial magnetic field, and the meteorological elements has been the conclusion of every research into this subject for the past 50 years. The elucidation of exactly what the connection is and the scientific proof of it is to be classed among the most difficult problems presented in terrestrial physics. The evidence adduced in favor of this conclusion is on the whole of a cumulative kind, since the direct sequence of cause and effect is so far masked in the complex interaction of the many delicate forces in operation as to render its immediate measurement quite impossible in the present state of science.

F.H. Bigelow

US Dept. Agriculture Weather Bureau

Bulletin No.21, 1898

This quotation [6] is from an article written over a century ago and yet it could be taken almost wholly from a contemporary paper. The observation that warm weather seems to coincide with high sunspot counts and cool weather with low sunspot counts was made as long ago as two hundred years by the astronomer William Herschel [7] who noticed that the price of wheat in England was lower when there were many sunspots, and higher when there were few. The most well-known example of a solar-climate effect is known as the Maunder Minimum [8], the period between 1645 and 1715—which ironically almost exactly coincides with the reign of Louis XIV, *le Roi Soleil*, 1643–1715—during which there was an almost complete absence of sunspots (Fig. 1). This marked the most pronounced of several prolonged cold spells in the period between about 1450 and 1890 which are collectively known as the Little Ice Age. During this period the River Thames in London regularly froze across and fairs

¹The Holocene is the present interglacial period—the previous 11.5 kyr since the end of the last ice age.

²CLOUD is an acronym for Cosmics Leaving OUtdoor Droplets.

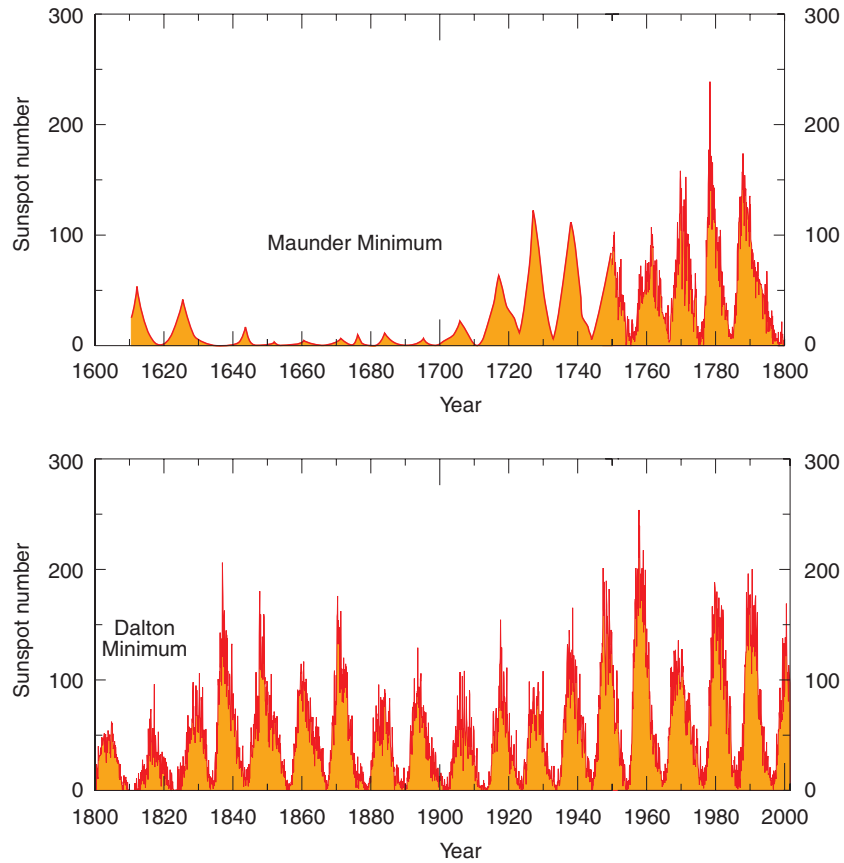


Fig. 1: Variation of the sunspot number from 1610 to 2001. The record starts 3 years after the invention of the telescope by Lippershey in Holland. The Maunder and Dalton Minima are two pronounced cold spells in the period between about 1450 and 1890 which are collectively known as the Little Ice Age.

complete with swings, sideshows and food stalls were a standard winter feature. Numerous studies of palaeoclimatic proxies have both confirmed that the Little Ice Age was a global phenomena and shown that it was but one of around 10 occasions during the Holocene when the Sun entered a grand minimum for centennial-scale periods and influenced the Earth's climate (§3). During the Holocene there have also been a similar number of extended periods of high solar activity, amongst which is the second half of the 20th century.

However, despite the evidence, solar variability remains controversial as a source of climate change since no causal mechanism has been established to link the two phenomena. The most obvious mechanism to suspect is a variation of the solar irradiance. Precision satellite measurements of the solar irradiance have indeed revealed a small variation of about 0.1% over the solar cycle [9] (§4.2). Together with observations of cyclic stars similar to the Sun, this has led to estimates of somewhat larger long-term variations of the solar irradiance, but these nevertheless appear to be too small to account for the observed climate variability. For example, it is estimated that the solar irradiance, I , was weaker by 3.3 Wm^{-2} ($\Delta I/I = 2.4 \cdot 10^{-3}$) during the Maunder Minimum [10], when globally-averaged temperatures were cooler by about 0.5–1K, after subtracting the estimated anthropogenic contributions during the last century. The relative temperature change is then $\Delta T/T = (0.5 - 1.0)/288 = (1.7 - 3.5) \cdot 10^{-3}$. This suggests that the Earth's temperature sensitivity, $\Delta T/T \simeq \Delta I/I$. Since a simple black body would respond as $\Delta T/T = \Delta I/4I$, this implies either that the Earth has a high sensitivity to irradiance changes or that other mechanisms exist that amplify the solar variations, or both. Indeed, the response of climate is complex and involves more than simple radiative heating and cooling (§3.3.2). Moreover there is in fact no direct evidence that the irradiance of the Sun is varying beyond the 0.1% solar cycle variations

(which can be quantitatively well-explained by sunspot darkening and facular brightening of the photosphere). So the magnitude of the long-term change in solar irradiance—if any—is speculative. The physical mechanism or mechanisms for solar-climate variability therefore remain a mystery.

However, the recent observation of correlations between the galactic cosmic ray (GCR) intensity and the fraction of Earth covered by low clouds [1, 2] (§2.1) may provide an important clue. Clouds cover a large fraction of the Earth’s surface—a global annual mean of about 65%—and exert a strong net cooling effect of about 30 Wm^{-2} , so long-term variations of only a few per cent could have a significant effect on the Earth’s climate. Since the GCR intensity is modulated by the solar wind, a GCR-cloud link could provide a sufficient amplifying mechanism for solar-climate variability. This would constitute a new *solar indirect* contribution to climate change, in addition to the direct contribution from irradiance changes (§3.3.5).

If a causal connection between GCR intensity and low cloud cover were to be confirmed, it could have profound consequences for our understanding of the solar contributions to the current global warming. During the 20th century the Sun’s magnetic activity increased dramatically and the solar wind more than doubled in strength [11] (§3.3.1). As a consequence, the mean GCR intensity on Earth diminished by about 15%. The implied reduction in low cloud cover by about 1.3% absolute could have given rise to a radiative forcing of about $+0.8 \text{ Wm}^{-2}$ ($3.3 \cdot 10^{-3}$), which is comparable to the estimated total anthropogenic forcing of about $+1.3 \text{ Wm}^{-2}$ (§3.3.5).

We can look further back in time for evidence of solar forcing³ of the climate. Detailed records of the magnetic variability of the Sun are preserved in the light radio isotope archives, notably the ¹⁴C content of tree-rings (for about the last 10 kyr) and the ¹⁰Be concentrations in ice-cores from Greenland and Antarctica (about 250 kyr) (§3.1.1). The light radio isotopes are produced by GCRs interacting with nitrogen, oxygen and argon nuclei in the atmosphere, and so they are a direct measure the prevailing GCR intensity. These records show the Sun to be a variable star, with both quasi-cyclic activity (11, 88, 208 yr. . .) and also periods of ‘grand-minima’ occurring on millennial-scale intervals. Diverse palaeoclimatic records have also shown that the Earth’s climate was not stable in the past and that large changes have occurred naturally. Comparisons between the solar and palaeoclimatic records reveal unmistakable evidence for a solar forcing of the climate (§3.2–§3.3).

However, in the absence of an established physical mechanism, even the evidence for solar-climate correlations accumulated in studies over the last two hundred years has not proved cause and effect. But now—and perhaps for the first time—we have a definite hypothesis for the mechanism that can be tested experimentally, namely: *are cosmic rays affecting cloud formation?* Since the energy flux from cosmic rays is tiny—about the same as that from starlight, and only a few parts per billion compared with the solar irradiance—a strong amplification mechanism would be required, i.e. some microphysical property or properties of clouds must be very sensitive to the ionisation or radicals produced by GCRs. Although a frequent criticism of the GCR-cloud hypothesis has been the absence of any microphysical mechanism, there are in fact several candidates, associated with aerosols, ice particles and cloud electricity (§4). However since none of these mechanisms is firmly established, they must be tested experimentally.

How can this best be done? In the atmosphere it is hard to establish cause and effect since it is difficult to measure all the variables and essentially impossible to adjust them. For these reasons the CLOUD experiment [3]–[5] (§5) proposes to perform the necessary measurements under controlled conditions in the laboratory. CLOUD plans to use a particle beam from an accelerator to provide a precisely controllable source of relativistic ionising radiation that closely duplicates cosmic rays in the atmosphere (§5.4). Processes can be studied with beams of varying intensity around naturally-occurring levels, and also with no beam present. The beam will pass through an expansion cloud chamber (§5.3.1) and a reactor chamber where the atmosphere is duplicated by moist air charged with selected aerosols and trace condensable vapours. The cloud chamber dynamically simulates the thermodynamic conditions, electric

³A climate forcing is a perturbation of the Earth’s radiative energy balance, with the convention that a positive forcing leads to a warming, and a negative forcing to a cooling.

fields and water vapour supersaturations within clouds throughout the troposphere and stratosphere. As well as *in situ* analysis of the cloud chamber contents, samples are extracted and sent to an array of external detectors and mass spectrometers where the physical and chemical characteristics of the aerosols and trace gases are analysed during beam exposure. Where beam effects are found, the experiment will evaluate their significance in the atmosphere by incorporating them into aerosol and cloud models, and by examining the sensitivity of clouds under atmospheric conditions to variations of the GCR intensity in the presence of other sources of natural variability. A close exchange is foreseen between CLOUD and the related field experiments so that, on the one hand, the laboratory results can be applied in the atmosphere and, on the other hand, new field work can help to shape the CLOUD experimental programme.

CLOUD is designed as a flexible ‘general-purpose’ detector, for which a wide range of experiments on ion-aerosol-cloud interactions is envisaged over several years (§5.5). Flexibility is required because this field is relatively unexplored but likely to develop rapidly in the coming years—and it is impossible to predict where these future experimental and theoretical developments may lead. For these reasons it is more appropriate to consider CLOUD as a facility than a one-off experiment. As well as its primary goal of investigating the effect of cosmic rays on clouds, the CLOUD facility will provide valuable experimental data on a broad range of important related aerosol and cloud properties, such as the optical reflectivities from liquid and ice clouds, and the dynamics of the activation of cloud condensation nuclei (CCN) into droplets.

2 OBSERVATIONS OF SOLAR-CLOUD VARIABILITY

2.1 Experimental observations

Although clouds have been routinely monitored from ground stations for more than a century, it was only over the last 20 years that global measurements became available from satellites. In 1997 Svensmark and Friis-Christensen reported a surprising correlation between global cloud cover and the GCR intensity [1]. Following their discovery, several papers pointed out important limitations in the satellite cloud data and its analysis (see, for example, refs. [12]–[14]). Among the concerns were the use of a composite of several independent satellite datasets with limited time coverage and with inter-calibration uncertainties; a spatial coverage limited to oceans and excluding the tropics and polar regions; a limited temporal coverage (mostly daytime only); and the absence of any indication of which type of cloud is affected (an increase in high clouds would result in a warming whereas an increase in low clouds causes a cooling). A frequent—but misplaced—criticism has also been the lack of any physical mechanism connecting cosmic rays and cloud cover (§4).

These limitations have largely been addressed with the recent release of the ISCCP-D2 cloud dataset [15] and its subsequent analysis [2]. The new cloud data (Fig. 2) comprise a single unified dataset over the period July 1983 to September 1994 and provide complete global coverage, day and night (at 10–12 μm IR wavelengths). As well as cloud frequency, the cloud-top temperatures and pressures are also determined. The temperatures and pressures are obtained by assuming an opaque cloud, i.e. an emissivity $\epsilon = 1$, and adjusting the cloud’s pressure level (effectively the cloud-top altitude) in the model until the reconstructed outgoing IR flux matches that observed. The clouds are classified into 3 altitude ranges according to the pressure at their top surface: low, >680 hPa (approximately <3.2 km); middle, 680–440 hPa (3.2–6.5 km); and high, <440 hPa (>6.5 km). The new cloud data indicate the presence of a solar modulation in the fraction of *low* clouds—but none for clouds at higher altitudes (Fig. 2). This establishes the *sign* of the GCR-cloud correlation: a higher GCR intensity is associated with increased low clouds and therefore with a cooler temperature (§2.2).

The global distribution of the correlation of GCR intensity and low cloud fraction is shown in Fig. 3a) [2]. The fraction of the Earth’s surface with a correlation coefficient above 0.6 is 14.2%. The regions of high correlation appear to be rather uniformly distributed—although there appears to be some preference for the oceans, where aerosol concentrations are generally lower than over land (§4.4.2). A

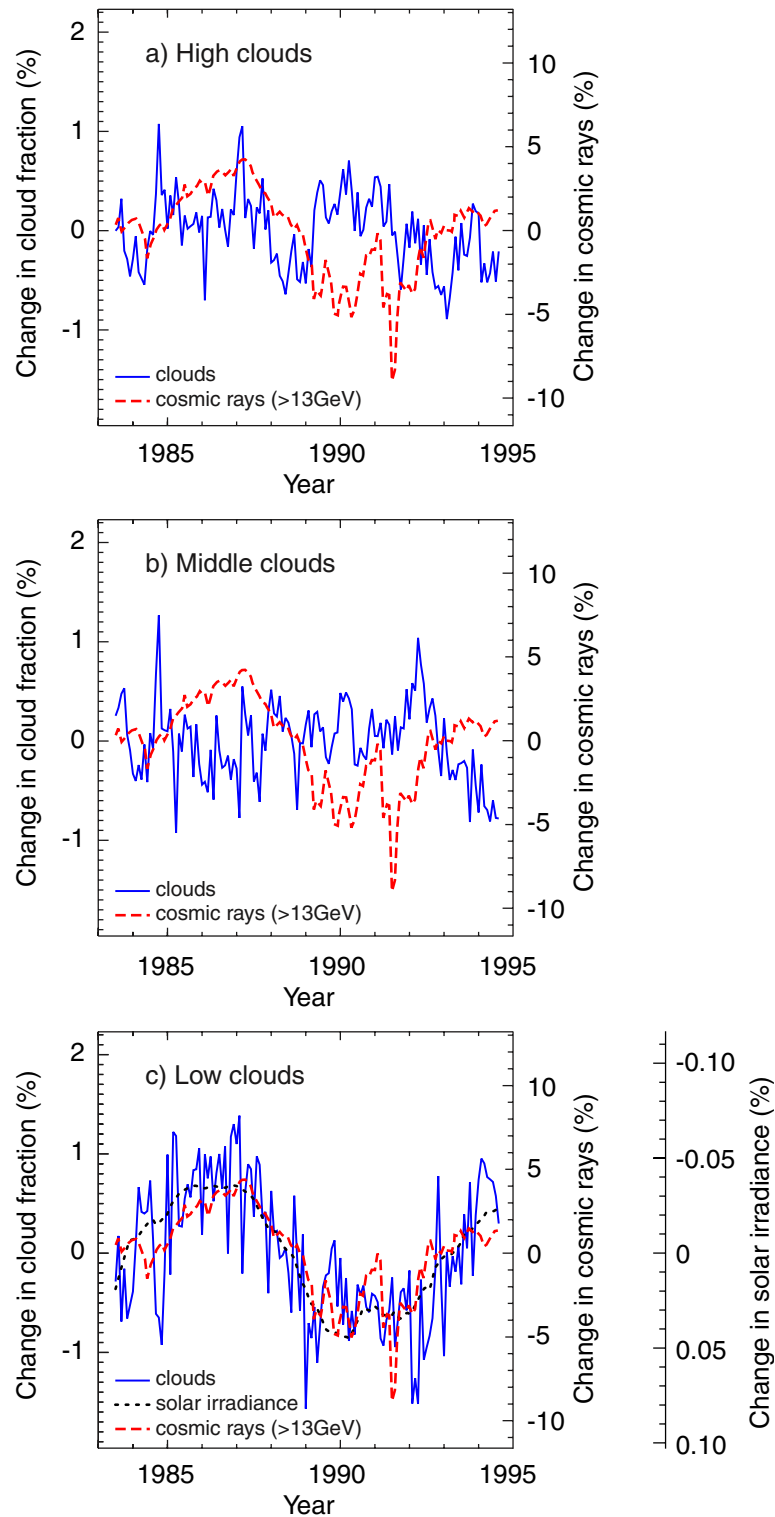


Fig. 2: Monthly mean values for the global absolute variations of infra red (IR) cloud coverage for a) high (cloud-top pressure <440 hPa), b) middle (440–680 hPa), and c) low (>680 hPa) clouds (solid lines) (adapted from refs. [2] and [16]). The ISCCP-D2 cloud data are obtained at infra-red (IR) wavelengths (10–12 μm) and have complete global coverage, day and night. The variations, over this period, of cosmic ray intensity (dashed line; 13 GeV/c rigidity cutoff) and solar irradiance (dotted line in lowest panel; note inverted scale) are also indicated. The mean global cloud fraction over this period for high, middle and low IR clouds is 13.5%, 19.9%, and 28.0% respectively. The cloud measurements are obtained from the ISCCP-D2 IR dataset [15].

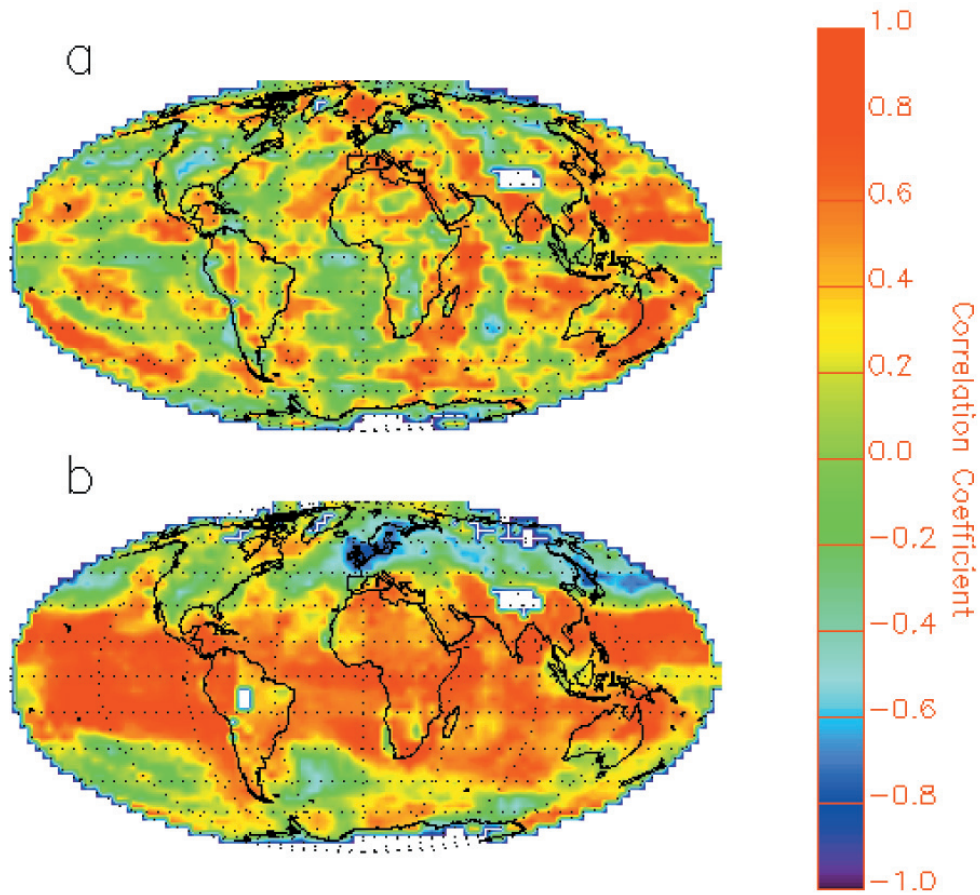


Fig. 3: Global maps of the correlation between cosmic ray intensity and a) low IR cloud fraction and b) low IR cloud-top temperature [2]. The low IR cloud fractions are calculated as in Fig. 2c), while the low cloud-top temperatures are obtained from the ISCCP-D2 IR model. White pixels indicate regions with either no data or an incomplete monthly time series. The correlation coefficients are calculated from the 12-month running mean at each grid point. Fractions of the Earth with a correlation coefficient ≥ 0.6 are a) 14.2%, and b) 29.6%, respectively. The probability of obtaining a correlation coefficient ≥ 0.6 from a random signal is $< 0.01\%$ per pixel.

few regions, such as North America, show a *negative* correlation. The lower map (Fig. 3b) shows the correlation of low IR cloud-top temperature and cosmic ray intensity. It shows a strong and continuous band of high correlation (>0.6) extending throughout the tropics, covering 29.6% of the globe. This is a counter-intuitive result since the solar modulation of the GCR intensity is a minimum near the geomagnetic equator. Nevertheless there is a finite GCR modulation of about 5% peak-to-peak at the equator.

The reason for the different global distributions of high GCR correlation in Figs. 3a) and 3b) is not known, although we note that these are two distinct cloud properties and so, in principle, they may be affected differently. Cloud frequency is a measure of cloud lifetime whereas cloud-top temperature measures the altitude at its upper boundary. The striking band of high correlation seen in the cloud-top temperatures over essentially the entire tropics may indicate an influence of GCRs on the convective activity of the Inter Tropical Convergence Zone (ITCZ)—the boundary between the northern and southern Hadley cells, where the Earth’s most intense convective transport of water vapour into the upper troposphere occurs. There is, in fact, some palaeoclimatic data to support a possible link between the ITCZ and solar activity (§3.2.4).

The cloud data of Fig. 2 span only a single solar cycle and one may speculate how the correlation

will develop in future. There have been numerous previous observations of solar cycle effects on the Earth's climate that have persisted for some decades and then apparently disappeared [17]. A notable example was the observation in 1923 [18] that the levels of the central African Lakes Victoria and Albert were highly correlated with solar activity (0.87 correlation coefficient) over the previous two solar cycles (1896–1922). The correlation broke down around this time, as did a number of other solar-climate relationships elsewhere. This may suggest the association was accidental. However—and perhaps more likely in view of the coincidental termination of several solar-climate observations—it may reflect the complexity of the Earth's climate, in which many factors are important and they interact in a complex way. The climate may have “stable” states where the conditions are favourable for solar forcing, and a correlation may persist for some decades. Then, at other times, the conditions are unfavourable and the correlations disappear.

Finally we note that there is some indication that the reflectivity (cloudiness) of Neptune may correlate with solar activity [19]. Measurements of the reflectivity at 472 nm and 551 nm from 1972 to 2000 show a 10% overall increase of brightness together with an apparent 2% residual solar modulation that is anti-correlated with solar activity over the three solar cycles spanned by the data. Since Voyager measured Neptune's magnetic field to be small—less than 1 Gauss—the cosmic ray intensity on Neptune is expected to be modulated by the solar wind. The conditions on Neptune are of course completely different than those on Earth—the solar irradiance is 0.1% of Earth's, and the clouds are probably liquid methane—so it is not possible to draw any conclusions about the Earth's climate from this observation.

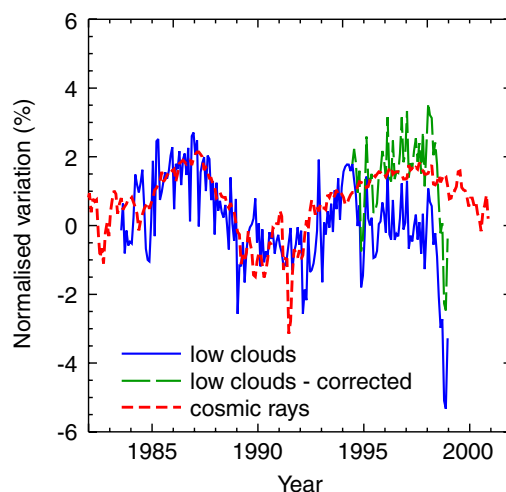


Fig. 4: Recent extension of the ISCCP-D2 low IR cloud data up to December 1998 (solid line). The broken line shows the ISCCP data after correcting by its difference with the SSMI cloud data after January 1994. The variations of cosmic ray intensity at Huancayo (13 GeV/c rigidity cutoff) are indicated by the dotted line. All curves have been normalised to their respective mean and variance over the period July 1987 - June 1990 [20].

2.1.1 ISCCP extension

Very recently, an extension of the ISCCP-D2 cloud data has been released for the period from January 1994 to December 1998. This shows a weakening of the correlation between low cloud amount and low cloud top temperature with cosmic rays after 1994 (Fig. 4) [20]. However comparison with independent cloud data from the SSMI instrument⁴ shows a good agreement with ISCCP low cloud amount until 1994, after which the two measurements seem to diverge. This suggests the possible presence of long term drifts in at least one of the satellite data sets. Indeed it is generally accepted that the long term stability and calibration of multi-satellite cloud detectors, such as ISCCP or SSMI, is challenging. Although it

⁴The SSMI (Special Sensor Microwave Imager) instrument is part of the DMSP (Defence Meteorological Satellite Program) satellites.

is not possible to resolve these discrepancies at present, an estimate of the uncertainties in the drift can be evaluated from the difference in the long term trends of the ISCCP and SSMI data. One limit is to correct the ISCCP cloud data after 1990 in direction and magnitude by its difference with the SSMI data, assuming the two datasets agree over the period 1987–1990. This is shown in Fig. 4 and indicates a good correlation between low cloud amount and cosmic rays over the full period of available cloud data (1983-1998). The uncertainties in cloud amount for the period after 1994 appear to be too large at present to draw any conclusion on either the absence or presence of a correlation with the GCR intensity.

2.2 Cloud radiative forcing

The observed variation of low cloud cover over the solar cycle of about 1.7% absolute corresponds to 6.0% relative. Since measurements by the Earth Radiation Budget Experiment (ERBE) indicate that low clouds contribute a global annual mean radiative forcing of about -17 Wm^{-2} (Table 1) this implies the cloud modulation corresponds to about $+1.0 \text{ Wm}^{-2}$, solar minimum-to-maximum. This is about a factor 5 larger than the solar cycle irradiance forcing at the Earth’s surface (0.2 Wm^{-2} ; §3.3.3), and in phase.

Table 1: Global annual mean forcing due to various types of clouds, from the Earth Radiation Budget Experiment (ERBE) [21].

Parameter	High clouds		Middle clouds		Low clouds	Total
	Thin	Thick	Thin	Thick	All	
Global fraction (%)	10.1	8.6	10.7	7.3	26.6	63.3
Forcing (relative to clear sky):						
Albedo (SW radiation) (Wm^{-2})	-4.1	-15.6	-3.7	-9.9	-20.2	-53.5
Outgoing LW radiation (Wm^{-2})	6.5	8.6	4.8	2.4	3.5	25.8
Net forcing (Wm^{-2})	2.4	-7.0	1.1	-7.5	-16.7	-27.7

3 RECORD OF SOLAR-GCR-CLIMATE CHANGE

3.1 Solar and palaeoclimatic records

3.1.1 Solar and GCR records

The flux of galactic cosmic rays reaching the Earth’s atmosphere is modulated by variations of the heliospheric magnetic field and of the Earth’s geomagnetic field. During times of high solar activity (sunspot maximum) there is an increase of the open magnetic flux and of the magnetic irregularities carried out into the heliosphere by the solar wind. These magnetic fields scatter the low-energy component of the incoming GCRs (below a few tens of GeV) and, in consequence, the flux reaching Earth is reduced. The global average modulation of the GCR intensity over the solar cycle is about 15%, but larger variations occur on longer timescales. During a reversal of the Earth’s dipole field, for example, it is estimated that the global GCR rate is enhanced by about a factor 2.5 relative to the present values.

An exquisite record of the variations of GCR intensity over the past 250 millennia is preserved in the light radio isotope records in ice cores [22]. These provide an essentially direct measurement of the prevailing GCR intensity and hence are a direct indication of variations of the solar *magnetic* activity. They are also frequently used as a proxy for putative changes of solar irradiance, although there exists no direct evidence for long-term variations of the solar irradiance. The light radio-isotopes are produced in spallation interactions of GCRs on nitrogen, oxygen and argon nuclei in the atmosphere. The two radioisotopes with the highest production rates are ^{14}C (half life = 5730 ± 40 yr and global mean

production rate $\sim 2.0 \text{ atoms cm}^{-2}\text{s}^{-1}$) and ^{10}Be (1.5 Myr; $\sim 1.8 \times 10^{-2} \text{ atoms cm}^{-2}\text{s}^{-1}$). The third most abundant isotope is ^{36}Cl , which is produced from GCR interactions with Ar nuclei (300 kyr; $\sim 1.9 \times 10^{-3} \text{ atoms cm}^{-2}\text{s}^{-1}$).

The ^{14}C is rapidly oxidised to $^{14}\text{CO}_2$. The turnover time of CO_2 in the atmosphere is quite short—about 4 years—mostly by absorption in the oceans and assimilation in living plants. However, because of recirculation between the oceans and the atmosphere, changes in the ^{14}C fraction on timescales less than a few decades are smoothed out. Plant material originally contains the prevailing atmospheric fraction of ^{14}C and, subsequently, since the material is not recycled into the atmosphere, the fraction decreases with the characteristic half life of ^{14}C . By analysing the ^{14}C content in the rings of long-lived trees such as the California bristlecone pine, a continuous yearly record of GCR intensity over the last 10–15 kyr has been assembled.

In the case of ^{10}Be , after production it rapidly attaches to aerosols and follows the motion of the surrounding air masses. Since the production of ^{10}Be follows the intensity profile of the cosmic ray hadronic showers, about 2/3 is produced in the stratosphere and 1/3 in the troposphere, globally averaged. Due to the tropopause barrier, aerosols in the stratosphere take about 1–2 years to settle on the Earth's surface, whereas the mean residence time in the troposphere is only about a week. If the sedimentation occurs in the form of snow in a permanently frozen and stable region such as Greenland or Antarctica then the subsequent compacted ice preserves a temporal record in layers according to their depth.

The measured variations of the light radionuclides are the product of two processes: 1) the production rate and 2) system effects i.e. transport, precipitation and exchange processes between the different reservoirs. Since the system effects are quite different for ^{14}C and ^{10}Be , it has been possible to reliably determine the production rates, and hence the GCR intensities. An advantage of ^{14}C is that it is well-mixed before storage in the tree-ring archives. In contrast, the short residence time of the tropospheric ^{10}Be fraction means that the measured concentrations are subject to possible variations of precipitation rate and of wind directions in carrying the radionuclide from where it is produced to where it is sedimented as snow. However, by using a minimum sample period of 1 or 2 years, the effects of variations of transport direction and efficiency are minimised. Since ^{10}Be is not recycled into the atmosphere, it has the important advantages of being able to record relatively short-term changes in the GCR intensity, and a freedom from uncertainties due to variations in the recycling processes.

3.1.2 Palaeoclimatic records

Many ingenious proxies have been developed to reconstruct the climate prior to the last two centuries, for which instrumental records are available. Cultural records over about the last millennium are an important source since humans are sensitive to climate change, especially when prolonged drought, cold or flooding is involved. These sources include documents recording the dates when the first cherry blossoms appeared each spring in China, as well as records of the grape harvests in Europe. Other records (with their approximate time span BP ⁵ in parentheses) are corals (400 yr), tree rings (10 kyr), mosses (10 kyr), pollen (1 Myr), ice cores (250 kyr), ocean sediments (>1 Myr) and geomorphology (3 Byr).

Ice cores are an especially valuable record of past climate [22]. As well as the solar-GCR record described above, the trapped gases preserve the atmospheric composition at earlier times, layer thicknesses measure precipitation rate, dust content measures wind speed and volcanic activity, sulphate measures sulphuric acid content of the atmosphere (volcanic and planktonic activity) and, of particular importance, H_2^{18}O measures past temperatures.

The physical basis for proxy temperature measurements from the stable ^{18}O isotope is that the vapour pressure of H_2^{18}O is lower than that of H_2^{16}O . Evaporation from the oceans thus produces water vapour that is ^{18}O -depleted (by about 1% relative); conversely, the remaining water is enriched in ^{18}O .

⁵BP signifies *before present*, where 'present' means 1950.

During condensation, the lower vapour pressure of the H_2^{18}O leads to preferential condensation, and so the water vapour becomes progressively more ^{18}O -depleted as it travels poleward. Because condensation is the result of cooling, the greater the fall in temperature, the lower is the heavy isotope concentration. Isotope concentration in the condensate is thus a function of the temperature at which condensation occurs. The relative proportion of ^{18}O and ^{16}O in an ice core sample, R_s , is expressed in terms of its fractional deviation, $\delta^{18}\text{O}$, from a standard value,

$$\delta^{18}\text{O}(\text{per mil}) = ((R_s/R_{\text{SMOW}}) - 1) \cdot 1000 \quad (1)$$

where $R = [^{18}\text{O}] / [^{16}\text{O}]$, for which the Standard Mean Ocean Water (SMOW) value is $R_{\text{SMOW}} = 2.0052 \cdot 10^{-3}$. It is found that a decrease of $\delta^{18}\text{O}$ by 1 per mil corresponds to a temperature decrease at the site of precipitation of between 1.5 K (polar regions) and 1.7 K (mid-latitudes). In addition the $\delta^{18}\text{O}$ value of sea sediments provides a measure of the global volume of water locked up in (^{18}O -depleted) ice sheets, since high ice volumes leave the oceans enriched in ^{18}O .

Deviations of other isotopes are defined in a similar way as $\delta^{18}\text{O}$. In the case of a radioisotope like ^{14}C , the final deviation is expressed as $\Delta^{14}\text{C}$ to signify correction of the measured value, $\delta^{14}\text{C}$, for radioactive decay and for isotopic ($\delta^{13}\text{C}$) fractionation (§3.2.7).

Many studies of past climate change show a correlation with changes of the GCR intensity and solar activity [17, 23]. In these cases a colder climate is found to correlate with low solar activity (high GCR intensity) and, conversely, a warmer climate correlates with high solar activity (low GCR intensity). The correlation with rainfall may be in either direction, depending on the region studied and the prevailing climatic conditions. In the remainder of this section we first present some examples of possible solar-influences on climate change during the late glacial and Holocene periods, and then close with a discussion of the solar contribution to the current global warming.

3.2 Solar-GCR-climate change during the late glacial and Holocene

3.2.1 The Younger Dryas (12,700–11,550 yr BP)

The Younger Dryas cold event (so called because it was marked by the spread of an Alpine flower known as *Dryas octopetala*) occurred between 12,700 and 11,550 years ago (Fig. 5a). For 3,000 years before the start of the Younger Dryas, the Earth had been gradually warming up after the end of the last ice age, but then the climate abruptly swung back into ice age conditions. During this warming period the first humans had entered the American continent by walking across the Bering land-bridge into Alaska. A settlement excavated from a peat bog at Monte Verde in southern Chile shows that they had rapidly migrated far south. However, around the start of the Younger Dryas, the Monte Verde water table rose and their settlement was flooded and abandoned. The cold Younger Dryas climate continued for about a thousand years before it abruptly switched back to warm conditions, marking the start of the Holocene. The temperature transitions were very rapid; the end of the Younger Dryas saw an increase of polar temperatures by about 15°C , with half that transition occurring in less than 15 years [24].

It is thought that this event was driven by changes of the ocean circulation. At present, northern maritime Europe is warmed by heat carried polewards by the Gulf Stream. When the warm water meets cold polar air in the North Atlantic, heat is released to the atmosphere and the water cools and sinks. This is reinforced by the increases in salinity, and therefore density, due to evaporation and to the formation of sea ice in the Arctic regions. The descending current is called the North Atlantic Deep Water (NADW). It flows southward through the western Atlantic where it joins the Southern Ocean Deep Water descending off the edges of Antarctica and flowing in an easterly direction. The deep water continues round South Africa and then into the Indian and northern Pacific Oceans, where it surfaces. The North Atlantic is warmer than the North Pacific. The increased evaporation therefore serves to increase salinity relative to the North Pacific, and it is this salinity gradient that is thought to drive the global thermohaline ocean circulation.

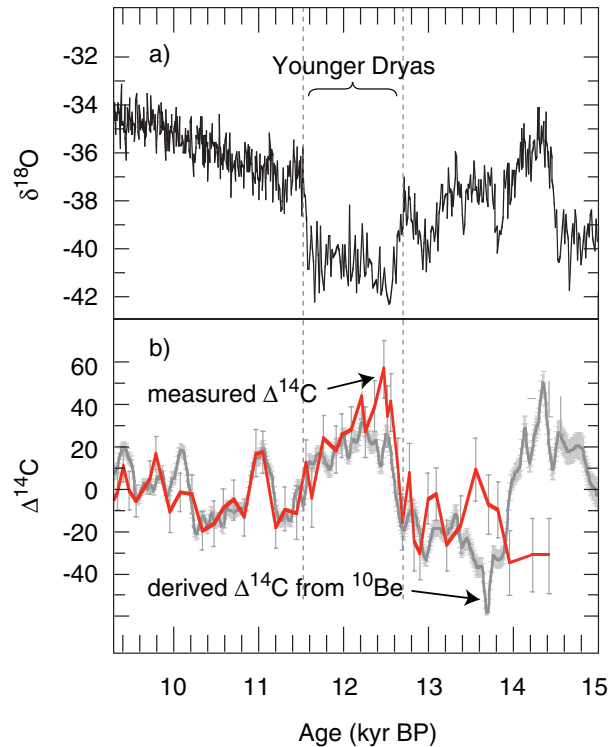


Fig. 5: The Younger Dryas cold event: a) the $\delta^{18}\text{O}$ variation over the period 15–9.4 kyr BP, as measured in the Greenland GRIP ice core, and b) the measured $\Delta^{14}\text{C}$ variation over this period (heavy curve) and the derived $\Delta^{14}\text{C}$ from ^{10}Be ice core data, taking into account some changes in deep water formation (light curve) [28].

This ‘heat conveyor belt’ is quite sensitive to climatic conditions—especially to the amount of fresh water entering the North Atlantic. During a glacial period, the formation of the NADW is thought to be much reduced or even shut down. At these times, the Arctic ice sheet extends much further south into the North Atlantic, pushing the position of the polar front southwards. Cooler sea surface temperatures reduce evaporation and therefore salinity, further weakening the thermohaline circulation. It has been suggested that the onset of the Younger Dryas was triggered by a sudden shutdown of NADW formation and therefore of the global thermohaline ocean circulation. Various causes have been proposed such as the presence of a large amount of fresh water from melting icebergs, as well as the melting and abrupt opening of the St. Lawrence waterway into the North Atlantic, diverting the drainage of fresh water over a vast region of North America away from the south and towards the north-east.

However shutdown of the NADW alone is considered insufficient to initiate global temperature changes and ice sheet development [25]. Other mechanisms would need to be invoked—and on a global scale since the Younger Dryas is also registered in the tropics and the Southern Hemisphere. Recent measurements of sea-surface temperatures (SSTs) in the mid southern latitudes over the period 40–10 kyr BP [26, 27] support the picture that North Atlantic thermohaline circulation is insufficient to drive the observed climate changes.

Interestingly, during the Younger Dryas a large increase occurred of atmospheric ^{14}C . It has been argued that this was due to the reduced circulation of (^{14}C -depleted) CO_2 from the oceans as the ice sheets advanced. However the increase of ^{14}C occurs abruptly at the start of the Younger Dryas and seems to be too sharp to be caused by changes of ocean circulation alone. Indeed a recent comparison with the ^{10}Be record during this period has concluded that the largest part of the increase of ^{14}C during the Younger Dryas can be attributed to a change in the production rate, i.e. to an increase of the GCR intensity (Fig. 5b) [28]. This suggests that solar forcing may have triggered and helped to sustain the Younger Dryas event [29].

3.2.2 *Ice-rafted debris in the North Atlantic (32,000 yr BP–present)*

Bond *et al.* have analysed sediments of ice rafted debris (IRD) in the North Atlantic [30, 31]. The latter are found in deep sea cores as layers of tiny stones and micro-fossils that were frozen into the bases of advancing glaciers and then rafted out to sea by glaciers. These reveal abrupt episodes when cool ice-bearing waters from the North Atlantic advanced as far south as the latitude of Britain, coincident with changes in the atmospheric circulation recorded in Greenland.

A quasi-cyclic occurrence of IRD events has been found, with a periodicity of 1470 ± 530 yr, during which temperatures dropped and glacial calving suddenly increased (Fig. 6). The underlying cause of these events is not yet known but the evidence tightly constrains the possibilities. First, the rafting icebergs are launched simultaneously from more than one glacier, so the driving mechanism cannot be ascribed to a single ice sheet but requires a common climate forcing mechanism. It points to a trigger that caused air temperatures to drop and induce the release of ice over a large region. Second, the events continue with the same periodicity through at least three major climate transitions: the Younger Dryas-Holocene transition, the deglaciation, and the boundary within the ice age between the marine isotope stages 2 and 3 (Fig. 6) [31]. Even though the ice conditions during these transitions were changing dramatically, the IRD events continued with the same periodicity.

Third, and especially surprising, is the evidence that the IRD cold events have continued through the Holocene (Fig. 6), with the same periodicity (but of course with a lower amount of IRD material). The events were abrupt during both the glacial and Holocene periods, generally switching on and off within one or two centuries. The estimated decreases in North Atlantic Ocean temperatures during the Holocene IRD events are 2 K, or about 15–20% of the full Holocene-to-glacial temperature difference. This observation questions the validity of the currently-held picture that the Holocene has been a period of exceptional climatic stability—and much more stable than previous interglacials. For the North Atlantic at least, the IRD data show that there has been much more climate change during the Holocene than previously thought.

The implication of these observations is the presence of a quasi-periodic climate cycle of about 1500 yr that occurs independently of the glacial-interglacial climate state. Furthermore, the IRD periodicity is suggestive of the pacing of the warm Dansgaard-Oeschger events during the ice age. These events (which are seen in stage 3 of Fig. 6a) are abrupt warmings of Greenland by about 5–10 K over a few decades, followed by gradual cooling over several hundred or thousand years. Presumably the warming of the waters far north leads to an increased calving of glaciers. Simulations [32] suggest that the cold stadial periods are the ‘stable’ mode of the glacial Atlantic Ocean circulation, with NADW formation south of Iceland—the so-called ‘cold’ conveyor mode. The warm Dansgaard-Oeschger events represent a temporary transition to the ‘warm’ conveyor mode with NADW formation further north, in the Nordic Seas. A small decrease of freshwater into the North Atlantic is sufficient to trigger these events. What causes these changes in freshwater production is not yet known, although there is increasing evidence that solar forcing is involved.

Until recently the origin of the quasi-1500 yr climate cycle was unknown. Ice sheet oscillations are ruled out as the forcing mechanism. Orbital periodicities around the Sun are too long to cause millennial-scale climate cycles. However a recent study has shown that solar variability is highly correlated with the ice rafted debris events during the Holocene (Fig. 7) [34]. The correlation embraces the Little Ice Age, which appears to be the most recent of these events. This rather convincing evidence implies that solar forcing has caused at least the Holocene section of the quasi-1500 yr climate cycle in the North Atlantic. It seems likely that solar forcing also caused changes in the hydrological cycle and North Atlantic Deep Water production, triggering the Dansgaard-Oeschger events and providing an additional mechanism for globally amplifying the solar signals. (Note that although the Dansgaard-Oeschger events show a strong correlation with decreased ^{10}Be concentration in Fig. 6, this is largely due to increased rainfall (dilution) rather than a change the production rate (GCR intensity) [33].)

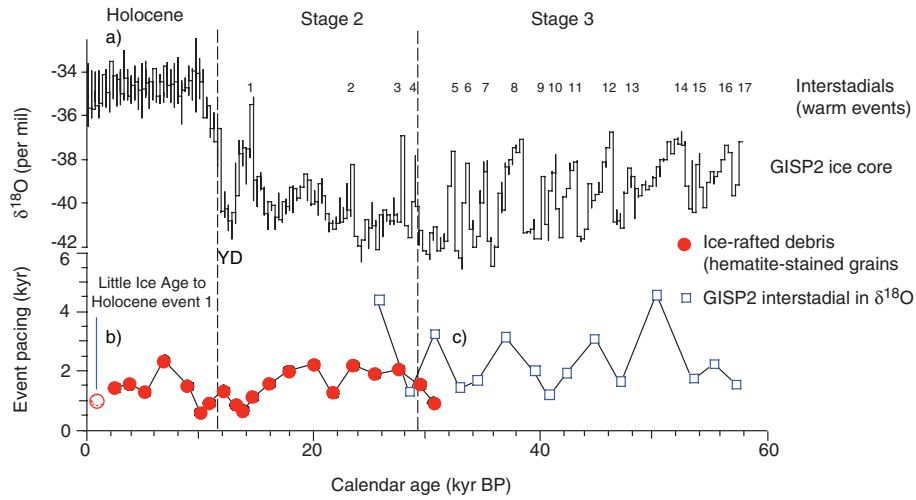


Fig. 6: Timing of ice-rafted-debris events in the North Atlantic [31]. The curves are a) the GISP2 Greenland ice core $\delta^{18}\text{O}$ record showing Greenland temperatures for the Holocene, the late glacial (Stage 2) and the mid glacial (Stage 3) periods, b) the periodicity of the ice rafted debris events from 32 kyr BP to the present, measured from haematite-stained grains (other tracers give similar results) and c) the periodicity of Dansgaard-Oeschger warm events in the GISP2 $\delta^{18}\text{O}$ data from 58 kyr to 26 kyr BP.

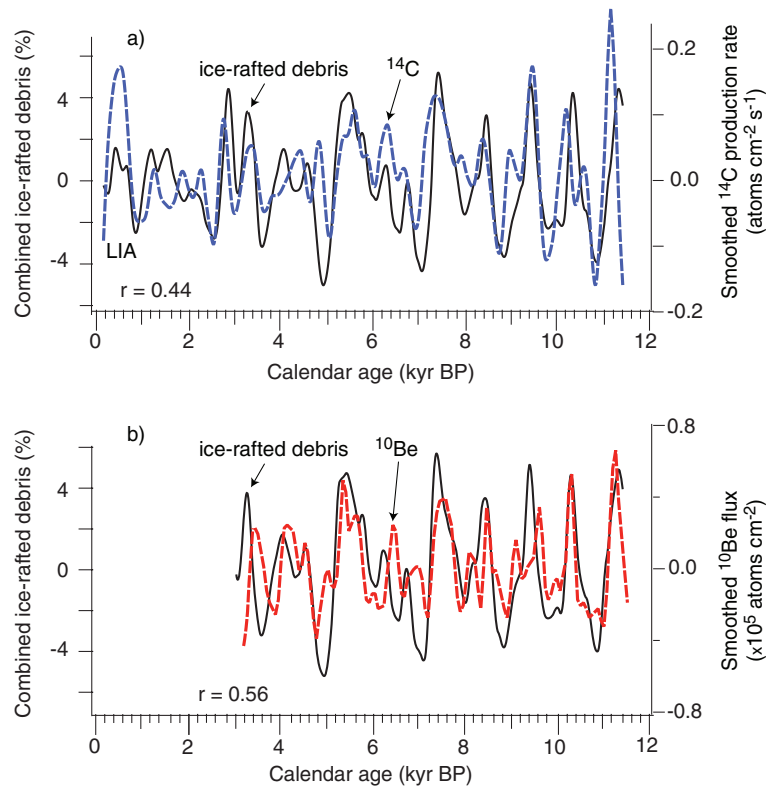


Fig. 7: Correlation of solar variability with ice-rafted debris events in the North Atlantic during the Holocene [34]: a) the ^{14}C record (correlation coefficient 0.44) and b) the ^{10}Be record (0.56), together with the combined ice-rafted-debris tracers. The Little Ice Age (LIA) is labelled in the upper figure.

3.2.3 Lake levels in the Jura Mountains (12,000 yr BP–present)

Magny has reconstructed the history of lake-levels in the French Jura Mountains which lie near the French-Swiss border [35]. These correlate well with the $\Delta^{14}\text{C}$ (GCR) record over the last 10 kyr (Fig. 8). This of course implies that the Jura lake levels also correlate with the periods of increased ice rafted debris in the North Atlantic [29] (§3.2.2).

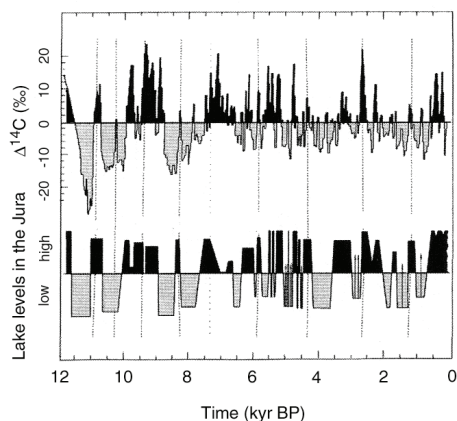


Fig. 8: Lake-Levels in the French Jura Mountains, and the $\Delta^{14}\text{C}$ variation over the last 12 kyr. (Fig. 8) [35].

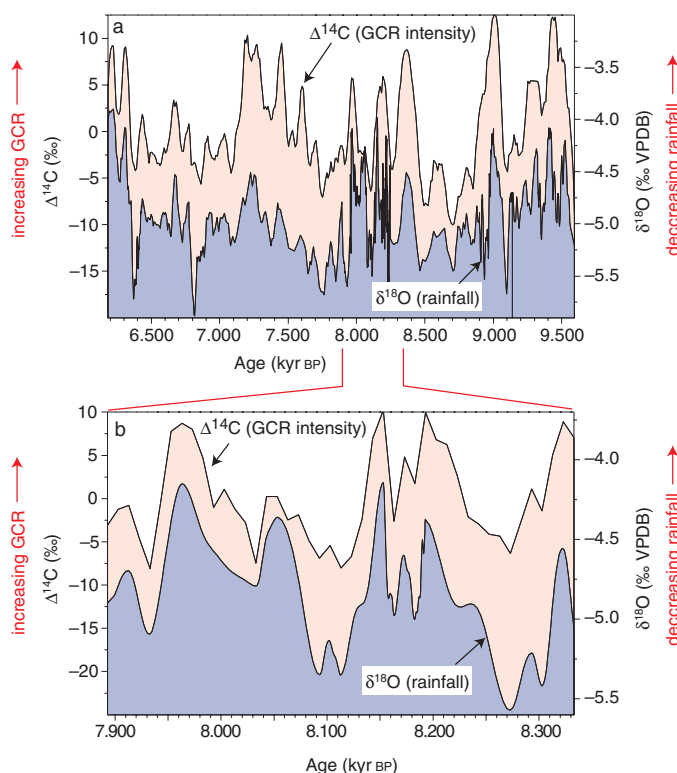


Fig. 9: Profiles of $\delta^{18}\text{O}$ from a U-Th-dated stalagmite from a cave in Oman, together with $\Delta^{14}\text{C}$ from tree rings, for a) the 3.4 kyr period from 9,600 to 6,200 yr BP and b) the 430 yr period from 8,330 to 7,900 yr BP [36].

3.2.4 Oman rainfall (9,600–6,200 yr BP)

Neff *et al.* [36] have recently measured the $\delta^{18}\text{O}$ content in the layers of a stalagmite from a cave in Oman, which are U-Th dated to cover the period from 9,600 to 6,200 yr BP. The $\delta^{18}\text{O}$ is measured in calcium carbonate, which is expected to be deposited in isotopic equilibrium with water. The data are

shown in Fig. 9 together with the $\Delta^{14}\text{C}$ obtained elsewhere from tree rings. The two timescales have been tuned to match bumps within the known experimental errors (smooth shifts have been applied to the U-Th dates up to a maximum of 190 yr). During a 430-yr period centred around 8.1 kyr BP, the stalagmite grew at a rate of 0.55 mm/yr—an order of magnitude faster than at other times—which allowed a high resolution $\delta^{18}\text{O}$ measurement to be made (Fig. 9b). It is interesting to note that this coincides with an ice rafting debris cold event in the North Atlantic (§3.2.2).

Oman today has an arid climate and lies beyond of the most northerly excursion of the inter tropical convergence zone (ITCZ), which carries with it the heavy rainfall of the Indian Ocean monsoon system. However there is evidence that the northern migration of the ITCZ reached higher latitudes at earlier times and, in consequence, that Oman had wetter climate. In this region, the temperature shifts during the Holocene are estimated to account for only 0.25 per mil variation in $\delta^{18}\text{O}$ of [36]. However the $\delta^{18}\text{O}$ values of monsoonal rainfall associated with the ITCZ show an inverse correlation with rainfall and so, for these data, the $\delta^{18}\text{O}$ variations are ascribed to changes of rainfall, as indicated in Fig. 9. Notice that the *sign* of the correlation is different from the other examples presented here; in this case a *high* GCR intensity is associated with a *low* rainfall. However, it is entirely plausible that a climate change can lead to different responses in different regions of the Earth. For example a globally-averaged increase of rainfall may still result in a decrease in certain regions due to effects such as a shift of the ITCZ.

The similarity between the $\delta^{18}\text{O}$ and $\Delta^{14}\text{C}$ curves in Fig. 9, both in their long-term and short-term variations, is striking. It suggests that solar-GCR activity controlled the pattern of tropical rainfall and monsoon intensity during this 3,000-year period on decadal to centennial timescales.

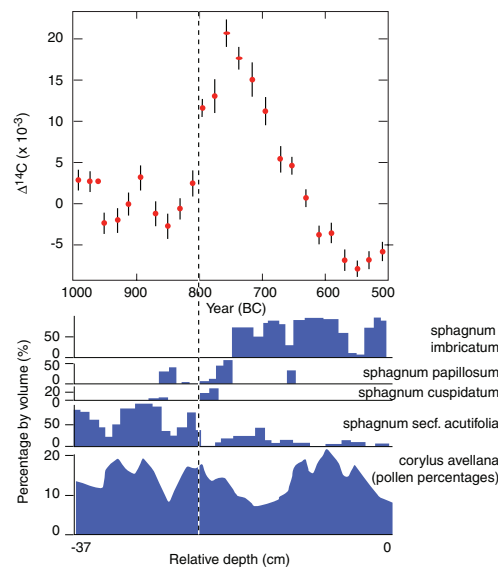


Fig. 10: Transitions in the fauna of a Netherlands peat bog core for the period 1000–500 BC [37]. Also shown is $\Delta^{14}\text{C}$ for this period. At the onset of the rise in ^{14}C around 800 BC, the peat-forming mosses shifted from those preferring relatively warm conditions to those preferring colder and wetter conditions.

3.2.5 Netherlands peat bog fauna (1000–600 BC)

Van Geel *et al.* [37] have studied peat-forming mosses in raised bogs in The Netherlands that were laid down in the period 1000–500 BC. They find an abrupt shift occurred around 800 BC from mosses preferring relatively warm conditions to those preferring colder and wetter conditions (Fig. 10). This coincides with a sharp rise in $\Delta^{14}\text{C}$ due to a decrease in solar activity. There is supporting evidence of a substantial climate shift at that time from archaeological remains of nearby Bronze Age settlements which had been continuously inhabited for more than a thousand years but were abandoned around that time, presumably as the ground became waterlogged.

There is extensive evidence that this solar-induced change to a colder and wetter climate around 800 BC was a global phenomenon. Some examples are as follows. Migrations of settlements are recorded in central Asia at this time [37]. An ice rafting debris event (§3.2.2) occurred in the North Atlantic around 2800 yr BP [31]. A substantial glacier advance took place at this time in the presently-arid south-central Andes Mountains of northern Chile, which has been attributed to a marked increase of precipitation [38]. A recent study of stalagmite growth rates in caves in the south-western United States shows that the period from 2800 to 2600 yr BP was the wettest for this (presently semi-arid) region during the last 4000 yr [39].

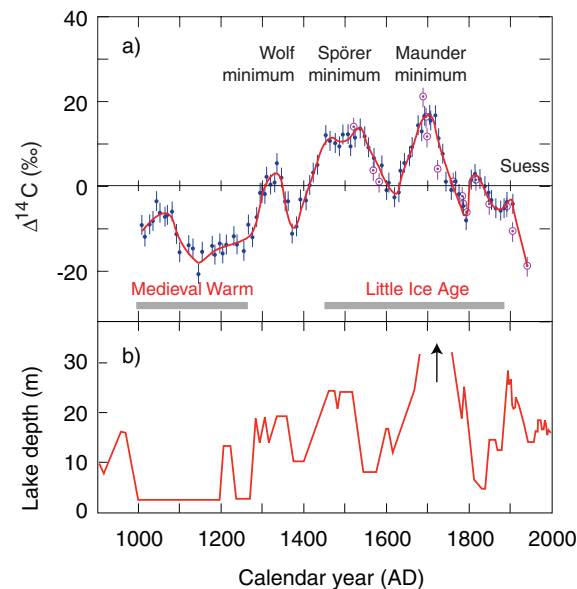


Fig. 11: a) History of $\Delta^{14}\text{C}$ from tree-ring analyses for the last millennium [40]. Recorded periods of climate change are indicated. The sharp negative ^{14}C deviation during the present century is the Suess effect, due to the burning of ^{14}C -depleted fossil fuels. b) History of rainfall and drought in equatorial East Africa during the last 1100 years [43]. The figure shows the reconstructed depth of Crescent Island Crater Lake, Kenya. The radiocarbon dating error for the lake data is ± 50 yr.

3.2.6 Kenyan lake levels (900–2000 AD)

The $\Delta^{14}\text{C}$ data for the last 1000 years reveal considerable solar variability (Fig. 11a) [40]. The periods of large ^{14}C deviation correspond to recorded climatic anomalies: a) 1000–1270, the so-called Medieval Warm period, b) 1280–1350, the Wolf Minimum, c) 1420–1540, the Spörer Minimum, and d) 1645–1715, the Maunder Minimum. Temperatures during the Medieval Warm epoch were elevated above normal, causing severe and extended droughts for the Anasazi in the south-western United States but allowing the Vikings to colonise Greenland and wine-making to flourish in England. It was followed by a period of about 4 centuries during which—save for a few short interruptions—the glaciers advanced and a cooler, harsher climate predominated. During this so-called Little Ice Age the River Thames in London regularly froze across, and fairs on the ice were a standard winter feature. The Little Ice Age was recorded in many parts of the world. For example, in China the rice crops of the Yellow River Valley were reduced from two to only one a year. Stalagmite studies in the south-western United States [39], Madagascar [41] and Nepal [42] have also recorded the Little Ice Age.

Evidence has also been found that the Medieval Warm and Little Ice Age climates extended into the equatorial regions, providing further support that they were global phenomenon. Figure 11b) shows the correlation of the ^{14}C record with the depth of a lake in equatorial East Africa over the last 1100 years [43]. The reconstruction is based on three independent palaeolimnological proxies: sediment stratigraphy

and species compositions of fossil diatoms and midges. These data not only confirm the presence of the major climatic anomalies associated with the Medieval Warm period and the Wolf, Spörer and Maunder Minima but also identify three extended drought periods between the minima: AD 1390–1420, 1560–1625 and 1800–1840. The cultural history of this region, preserved in records and oral tradition, has recorded alternating periods of drought and prosperity that coincide with the lake-level reconstruction.

3.2.7 Ionian Sea sediments (1750–1975 AD)

Biological organisms frequently preferentially use light isotope species because of the lower internal energy ‘costs’ to the organism associated with breaking the bonds in these molecules—so-called *kinetic* isotope fractionation. The result is significant fractionation between the substrate (heavier) and biologically-mediated product (lighter). The magnitude of the kinetic fractionation depends on reaction rates, concentrations of products and reactants, and environmental conditions such as light and temperature.

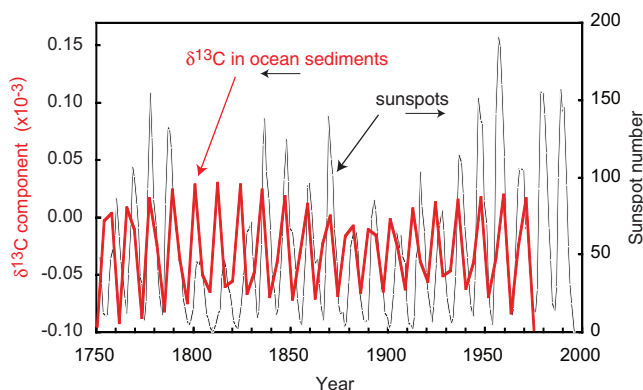


Fig. 12: The $\delta^{13}\text{C}$ content of *Globigerinoides ruber* skeleton sediments in the Ionian Sea, Italy, over the period 1750–1975 [44] (thick line). Also shown is the sunspot series over the period 1750–1995 (thin line). The $\delta^{13}\text{C}$ amplitude is obtained from the 11.3 yr component found in a singular spectrum analysis (SSA).

One such study that has been made is the $\delta^{13}\text{C}$ content of skeleton sediments of *Globigerinoides ruber*, a symbiotic planktonic foraminifera, in the Ionian Sea, Italy [44]. The $\delta^{13}\text{C}$ variations in symbiotic foraminifera mainly measure the symbiont density and the photosynthetic activity, which varies with incident light level. Analysis of the time series for the period 1147–1975 AD has revealed an 11-year component, with high significance, that is in phase with the solar cycle and has an average amplitude of 0.04 per mil. The data for the last 250 yr period are shown in Fig. 12 and show a higher amplitude of about 0.08 per mil in recent solar cycles. Estimates indicate that this amplitude is compatible with the variation of sunlight expected from a relative change in cloud cover of about 3% over the solar cycle [44]. This is consistent with the satellite-observed value of $1.7\% / 61\% = 2.8\%$.

It is interesting to note that the solar modulation of *Globigerinoides ruber* continued through the Maunder Minimum when the sunspots disappeared (and with them, disappeared also the sunspot /facular irradiance modulation). However the solar cycle modulation of ^{10}Be continued through the Maunder Minimum, as shown by the ^{10}Be ice core measurements [45]. These observations are consistent with a GCR interpretation for the change of cloud cover, but not with a solar irradiance variation interpretation (§4.1).

3.3 Solar-GCR-climate change in the Industrial Age

3.3.1 Solar-GCR change

The variation of ^{10}Be concentration in the Greenland ice core over the last 300 years (Fig. 13) [46] reveals considerable changes of solar magnetic activity have occurred in recent times. The peaks in GCR

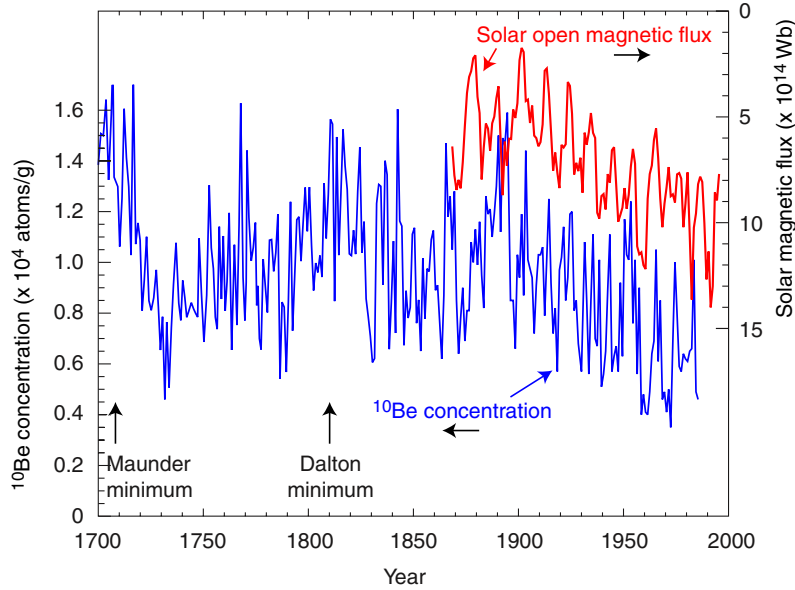


Fig. 13: Variation of ^{10}Be concentration in the Greenland ice core over the last 300 years [46], due to changes of solar magnetic activity (thin line). The variation of the solar coronal source flux, F_S , over the last 140 years is shown by the thick line [11] (note inverted scale and unsuppressed zero). In the period since 1901, the increase of the solar open magnetic flux has been a factor 2.3.

intensity over this period coincide with the cold spells of the Maunder Minimum and Dalton Minimum.

More recently, during the 20th century, the ^{10}Be data show a substantial reduction of the GCR intensity due to a marked increase in the strength of the solar wind. The latter is independently confirmed by the $\langle aa \rangle$ geomagnetic index,⁶ for which there is a continuous record extending back to 1868 and covering 12 sunspot cycles. From the level of geomagnetic activity seen at Earth in the $\langle aa \rangle$ index, Lockwood *et al.* [11] have estimated the source magnetic flux, F_s , that leaves the corona and enters the heliosphere. Their method to derive the coronal source flux has been successfully tested against near-Earth interplanetary space measurements made since 1963, during which time the coronal source flux has been observed to rise a factor 1.4. In the period since 1901, the calculated increase has been a factor 2.3 (indicated by the thick line in Fig. 13).

The open solar flux shows a highly significant anti-correlation with the GCR intensity and so can be reliably used to reconstruct the global GCR intensity over the last 140 years [47]. The data are shown in Fig. 14 (left-hand axis) and indicate a reduction of GCR intensity during the last century by about 20% for the Climax neutron monitors (3 GeV/c cutoff). This implies global average reductions of about 15% at the top of the troposphere, or about 10% at 3 km altitude (§4.3).

Assuming for the moment the existence of a linear relationship between GCR intensity and the low cloud absolute fraction (Fig. 2), then the open solar magnetic flux can also be used to reconstruct the change in low cloud fraction over the same period. The result is shown in Fig. 14 (right-hand axis) [16] and indicates a reduction of low cloud fraction since 1900 by about 1.3% absolute (4.6% relative). Since low clouds are estimated to contribute a net radiative forcing of -17 Wm^{-2} (Table 1), this corresponds to a forcing of about $0.046 \times 17 = +0.8 \text{ Wm}^{-2}$, which is climatically significant (§3.3.5). Overall, these data suggest that a 10% reduction of GCR flux at 3 km altitude is associated with a 4.6% relative reduction of low cloud cover. This would imply a rather high sensitivity of low clouds to GCR intensity, if we assume they are linked.

⁶The $\langle aa \rangle$ geomagnetic index is a sensitive measurement by two antipodal stations of short-term (3-hour interval) variations of the geomagnetic field at the Earth's surface, which is affected by the interactions of the solar wind with the Earth's magnetosphere.

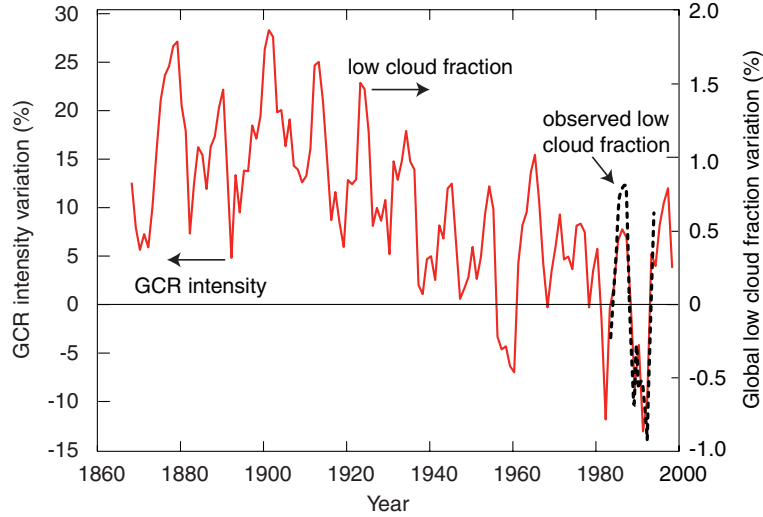


Fig. 14: Estimates of the variations of GCR intensity (3 GeV/c cutoff) and low cloud absolute fraction from 1870 to 2000 (thin line) [16]. The observed variation of low cloud fraction for the period 1983-1994 is also shown (thick dashed line). The estimates are based on the coronal source flux (F_S) measurements over this period obtained from the $\langle aa \rangle$ geomagnetic index [11], and use linear fits between F_S and the observed variations of GCR intensity for the Climax neutron monitor and of low cloud fraction.

There is some evidence that solar forcing may also trigger oscillations in the Earth's climate system, perhaps as secondary processes after changes in the clouds and hydrological cycle. Evidence has been discussed above for a solar forcing of the thermohaline circulation (§3.2.2). Another example may be the anomalous warming (El Niño) or cooling (La Niña) of the surface water in the eastern equatorial Pacific Ocean, which occurs in conjunction with the Southern Oscillation, a see-sawing of atmospheric pressure between the eastern and western tropical Pacific. The combined El Niño Southern Oscillation (ENSO) together with La Niña is the strongest source of natural variability in the Earth's climate system on short timescales, and dominates short-term global temperature anomalies. ENSO is widely viewed as an example of a free internal oscillation of the Earth's climate system, independent of any external forcing. However, Landscheidt [48] has presented evidence that the timings of El Niño, La Niña and the Southern Oscillation over the last 50 years can be linked with phases within the ascending and descending parts of the solar cycle.

3.3.2 Climate sensitivity to radiative forcing

Climate model calculations indicate an approximately linear relationship between global mean radiative forcing, ΔF (Wm^{-2}), and the equilibrium global-mean surface temperature change, ΔT (K),

$$\Delta T = \lambda \Delta F \quad (2)$$

where λ (K/Wm^{-2}) is the climate sensitivity parameter. This parameter is relatively insensitive to the nature of the forcing, for example, greenhouse gases or solar irradiance, provided the forcing agent is not highly variable spatially (like, for example, aerosols). All climate feedback processes, such as changes in water vapour, clouds or ice sheet albedo, are implicitly included in λ .

The value of λ can be inferred from past climate change and from climate models. For example, using ice core samples, between glacial and interglacial periods it is estimated that $\lambda \simeq 5 \text{ K} / 7 \text{ Wm}^{-2} = 0.7 \text{ K/Wm}^{-2}$. Climate models indicate a doubling of the concentration of atmospheric CO_2 from pre-industrial levels (280 ppm) produces $+4 \text{ Wm}^{-2}$ forcing and a mean temperature rise ranging from 1.5 K to 4.5 K, with a central value of 3 K [49]. Therefore $\lambda \simeq (3 \pm 1.5)/4 = (0.75 \pm 0.4) \text{ K/Wm}^{-2}$, in agreement with the previous estimate.

These figures can be compared with the response of the Earth if it were to act as a simple black body. In this case the radiant emittance is $R = \sigma T^4$, where σ is the Stefan-Boltzmann constant. The radiation from a black body varies as $\Delta R/R = 4\Delta T/T$, so that $\Delta T = (T/4R) \Delta R$. Since $\Delta R/R = \Delta I/I$, the fractional change in solar irradiance, it follows that $\lambda_0 = T/4I$. The effective radiating temperature of the Earth is $T \simeq 266$ K, and the global mean solar irradiance reaching the lower troposphere is $I \simeq 0.7 \times 1366/4 \simeq 240 \text{ Wm}^{-2}$ (the factor 0.7 accounts for shortwave albedo and the factor 4 averages the solar irradiance of 1366 Wm^{-2} over the full surface area of the Earth). We thereby estimate $\lambda_0 = 266/(4 \cdot 240) \simeq 0.3 \text{ K/Wm}^{-2}$ for the Earth in the absence of any feedbacks. Therefore the climate feedback factor of the Earth is between a factor of about 1.2 and 4, with a central value of 2.5, and it is greater than one, amplifying the temperature response to a radiative forcing compared with that for a simple black body.

3.3.3 Solar contribution to the current global warming

The reconstructed global mean surface temperature of the Earth between 1860 and 2001 (Fig. 15) [49] indicates a warming of about 0.6 K over this period. A notable feature of the warming is that it did not rise smoothly along with the steadily increasing emissions of anthropogenic greenhouse gases but seemed to flatten, or even reverse sign, during the period 1945–1980.

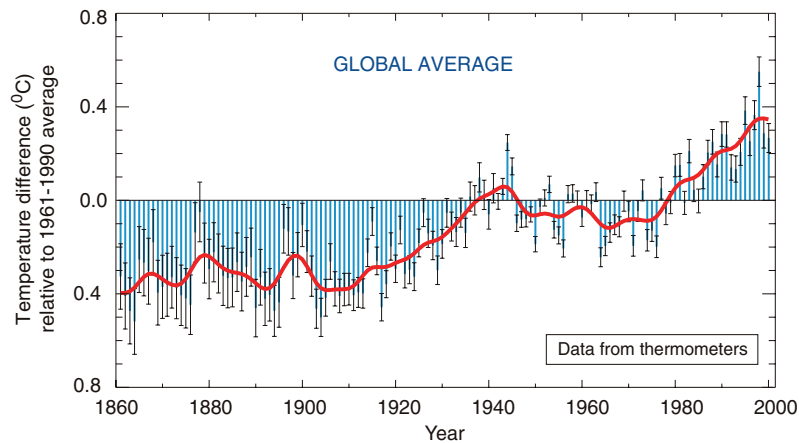


Fig. 15: The global mean surface temperature of the Earth, 1860–2001, relative to the 1961–1990 average [49].

A component of this temperature reconstruction is sea-surface temperatures (SSTs), which have been measured on a routine basis by ocean-going ships since the mid-19th century. The SST record is a particularly valuable measure of global climate since it represents over 70% of the Earth’s surface and is much more spatially and temporally homogeneous than the land surface, as well as being free of such problems as the warming from ‘urban heat islands’. The mean SSTs over the period 1860–1985, for the Atlantic, Pacific and Indian Oceans, are shown in Fig. 16, together with the global mean SST [50]. All of these oceans show a temperature rise that levels off in the same period around 1945–1980, as well as a cooling around the beginning of the last century. Both of these features are characteristic of solar activity, as can be seen in the smoothed sunspot number (Fig. 16) and in the GCR intensity (Fig. 17).

A world-wide simultaneous variation of SST puts severe constraints on a possible forcing mechanism. Since the same characteristic features are seen in all oceans, they are unlikely to be caused by changes such as El Niño events, shifts in wind patterns or changes in the thermohaline circulation, which would lead to differences between the oceans. The mechanism could in principle be increases of anthropogenic greenhouse gases, but the variation in the first half of the 20th century occurred before these were significant. There were insufficient volcanic events to account for the mid-century cooling. The inescapable conclusion is that these data provide quite strong evidence that solar variability was the pri-

mary cause of the warming during at least the first half of the last century. It remains an open question as to what was the solar contribution to the warming during the second half of the 20th century.

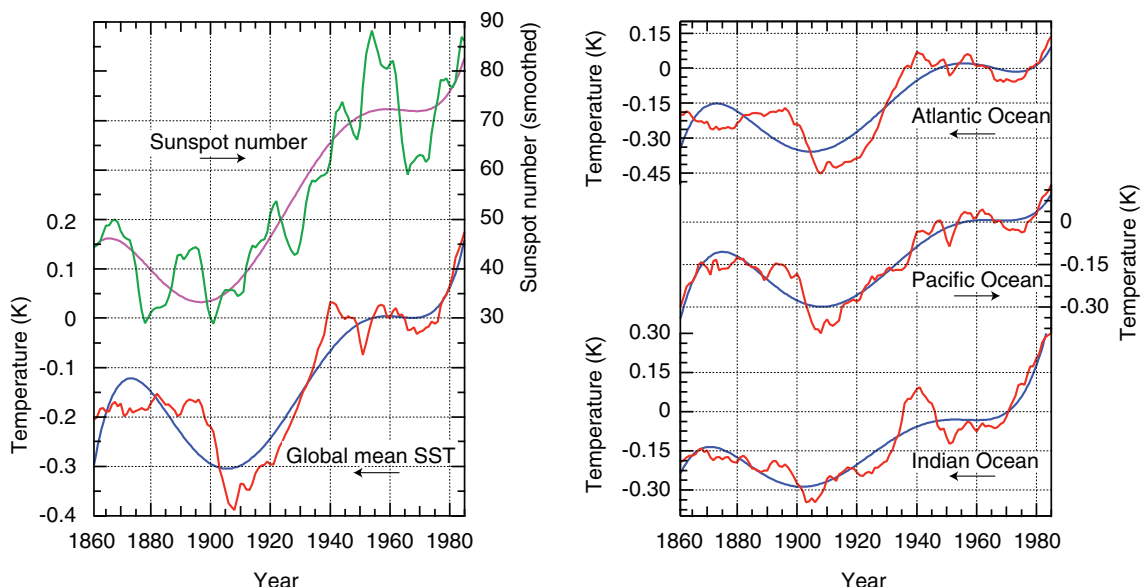


Fig. 16: Annual mean sea-surface temperatures (SST), 1860–1985, for the Atlantic, Pacific and Indian Oceans (right-hand panel) and the global mean (lower curve in the left-hand panel) [50]. The temperatures are shown relative to their 1951–1980 averages. Also shown is the 11-year running mean of the annual sunspot numbers (upper curve in the left-hand panel). The smooth curves are 7th order polynomial fits to the data.

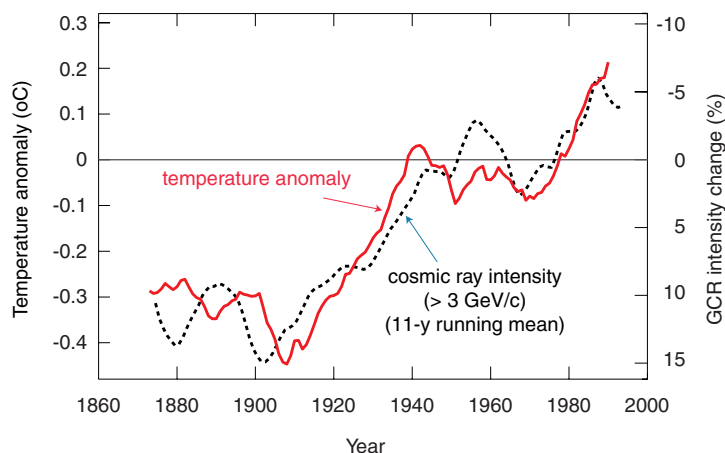


Fig. 17: The mean GCR intensity (3 GeV/c cutoff) over the period 1870–1990, smoothed with an 11-year running mean, together with the global mean temperature anomaly in the same period. The GCR intensity is based on the correlation between directly measured values and the coronal source flux estimates of Lockwood *et al.* (§3.3) [11].

3.3.4 Solar signal in the temperature record

If there is a solar contribution to the current global warming then a solar cycle signal should be present in the temperature record. However the temperature variation is expected to be quite small. The 1.1 Wm^{-2} (0.08%) variation of the solar irradiance (§4.2) corresponds to a global mean variation of $0.7 \times 1.1/4 = 0.2 \text{ Wm}^{-2}$ at the Earth’s surface. This would be expected to produce an *equilibrium* temperature change $\Delta T = 0.7 \times 0.2 = 0.14 \text{ K}$ (i.e. a sinusoidal amplitude of 0.07 K). However the large thermal mass of the surface layer of the oceans will reduce the actual temperature response.

Table 2: Estimation of the ocean's temperature response to a solar cycle (11-yr sinusoidal) forcing of amplitude 0.1 Wm^{-2} , using an RC-circuit equivalent (§3.3.4). The attenuation factor is the ratio of the maximum amplitudes for an 11-yr sinusoidal forcing to a constant forcing.

Climate sensitivity, λ (K/Wm^{-2})	Time constant (yr)	ΔT for constant forcing (K)	Attenuation factor	Solar T amplitude (K)	Phase lag (degrees)
0.3	3.6	0.03	0.43	0.013	64
0.5	6.0	0.05	0.28	0.014	74
0.7	8.4	0.07	0.20	0.014	78

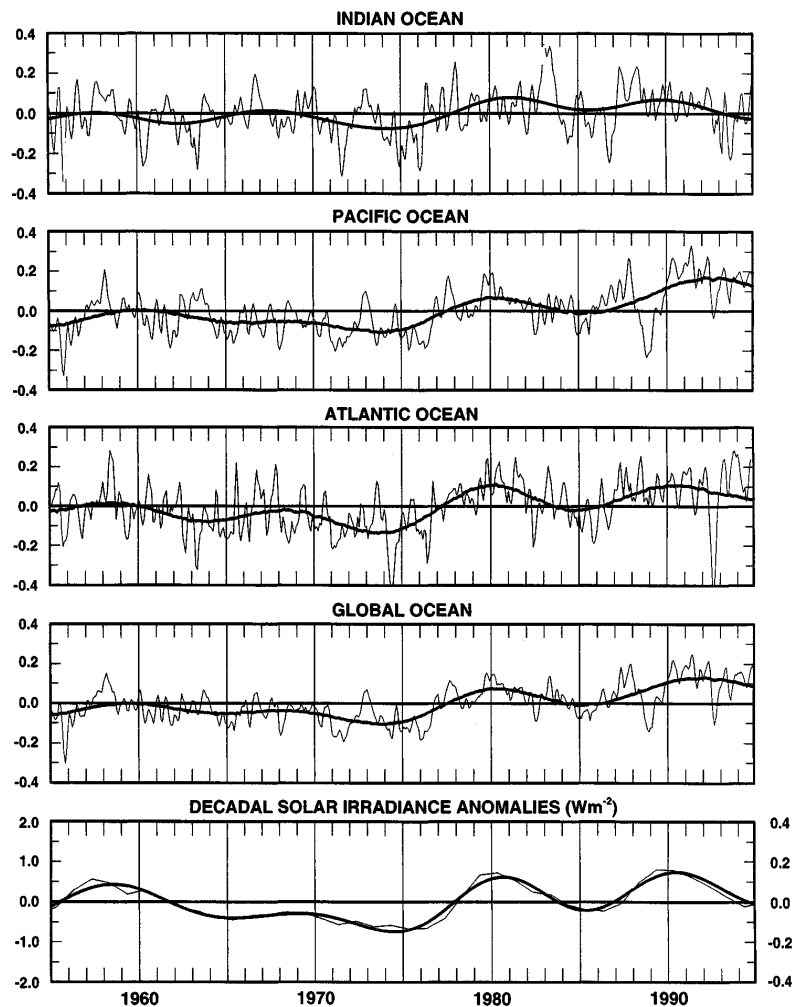


Fig. 18: Sea surface temperature anomalies (Kelvin) from bathythermograph measurements collected from 1955 to 1994 [52]. The light curves in the upper four panels show monthly mean values and the heavy curves show low-pass filtered values (with half-power points at 7 yr). The lowest panel shows the reconstructed solar irradiance over the same period, including the low-pass filtered values (the left hand axis gives the value at the top of the atmosphere and the right hand axis gives the global mean values at the sea surface).

We can estimate the damping effect of the oceans by considering an analogous electromagnetic RC-circuit equivalent for the ocean surface layer: a resistor and capacitor in series, which are set into forced oscillation by a sinusoidal voltage [51]. In this analogy, forcing heat fluxes are analogous to current, temperature to voltage, the ocean to a capacitor and the climate sensitivity parameter to a resistor. The value of the capacitor can be estimated by assuming the ocean can be represented by a well-mixed upper layer of about 90 m depth that is effectively isolated from thermal exchange with deeper water, except by relatively slow diffusion. Table 2 summarises the expected ocean response to a solar cycle (11-yr sinusoidal) forcing of amplitude 0.1 Wm^{-2} . The estimated solar cycle temperature amplitude is 0.014 K , with a phase lag of about $65\text{--}80$ degrees. The temperature response is essentially independent of the climate sensitivity, λ , since the larger constant-forcing temperatures at higher λ are compensated by longer time constants and therefore larger attenuation factors for the short solar cycle. The attenuation factors are large, e.g. a factor 5 for $\lambda = 0.7 \text{ K/Wm}^{-2}$, corresponding to an 8.4 yr time constant.

White *et al.* [52, 53] have analysed the SST data for the period 1900–1991 and the bathythermograph data for 1955–1996. Their analyses reveal convincing solar signals in the Indian, Pacific and Atlantic Oceans which all show comparable amplitudes and phases (Fig. 18). The solar cycle amplitude is $(0.03 \pm 0.005) \text{ K}$ and it lags the changes in solar irradiance by about $0\text{--}65^\circ$. This amplitude is twice the expected value (Table 2). An inter-decadal (18–25 yr) solar signal is also observed with $(0.04 \pm 0.005) \text{ K}$ amplitude and $15\text{--}50^\circ$ phase lag, which is consistent with the expected value from estimated longer-term changes in solar irradiance [53]. In summary, there is indeed clear evidence of a solar signal in the temperature record but the 11-yr oscillation appears to be about a factor two larger than expected, suggesting either an error in the modelling, or else the presence of a mechanism that amplifies the solar forcing.

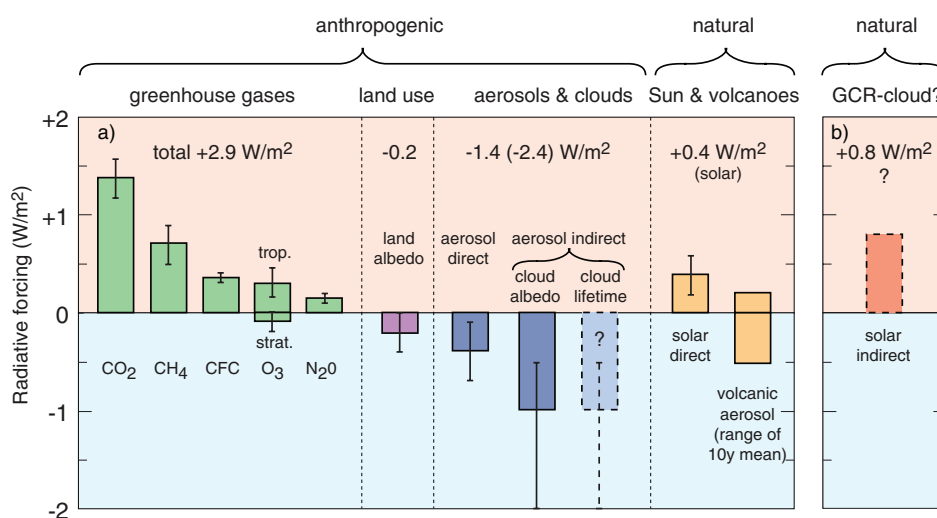


Fig. 19: a) The global mean radiative forcings of the climate for the period from pre-industrial (1750) to present, as estimated by the Intergovernmental Panel on Climate Change (IPCC) [49]. A positive forcing causes a global mean warming, and a negative forcing causes a cooling. The *aerosol indirect* contributions are poorly known, especially the second one (changes of cloud lifetime), and have large uncertainties. b) Estimated radiative forcing due to the putative *solar indirect* effect if GCRs and low cloud cover are causally linked.

3.3.5 Radiative forcings

The global mean radiative forcings of the climate in the Industrial Age, as estimated by the Intergovernmental Panel on Climate Change (IPCC), are shown in Fig. 19a) [49]. Whereas the forcings due to greenhouse gases are known quite well, there are large uncertainties associated with the forcings due to anthropogenic aerosols. The latter are separated into *aerosol direct* effects (albedo and absorption of solar radiation) and *aerosol indirect* effects due to their influence on clouds (the so-called *first indirect*

type is the change in cloud albedo and the *second indirect type* is the change in cloud lifetime). All three of these processes are estimated to contribute a net negative forcing. The change in cloud lifetime is indicated in the figure as being of roughly equal importance as the change in cloud albedo. However both of the indirect aerosol effects are poorly known, and, moreover, new effects have recently been reported [54] that make their large uncertainties even larger. These relate to the effects of surface tension depression of water droplets by anthropogenic organic surfactants, which lead to an increase of droplet number concentration at a given water vapour supersaturation (§4.4).

Ignoring cloud lifetime changes, the total estimated anthropogenic forcing is about $+1.3 \text{ Wm}^{-2}$. In comparison, the estimated solar direct contribution is $+0.4 \text{ Wm}^{-2}$. These figures can be converted to an approximate expected change in equilibrium temperature using Eq. 2: $\Delta T(\text{K}) = 0.7 \cdot \Delta F(\text{Wm}^{-2})$. However Fig. 19 takes no account of the spatial and temporal distribution of the forcings, which are highly non-uniform in the case of aerosols, and so the actual temperature response may be quite different. Nevertheless it is probably reasonable to conclude from Fig. 19a) that the residual anthropogenic forcing is the small difference between two relatively large numbers—a positive forcing from greenhouse gases and a negative forcing of uncertain magnitude from anthropogenic aerosols.

If GCRs are indeed the causal mechanism for the observed changes in cloud cover, then we can estimate the resultant forcing since 1900 to be about $+0.8 \text{ Wm}^{-2}$ (§3.3.1), as shown in Fig. 19b). This *solar indirect* effect is potentially a sizeable forcing—about a factor two larger than the supposed changes in solar irradiance over the same period, and with the same sign. Together with the previous contributions, this would imply a total mean forcing during the 20th century of about $+1.3 \text{ Wm}^{-2}$ anthropogenic and $+1.2 \text{ Wm}^{-2}$ natural (solar). From these figures should be subtracted the cooling effects of the anthropogenic increases in cloud lifetimes, and of volcanoes, respectively. If we take the central value for the climate sensitivity, $\lambda = 0.7 \text{ K/Wm}^{-2}$, then it would imply a larger warming than the 0.6 K observed. This discrepancy could be due to several reasons. On the one hand some of these contributions are poorly known and their estimated magnitudes may change. This includes the possibilities that there is a significant anthropogenic effect on cloud lifetimes and, of course, that the GCR-cloud effect may not exist. Alternatively, or in addition, the climate sensitivity parameter may be less than 0.7 K/Wm^{-2} , in which case the projected anthropogenic temperature increase during the present century would be reduced.

In conclusion significant uncertainties remain in estimating the radiative forcings from anthropogenic and natural sources. The largest uncertainties concern the microphysics of clouds and aerosols. Among these is the possible new contribution due to cosmic ray-cloud interactions. It is clearly important to either confirm or rule out this hypothesis as a natural mechanism for climate change. In the remainder of this paper we will first look at the possible microphysical mechanisms that could be responsible for solar-cloud variability and then describe the proposed CLOUD experiment to test the GCR-cloud mechanism.

4 PHYSICAL MECHANISMS FOR SOLAR-CLOUD VARIABILITY

4.1 The Sun-Earth link

There are only three physical paths that could connect variations of the Sun to the Earth's climate (since neutrinos can be safely ignored!):

1. Solar electromagnetic radiation.
2. Galactic cosmic rays, whose intensity is modulated by the solar wind.
3. Solar wind, and its direct interaction with the troposphere.

The third option is probably not important since the charged particles of the solar wind generally have very low energy (few keV) and so they are easily shielded by the Earth's magnetosphere—that is except over the polar regions, where they range out in the thermosphere at an altitude of about 100 km. This

is far from the tropopause (which lies at an altitude of about 8 km over the poles and 18 km over the tropics). Large coronal mass ejections (CMEs) that aim towards the Earth’s magnetosphere can generate severe magnetic disturbances and cause electron precipitation events in the polar regions. There is an appreciable rate of such events, about 60–80 per year, concentrated about 3 years after the peak of the solar cycle, i.e. with approximately the opposite solar phase as that of the GCR flux [55]. These few-MeV electrons reach altitudes of about 25 km over the polar regions and can influence processes in the the polar stratosphere. Occasionally, very energetic CMEs give rise to so-called solar cosmic rays (SCRs; also known as solar energetic particles, SEPs) of a few $\times 100$ MeV maximum energy, which are thought to be generated by a linear shock acceleration mechanism. These may penetrate to ground level at high geomagnetic latitudes. They are relatively rare, however, occurring at a peak rate of about 3–8 per year around solar maximum (preferentially during the rising and falling part of the cycle), with almost none around solar minimum [55]. During these infrequent events, SCR’s could affect the atmosphere via the same microphysical interactions as GCRs.

In summary, there are only two plausible paths that could connect variations of the Sun with the Earth’s global clouds, namely: 1) solar electromagnetic radiation and 2) GCRs, via solar-wind modulation. We will consider these two candidates in more detail below.

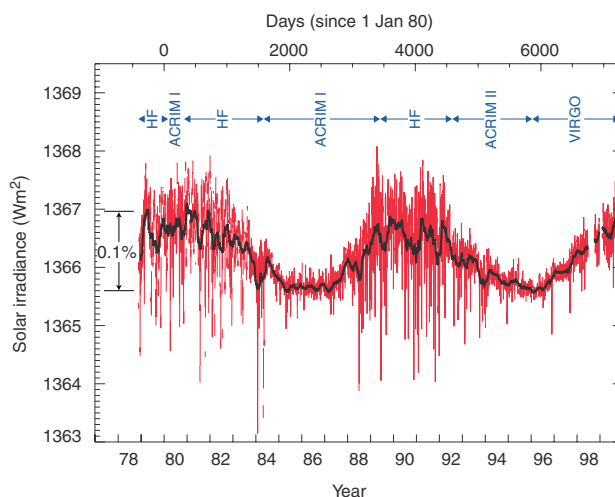


Fig. 20: Total solar irradiance at the top of the Earth’s atmosphere over the last two solar cycles [9]. Sunspot maximum corresponds to peak irradiance. The relatively large and rapid fluctuations are due to sunspots rotating into and out of the field of view.

4.2 Solar electromagnetic radiation

It is natural first to consider variations of the solar irradiance—either in overall intensity or in the distribution of insolation in space and time—as a possible cause of climate change. Indeed there is strong evidence that the Milankovitch theory of climate forcing, due to variations of the Earth’s orbit around the Sun, plays an important role in long-timescale (10–100 kyr) climate change. Milankovitch identified three types of orbital variation that could act as climate forcing mechanisms: tilt of the Earth’s axis, precession of the equinoxes, and eccentricity of the Earth’s orbit around the Sun. Each has its own characteristic periodicity and phase, and these are seen in palaeoclimatic studies. Nevertheless, it seems that orbital forcing mechanisms alone could not account for the magnitude of the observed climatic variations over the past 2 million years. Other mechanisms—such as positive feedbacks or perhaps entirely new mechanisms—need to be invoked

An additional possibility is a variation of the solar irradiance itself. Satellite data (Fig. 20) [9] have shown that the total solar irradiance is indeed varying over the course of solar cycle, but by a tiny amount of about 0.08%. This can be quantitatively well-explained by sunspot darkening and facular brightening

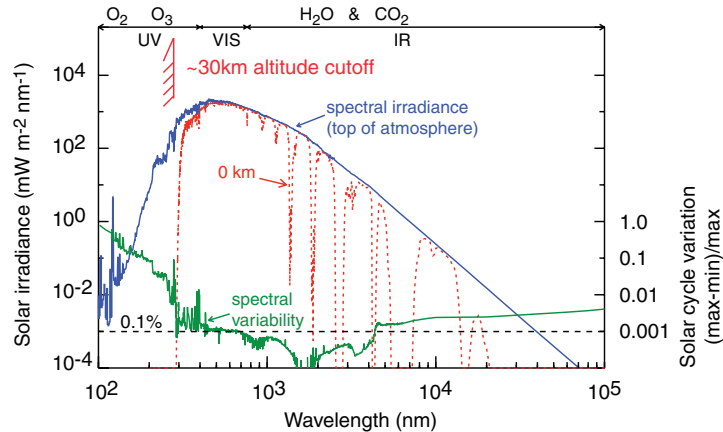


Fig. 21: Wavelength dependence of the solar irradiance at the top of the atmosphere, and its variation from sunspot minimum to maximum. Also shown is the solar spectrum at the Earth’s surface after absorption by the atmosphere.

[56]. Together with measurements of cycling stars similar to our Sun, this had led to estimates of longer-term changes of solar irradiance that appear to be too small to account for the observed climate changes (§1). Attention has therefore focussed on changes of the ultra violet (UV) component of the solar spectrum [57] which, although it carries only a small fraction of the total energy (about 0.1%), shows a much larger variation of several per cent over the solar cycle (Fig. 21).

The UV wavelengths are absorbed at altitudes above 30 km by oxygen (<240 nm wavelength) and ozone (200–300 nm), and cause measurable heating of the thin atmosphere in the upper stratosphere. A positive feedback mechanism exists since the increased UV creates more ozone, although the fractional change is small (about 1–2% from solar minimum to maximum). Modelling reproduces these changes, and studies (e.g. ref. [58]) suggest that circulation changes initially introduced in the stratosphere by this heating can affect circulation at lower altitudes in the troposphere, and therefore can in principle influence cloudiness. It is clearly important to investigate this mechanism further with more experimental and modelling studies.

4.3 Cosmic rays, via solar wind modulation

The other candidate link between solar variability and the Earth’s climate is via GCRs, which are modulated by the solar wind. In contrast with solar UV radiation, GCRs directly penetrate the lower troposphere where the cloud variation is observed, and they have an appreciable intensity variation over the solar cycle.

4.3.1 Solar wind characteristics:

The solar wind is a continuous outward flow of plasma (mainly protons and electrons, with about 5% heavier ions) from the Sun’s corona. As a consequence of its high electrical conductivity, a weak magnetic field is ‘frozen’ into the plasma. The solar wind follows the Parker spiral trajectories out over the huge volume of the heliosphere to distances of 50–100 AU, well beyond the orbit of Neptune. At the Earth’s orbit it has a velocity of 350–800 km s⁻¹ ($\beta = 0.001$ –0.003), an intensity of $(0.5$ – $5) \cdot 10^8$ particles cm⁻² s⁻¹, and a magnetic field of about $5 \cdot 10^{-5}$ Gauss. The main sources of the solar wind on the Sun’s surface include large regions of open magnetic flux known as coronal holes; regions on the photosphere at the boundaries of the supergranulation cells, where magnetic reconnections occur; and coronal mass ejections.

The sunspots themselves are but a visible indication of a high state of magnetic activity of the Sun, when the solar wind is strong. Sunspots are areas of the Sun’s photosphere where strong local

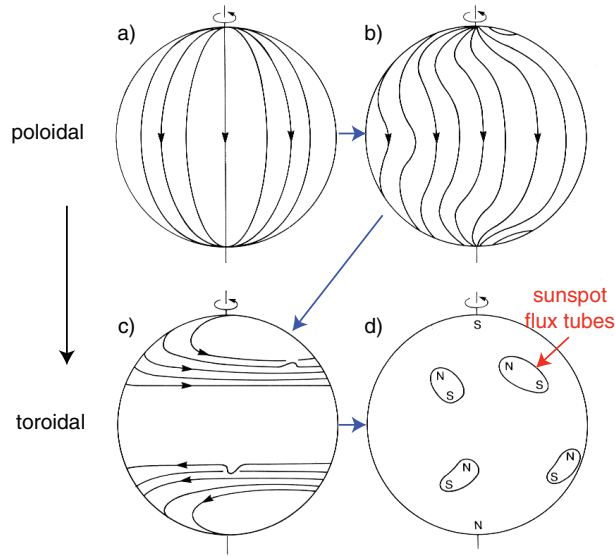


Fig. 22: Babcock's model for generation of sunspot magnetic fields [59]. An initial weak dipole field (a) in the convective zone is wound up into a toroidal field (b) by the Sun's differential rotation. Eventually the field becomes strongly toroidal (c) and magnetic flux tubes rise through the convective zone where they break through the photosphere to form sunspots with opposite polarity in the N and S hemispheres (d).

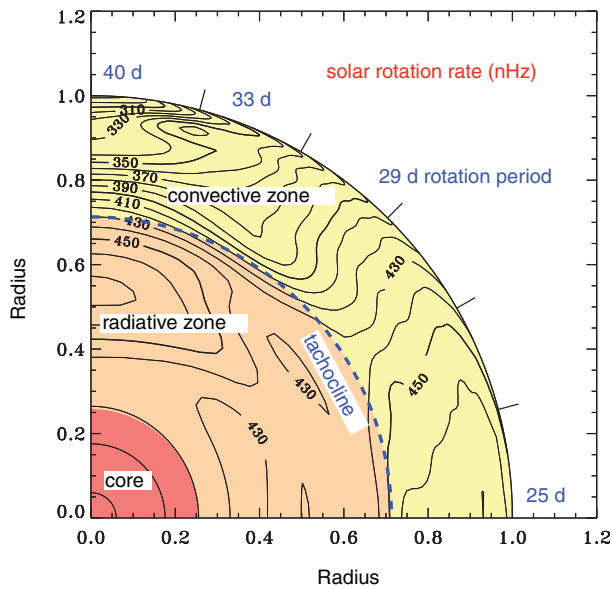


Fig. 23: The interior rotation of the Sun from helioseismology measurements [60, 61]. The fastest rotation is 25 days at the equator and the slowest is over 35 days near the poles. Each contour is separated by 10 nHz (about 0.7 d). The tachocline is the shear region between the radiative and convective zones and is thought to be where the surface magnetic fields originate.

magnetic fields emerge vertically. The fields are typically about 2500 Gauss, to be compared with a mean quiescent photospheric field of below a few Gauss. They appear dark because their temperature is about half of the surrounding photosphere (3,000 K compared with 5,800 K). They are generated (Fig. 22) by the differential rotation of the Sun with respect to latitude and depth: at the surface, one revolution takes 25 days at the equator and over 35 days near the poles (Fig. 23). This transforms the quiescent dipole field into a toroidal field and eventually creates ‘knots’ of strong localised fields. These knots may penetrate the photosphere to form sunspots, which appear cooler due to modification of the normal convective motions of the plasma by the strong magnetic fields. The sunspots first appear at high latitudes and then gradually migrate towards the equator. They eventually disappear by magnetic recombination, leaving a quiescent dipole field once more (but of opposite polarity). The half-cycle from dipole to toroidal and back to (reversed) dipole field is termed the solar, or sunspot, cycle and takes about 11 years on average.

The key to solar variability is a fundamental understanding of the complex solar magnetic fields. How are they generated by the dynamo and what causes their quasi-periodic behaviour? Dynamo action involves the conversion of kinetic energy into magnetic energy by the inductive effects of fluid motion in an electrically conducting fluid—the solar plasma. Babcock’s qualitative picture of the solar dynamo (outlined in Fig. 22) has been known for over 40 years. But it has only been recently with the exquisite helioseismology measurements of GONG and other experiments, and with data from the high-precision spectrometers and detectors on board SOHO, Ulysses, Yohkoh, TRACE and other satellites, that great advances are being made. For example, it appears that the tachocline (Fig. 23) plays an essential role in the solar dynamo, and is the primary region for generation of the magnetic flux before it rises through the convective zone. Also, from high-resolution movies taken of the photosphere and corona, it appears that the dynamics of magnetic flux bundles, and their reconnections, are key to understanding the energy source that heats the solar corona and accelerates the solar wind.

4.3.2 *Modulation of galactic cosmic rays by the solar wind:*

Cosmic rays are generated by supernovae and other energetic sources in our galaxy and beyond. On entering the heliosphere, charged cosmic rays are deflected by the magnetic fields of the solar wind. The transport problem of the GCRs through the heliosphere was first solved by Parker [62] and involves several processes of which the dominant is scattering off the magnetic irregularities, which produces a random walk or diffusion effect. It has been shown theoretically [63] that the effect on the energy of a charged cosmic ray particle in passing through the heliosphere is equivalent to that produced by a heliocentric retarding electric potential with a magnitude at the Earth’s orbit equal to the energy lost by the cosmic rays in interacting with the solar wind. This retarding potential varies between about 1000 MV during periods of very high solar activity and zero during grand minima such as the Maunder Minimum. The solar wind therefore partly shields the Earth from the lower energy GCRs and affects the flux at energies below about 10 GeV. The effective retarding potential over the present eleven-year solar cycle averages about 550 MV, ranging from about 450 MV at the minimum to 850 MV at maximum. This leads to a distinct solar modulation of the GCR intensity (Fig. 24).

The geomagnetic field also partially shields the Earth from GCRs. The dipole field imposes a minimum vertical momentum of about 13 GeV/c at the equator, 3 GeV/c at mid latitudes, and falling essentially to zero at the geomagnetic poles. In consequence, the GCR intensity is about a factor 3.6 higher at the poles than at the equator, and there is a more marked solar cycle variation at higher latitudes. Over the solar cycle, the variation of GCR intensity at the top of the atmosphere is about 15%, globally averaged, and ranges from ~5% near the geomagnetic equator to ~50% at the poles.

At lower altitudes both the GCR intensity and its fractional solar modulation decrease. These are consequences of the absorption of low energy GCRs and their secondary particles by the atmospheric material, which totals about 11 nuclear interaction lengths. Balloon measurements (Fig. 25) show solar cycle variations of about 10% at low altitudes around 3 km (for a 2.4 GeV/c rigidity cutoff).

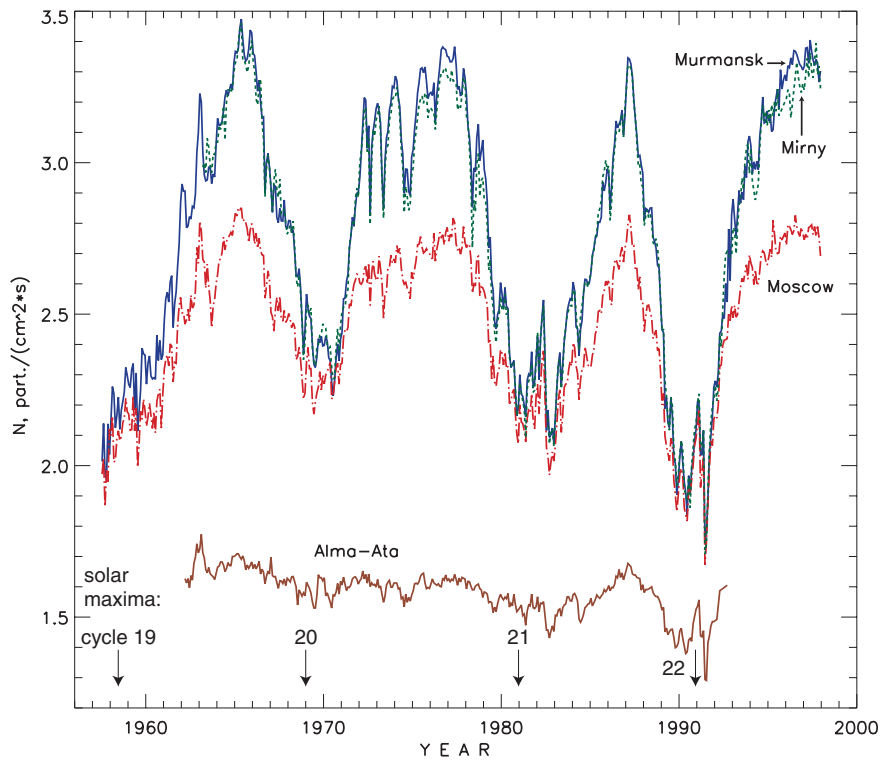


Fig. 24: Balloon measurements of the cosmic ray intensity at shower maximum (15–20 km altitude) for the period 1957–1998, measured by the Lebedev Physical Institute. The curves correspond to four different locations for the balloon flights: Mirny-Antarctica (0.03 GeV/c rigidity cutoff), Murmansk (0.6 GeV/c), Moscow (2.4 GeV/c) and Alma-Ata (6.7 GeV/c). Due to atmospheric absorption, the data of Murmansk and Mirny practically coincide with each other. The approximate times of the sunspot maxima for the last 4 solar cycles are indicated.

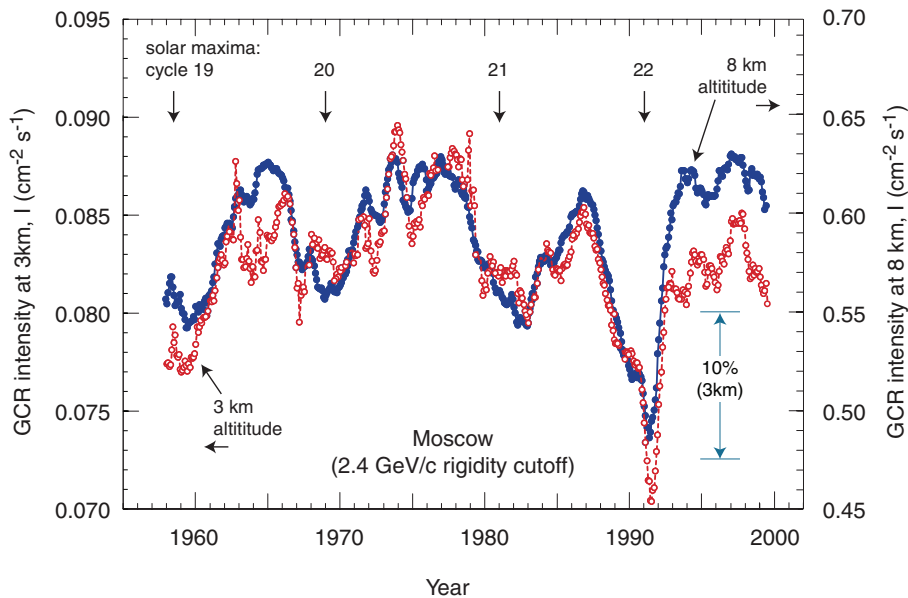


Fig. 25: Balloon measurements of the GCR intensity at 3 km and 8 km altitudes from 1957 to 2000 (2.4 GeV/c rigidity cutoff), measured by the Lebedev Physical Institute. The approximate dates of the sunspot maxima for the last 4 solar cycles are indicated.

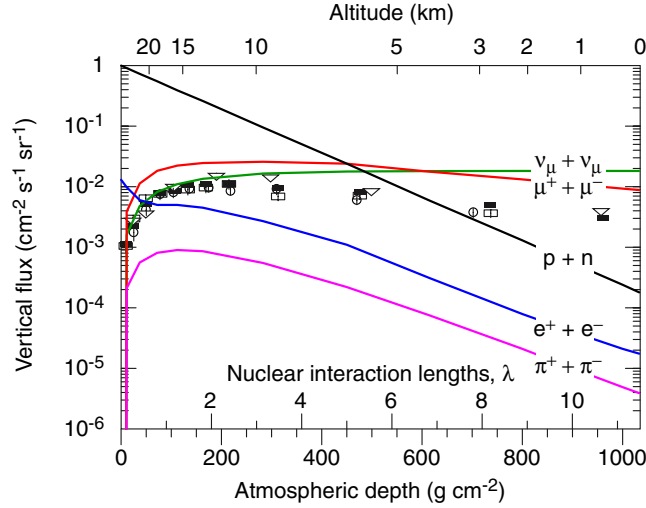


Fig. 26: Vertical fluxes of cosmic rays and secondary particles ($E > 1$ GeV) vs. altitude [64]. The primary nucleons (p and n) include protons, He and heavier nuclei. The points show measurements of μ^- with $E_\mu > 1$ GeV.

4.3.3 Interactions of cosmic rays in the atmosphere:

The composition of the charged primary cosmic rays at the top of the atmosphere is about 98% protons and heavier nuclei, and 2% electrons. Of the former, about 87% are protons, 12% are He nuclei and the remaining 1% are heavier nuclei (especially C, N, O and Fe). The incident nucleons interact with nuclei in the atmosphere and produce secondary particles in hadronic cascades. The initial secondaries are mainly p, n, π^\pm and γ (from π^0 decay), and these subsequently produce μ and e (Fig. 26). Below about 6 km altitude, muons from the decay of π mesons dominate the cosmic ray flux.

The maximum cosmic ray fluxes occur at altitudes of 15–20 km, where the charged particle intensities vary between about 0.8 and $2.3 \text{ cm}^{-2}\text{s}^{-1}$ (at solar maximum, i.e. GCR minimum), depending on geomagnetic latitude (Fig. 27a) [65]. Most of the primary cosmic rays interact in the first $\sim 2\lambda$ material above the tropopause, so the heavily ionising primaries (heavy nuclei) are screened from reaching lower altitudes. Therefore, essentially throughout the troposphere, charged cosmic rays are mostly moderately relativistic singly-charged particles. These particles (other than electrons) lose energy primarily by ionisation of the air molecules, with an ionisation energy loss rate, $dE/dx \sim 1.7 \text{ MeV/g cm}^{-2}$, characteristic of so-called *minimum ionising particles*. In air at stp, the number of ion pairs produced by a minimum ionising particle is $60 \text{ ion-pairs cm}^{-1}$ or, equivalently, the ionisation energy loss is 34 eV per ion pair created. At high altitudes near the GCR maximum the fraction of heavily-ionising non-relativistic particles becomes significant and the mean ionisation density is about $110 \text{ ion pairs cm}^{-1}$, corrected to one atmosphere pressure [65]. At 15 km the density of air is $0.20 \times 10^{-3} \text{ gm cm}^{-3}$ and the mean ionisation density is therefore about $18 \text{ ion-pairs cm}^{-1}$ per charged particle. Therefore the ion pair production rate by cosmic rays at 15 km altitude is $I = 18 \times (0.8\text{--}2.3) = (14\text{--}41) \text{ cm}^{-3}\text{s}^{-1}$, depending on geomagnetic latitude. At 3 km altitude, the GCR flux is about $0.08 \text{ cm}^{-2}\text{s}^{-1}$ at 2.4 GeV/c cutoff (Fig. 25), producing about $3.5 \text{ ion-pairs cm}^{-3}\text{s}^{-1}$. At ground level these are $0.02 \text{ cm}^{-2}\text{s}^{-1}$ and $1.2 \text{ i.p. cm}^{-3}\text{s}^{-1}$, respectively.

Natural radioactivity also contributes to atmospheric ionisation over land. The relative contribution from radioactivity and GCRs as a function of altitude is shown in Fig. 28. During rainfall, radon and its daughter radioisotopes are sedimented from the air, and the ionisation rate measured by plastic scintillation counters close to the ground can increase by up to about 25% [66]. This increase reflects only the γ ray component since α particles are generally not detected because of their short range. However, α particles represent an important component of the radon decay chain. Over oceans, the contribution of radioisotopes is negligible and so, averaged over the entire troposphere, GCRs are by far the dominant source of ionisation (more than 99%).

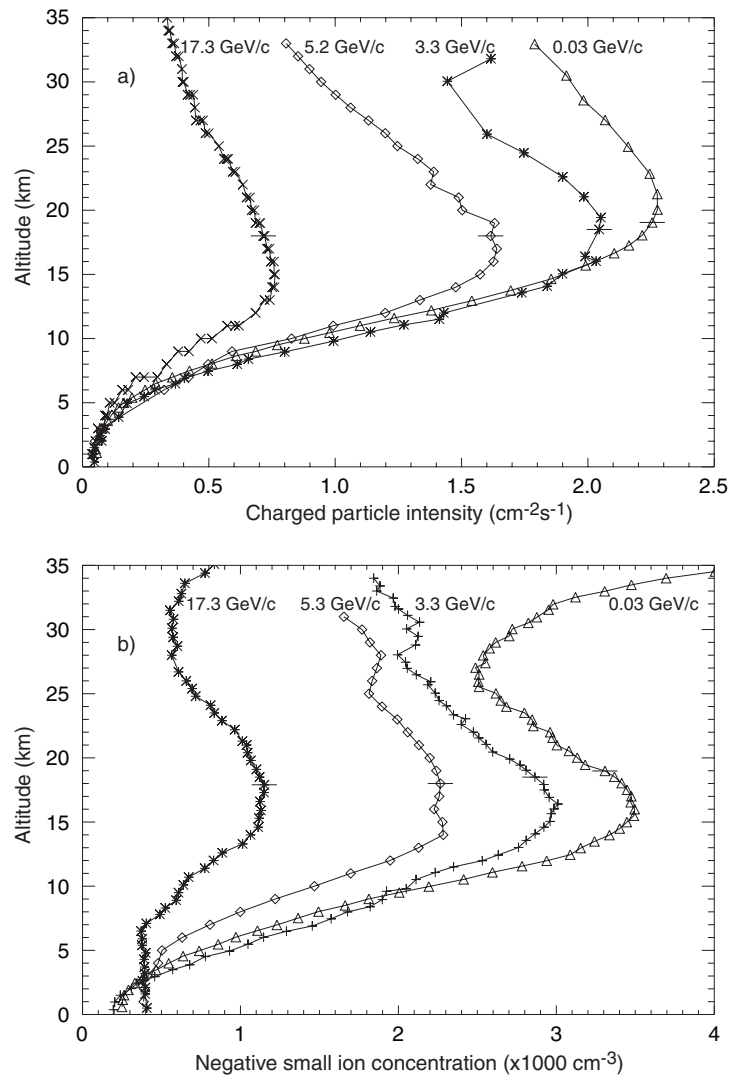


Fig. 27: a) The charged particle intensity and b) the negative small ion concentration vs. altitude, measured at several latitudes with cutoff rigidities, R_c , as indicated. The data were recorded by Lebedev Physical Institute [65] in or near 1990, corresponding to a sunspot maximum (but without solar proton events), i.e. during a cosmic ray *minimum*. The horizontal bars show the typical experimental statistical errors.

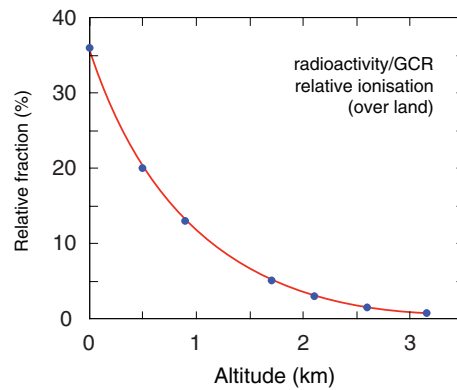


Fig. 28: Relative fraction of atmospheric ionisation from radioactivity and from GCRs as a function of altitude, over land.

Free radicals are also created by galactic cosmic rays, which may lead to a significant source of chemically-reactive molecules in certain regions of the atmosphere. As examples, about 1–2 OH radicals [67] and 1.5 NO molecules [68, 69, 70] are estimated to be produced per ion-pair. Mixing ratios of about 1 pptv OH or NO are therefore generated by cosmic rays per day in the upper troposphere.

4.3.4 Evolution of ions in the atmosphere:

The ions and free electrons created by cosmic rays rapidly interact with molecules in the atmosphere and convert to complex positive and negative cluster ions [71]. Primary positive ions are mostly N_2^+ , O_2^+ , N^+ , and O^+ . Free electrons rapidly ($\tau \sim 200$ ns) attach to O_2 , leading to O_2^- as the most important primary negative ion. Both positive and negative primary ions experience rapid ion-molecule reactions with relatively abundant atmospheric gases, leading to the cluster ions $\text{H}_3\text{O}^+(\text{H}_2\text{O})_n$ (formation time ~ 1 ms and n peaking around 4–6) and $\text{CO}_3^-(\text{H}_2\text{O})_n$ (formation time ~ 10 ms).

More complex cluster ions form on slightly longer timescales. The positive ions react further with basic molecules B possessing proton affinities larger than that of H_2O , leading to $\text{H}^+\text{B}(\text{H}_2\text{O})_n$. Important examples for B are ammonia, forming $\text{NH}_4^+(\text{NH}_3)_m(\text{H}_2\text{O})_n$, and acetone $(\text{CH}_3)_2\text{CO}$. Negative ions react with acidic molecules, particularly H_2SO_4 and HNO_3 , leading to $\text{HSO}_4^-(\text{H}_2\text{SO}_4)_l(\text{HNO}_3)_m(\text{H}_2\text{O})_n$ and $\text{NO}_3^-(\text{HNO}_3)_m(\text{H}_2\text{O})_n$. The above species have been observed in the upper troposphere and lower stratosphere by aircraft-based ion mass spectrometers [72, 73]. As a consequence of the chemical differences between positive and negative ion clusters in the atmosphere, the latter have a slightly higher (20%) electrical mobility.

These ion clusters form on timescales of order 1–100 s, depending on the trace gas concentrations. The formation time can be simply estimated as follows. At stp, the mean free path of air molecules is $0.067 \mu\text{m}$ and the rms velocity is 500 ms^{-1} , so the collision rate per molecule is about $500/(0.067 \cdot 10^{-6}) \sim 10^{10} \text{ s}^{-1}$. Therefore the mean time interval for a trace gas molecule to collide with an ion is about 0.3 s at 1 ppbv concentration ($3 \cdot 10^{10} \text{ molecules cm}^{-3}$ at stp), and about 300 s at 1 pptv ($3 \cdot 10^7 \text{ molecules cm}^{-3}$).

The evolution of an embryonic cluster ion competes with ion “loss” mechanisms such as ion-ion recombination, ion-aerosol attachment and, in clouds, ion-droplet attachment. Away from clouds, the production rate of ions by cosmic rays, I [ion-pairs $\text{cm}^{-3}\text{s}^{-1}$] is in equilibrium with the loss rate according to

$$I = \alpha n^2 + \beta nN \quad (3)$$

where n [cm^{-3}] is the small ion concentration of one sign (and we assume $n \simeq n_+ \simeq n_-$), α [cm^3s^{-1}] is the ion-ion recombination coefficient (about $1.6 \times 10^{-6} \text{ cm}^3\text{s}^{-1}$), β [cm^3s^{-1}] is the ion-aerosol attachment coefficient (which varies with aerosol size and charge) and N [cm^{-3}] is the aerosol number concentration. If we assume for the moment that the principal removal mechanism is ion-ion recombination, then the expected equilibrium ion density (of one sign) at 15 km altitude is $n = \sqrt{I/\alpha} = \sqrt{(14 - 41)/(1.6 \times 10^{-6})} = (3 - 5) \cdot 10^3 \text{ cm}^{-3}$. The measured negative small ion concentrations vary between 1000 and 3500 cm^{-3} at 15–20 km altitude (Fig. 27b), depending on cutoff rigidity. These values are between a factor 2–3 smaller than the estimated ion concentrations assuming loss by recombination alone. Similar factors of about 3 occur at lower altitudes in unpolluted air, which indicates that the scavenging of small ions by aerosol particles is an important loss mechanism at all tropospheric altitudes.

From Eq. 3, the *recombination* lifetime of an ion is $\tau = 1/\alpha n$. This implies ion lifetimes due to recombination of about 3–10 min, depending on altitude. When attachment dominates, the ion lifetime is given by $\tau = 1/\beta N$, and typical values are about 100 s. These lifetimes set the timescale within which processes such as ion-induced nucleation must take place if they are significant. The small ions drift vertically in the electric field created by the negatively-charged Earth and the positively-charged ionosphere. The field strength is $E \sim 100 \text{ V/m}$ at ground level, producing an drift velocity for small

ions of about 1.5 cm s^{-1} and a drift distance of up to about 10 m over their lifetime. At an altitude of 15 km, however, $E \sim 2 \text{ V/m}$ due to the higher conductivity of the air. Here the drift velocity is only 0.1 cm s^{-1} and the drift distance is less than 1 m. Ions can be transported substantially further by winds, both vertically and horizontally.

Small ions are very efficiently scavenged by cloud droplets in a similar way to the aerosol attachment described above. This results in a sharp reduction of the electrical conductivity, σ , inside clouds (since $\sigma = n e \mu$, where e is the electronic charge and μ the electrical mobility). Measured conductivities inside non-electrified clouds are reduced by a factor 10–40 relative to clear air, and by a factor 200–500 inside electrically active clouds ($E = 30 \text{ kV m}^{-1}$) [102]. Therefore, to a good approximation, almost all the charge in a cloud can be assumed to reside on the droplets. The difference between the conductivities of clear air and of clouds causes a layer of space charge to form at the cloud boundary regions, which generates a relatively high vertical electric field, E_z , inside the cloud and maintains continuity of the vertical conduction current, $J_z = \sigma E_z$. Typical equilibrium droplet charges at cloud boundaries are quite large—about $100 e$ —and take about 13 min to be established by the positive and negative ions drifting into the cloud from below and above, respectively. Updrafts and downdrafts can carry these highly charged droplets and aerosols deeper inside the clouds.

4.3.5 Physical mechanisms

Having summarised the general characteristics of the interaction of cosmic rays with the atmosphere, we will now consider how these interactions may influence cloud microphysics. These processes fall into three categories, as shown schematically in Fig. 29:

1. Aerosols.
2. Ice particles.
3. Cloud electricity.

Each of these processes is discussed in detail below.

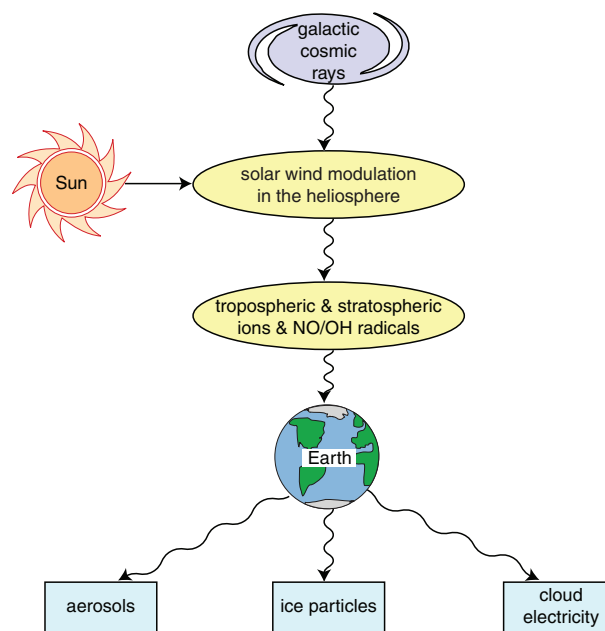


Fig. 29: The three categories of cloud processes that may be affected by galactic cosmic rays, whose intensity is modulated by the solar wind.

4.4 Aerosols

4.4.1 Cloud condensation nuclei

Atmospheric aerosols are liquid or solid particles suspended in the air. The atmosphere contains significant concentrations of aerosols, sometimes as high as 10^6 cm^{-3} . Aerosol composition varies significantly with respect to location and size distribution, with the smallest aerosols often being clusters of volatile species such as sulphuric acid and water (formed from gas-to-particle conversion) and the largest often being inorganic salts and dust particles (§4.4.3). Aerosol sizes can often be described by three quasi-distinct modes comprising a *nucleation* mode (diameter range $\sim 1\text{--}100 \text{ nm}$), an *accumulation* mode ($\sim 0.1\text{--}1 \mu\text{m}$) and a *coarse* mode ($>1 \mu\text{m}$).

Many different kinds of aerosol are capable of acting as condensation nuclei (CN) but only a subset constitute cloud condensation nuclei (CCN). These can activate into cloud droplets when the relative humidity exceeds 100% (or, equivalently, when the water vapour supersaturation, S , exceeds 0%). The presence of a largely abundant supply of CCN ensures that the maximum water vapour supersaturations in the atmosphere rarely exceed values of about 1% since higher values are arrested by the removal of water vapour during droplet growth. These values are far below the supersaturations ($\sim 500\%$) required to activate small ions into droplets, as in a classical Wilson cloud chamber [74]. Therefore, if GCRs can affect clouds, it is *a priori* likely to be through some influence on CCN.

The activation process can be understood from the Köhler curves (Fig. 30), which show the equilibrium S (and therefore equilibrium vapour pressure) over droplets of various sizes and containing various masses of dissolved salts. The equilibrium S of pure water droplets increases with decreasing radius due to the effect of curvature (Kelvin's equation; $\ln(p/p_0) \propto \sigma/r$, where p/p_0 is the water vapour saturation ratio, σ is the air-water surface tension and r is the droplet radius). However dissolved salts reduce the equilibrium S due to a reduction of the molar concentration of the water (Raoult's law; $p/p_0 \propto -1/r^3$). The latter effect dominates at small radii, i.e. at high solute concentrations. Recent studies [54] have also indicated the importance of organic surfactants in reducing the surface tension of cloud droplets, resulting in an increase of droplet number concentrations at lower supersaturations

The droplet number density in liquid water clouds depends upon the cooling rate of the air as it enters the cloud (since this affects the peak S that is reached) and upon the concentration, size and chemical composition of the CCN. Although highly variable, typical number densities are a few $\times 100 \text{ cm}^{-3}$ in continental clouds and a few $\times 10 \text{ cm}^{-3}$ in marine clouds (Fig. 31). Number densities are usually higher in convective clouds than in stratiform clouds.

Once activated, droplets grow by diffusion of water vapour. Diffusional growth is rather slow and it is unusual for droplet radii to exceed $20\text{--}30 \mu\text{m}$ by this process. Cloud droplets typically attain sizes of $10 \mu\text{m}$ within a few minutes but take over an hour to reach $100 \mu\text{m}$ (since the growth time $\propto r^2/S$). Droplet collision and coalescence (which occurs when droplets collide while falling under the influence of gravity) takes over as the principal growth mechanism for radii above about $20 \mu\text{m}$.

For clouds to generate rainfall, some drops must grow to precipitable sizes of 1 mm or greater. This is achieved either by collision and coalescence of droplets or by ice formation (glaciation). Ice formation usually occurs in only a small fraction of the cloud droplets, allowing these to preferentially grow by vapour diffusion due to the lower vapour pressure of ice compared with water droplets (§4.5.1). The ability of a cloud to generate rain is an important factor in determining its lifetime.

4.4.2 Effect of CCN changes on cloud radiative properties and lifetime

Cloud radiative properties: The effect that a change in the CCN number concentration has on the radiative properties of a cloud can be quantitatively estimated as follows [76, 77]. Assuming the liquid water content and depth of the cloud is fixed, then its optical thickness, τ , is given by $\tau \propto Nr^2$, where N is the droplet number concentration and r the mean droplet radius. Since $N \propto r^{-3}$, this indicates $\tau \propto N^{1/3}$. Therefore a change of the droplet number concentration by ΔN leads to a change of the

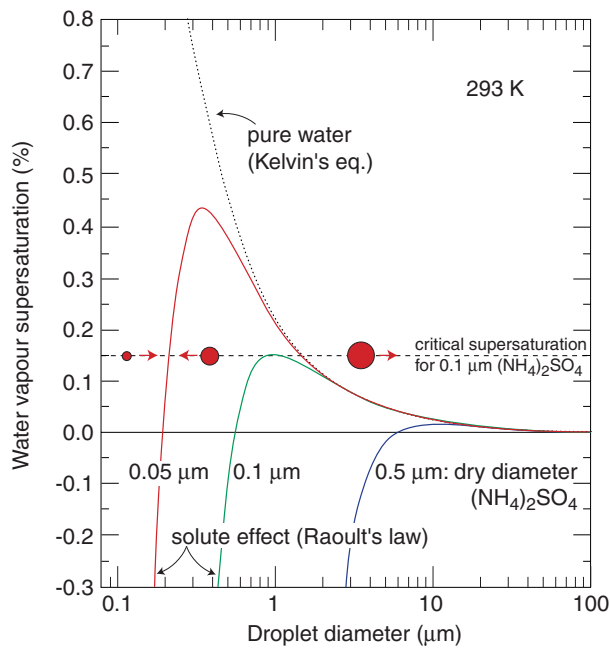


Fig. 30: Köhler curves showing the equilibrium water vapour supersaturation at 293 K for droplets of pure water (dotted curve) and for droplets containing various masses of dissolved (NH₄)₂SO₄ (solid curves) vs. diameter of the droplet [75]. The water vapour supersaturation, S (%) = $(p/p_0 - 1) \cdot 100$, where p is the partial pressure of the water vapour and p_0 is the saturated vapour pressure over a plane surface of water at this temperature. In the indicated example, an ambient water vapour S of 0.15% (dashed line) exceeds the critical value for all ammonium sulphate aerosols with dry diameter $\geq 0.1 \mu\text{m}$. These aerosols will therefore activate and grow into cloud droplets, whereas smaller aerosols remain as unactivated haze particles. Droplets below their corresponding equilibrium curve will shrink by evaporation whereas those above will grow by condensation (the indicated droplets correspond, for example, to a dry diameter of $0.05 \mu\text{m}$).

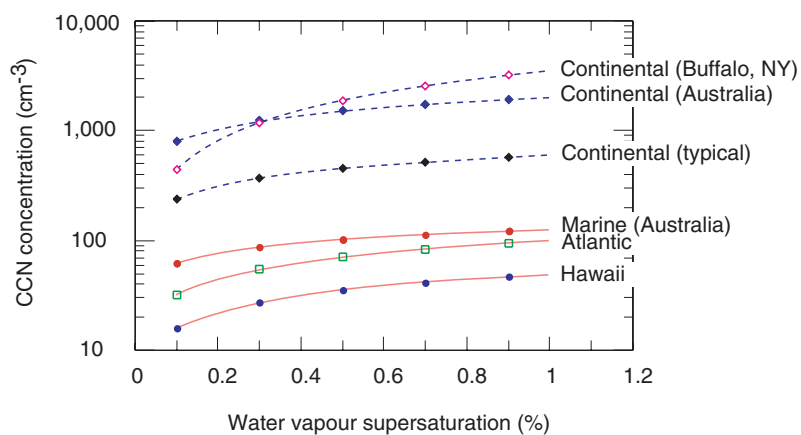


Fig. 31: Measurements of CCN concentrations at several sites: marine (solid curves) and continental (dashed curves) as a function of the water vapour supersaturation [75]. The CCN concentrations are equal to the cloud droplet concentrations at a given supersaturation.

optical thickness by $\Delta\tau$ given by

$$\frac{\Delta\tau}{\tau} = \frac{1}{3} \cdot \frac{\Delta N}{N} \quad (4)$$

The albedo (reflectivity), A , of a cloud is the fraction of incident radiation that is reflected into the backward hemisphere. For the scattering of solar radiation by clouds [76, 77],

$$A \approx \frac{\tau}{\tau + 6.7} \quad (5)$$

Differentiating Eq. 5 and combining with Eq. 4 gives,

$$\frac{\Delta A}{A} = (1 - A) \cdot \frac{\Delta N}{N} \quad (6)$$

The rather thin stratiform clouds that cover an appreciable fraction of the Earth's surface, and especially marine regions, have an albedo of about 0.5 and a droplet number concentration of about 100 cm^{-3} or less (Fig. 31). Equation 6 shows that these clouds are very sensitive to changes in the CCN number concentration; their reflectance changes by 0.5% or more per single additional cloud droplet per cubic centimetre of air! This in turn indicates that GCR-induced changes in the CCN number concentration of only a few per cent could produce significant effects on the radiative properties of such clouds. As a numerical example, Figure 32 shows the variation of cloud reflectivity with cloud depth and cloud droplet number density for a fixed liquid water content of 0.3 g m^{-3} (most data confirm that there is little or no dependence of liquid water content on cloud droplet number density).

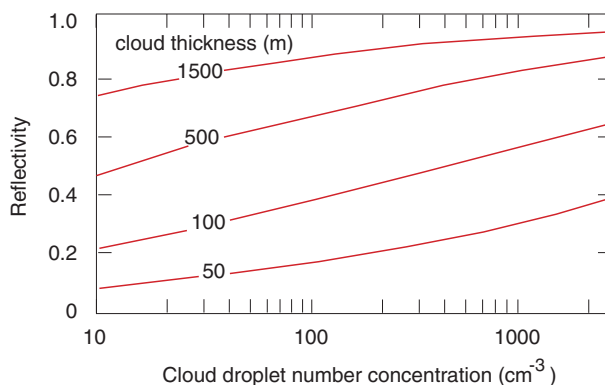


Fig. 32: The variation of cloud albedo with cloud thickness and droplet number concentration for a fixed liquid water content of 0.3 g m^{-3} [78].

Cloud lifetime: The second effect of an increase in CCN concentration is to suppress rainfall and thereby to increase cloud lifetime. This has been observed over oceans in ship tracks [79] and, recently, also over land [80]. The latter study used NOAA satellite data to investigate clouds that formed downwind of industrial sites located in pristine areas. The otherwise uniform cloud data from these regions was streaked with bright (highly reflective) cloud plumes from the industrial sites. The droplets in these plumes were found to be more numerous than the nearby regions and of a smaller diameter—typically less than $10 \mu\text{m}$ and therefore below the threshold size for them to coalesce efficiently and precipitate. In contrast the droplets outside the plumes measured more than $25 \mu\text{m}$ in diameter. The high reflectivity of the plumes resulted from the high droplet number density at fixed liquid water content (Fig. 32). Independent analysis of data from the Tropical Rainfall Measuring Mission confirmed that these plumes did indeed produce less rain and therefore had a longer lifetime than clouds in the nearby regions.

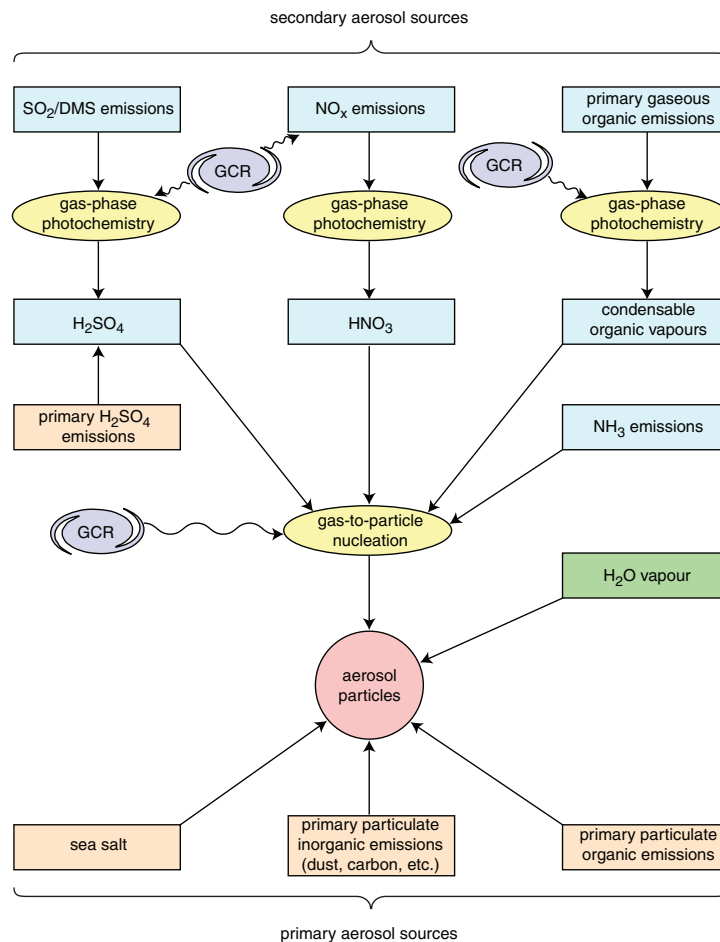


Fig. 33: The main sources of atmospheric aerosols. These are classified as *primary* if they are injected directly into the air or *secondary* if they result from gas-to-particle conversion in the atmosphere. The processes that may be affected by galactic cosmic rays (GCRs) are indicated by wavy arrows.

4.4.3 Production and loss of atmospheric aerosols

The main sources of atmospheric aerosols are summarised in Fig. 33. Aerosols are classified as either *primary* or *secondary*, and they may be of either natural or anthropogenic origin. Primary aerosols are those injected directly into the air (e.g. by wind erosion, sea spray, pollen, etc.). Secondary aerosols are those created by gas-to-particle nucleation of vapour molecules. Inorganic aerosols are usually weakly acidic, with the most common aqueous cationic components being H⁺, ammonium and sodium, and with common anionic components being sulphate, chloride and nitrate. Such aerosols are hygroscopic. Aerosols can also be partly or wholly composed of organic compounds derived from plant waxes and combustion sources. These aerosols may be either hygrophobic or hygroscopic. Secondary aerosols may originate from emissions of non-condensable vapours followed by gas-phase chemical conversion into condensable (i.e. low vapour pressure) aerosol precursors, e.g. SO₂ oxidation to H₂SO₄. Clouds are an important source as well as sink of aerosols since they provide efficient sites for the scavenging and chemical processing of aerosols and aerosol precursors. These pathways for aerosol production are shown in Fig. 34.

Once formed, aerosols have both a *direct* and an *indirect* effect of the climate by their influence on radiative forcing. The direct effect is due to scattering and absorption of the incoming shortwave radiation. Absorptive aerosols such as black carbon have a net positive forcing (warming) and reflective ones such as hygroscopic aerosols have a net negative forcing. The indirect effect is due to a change

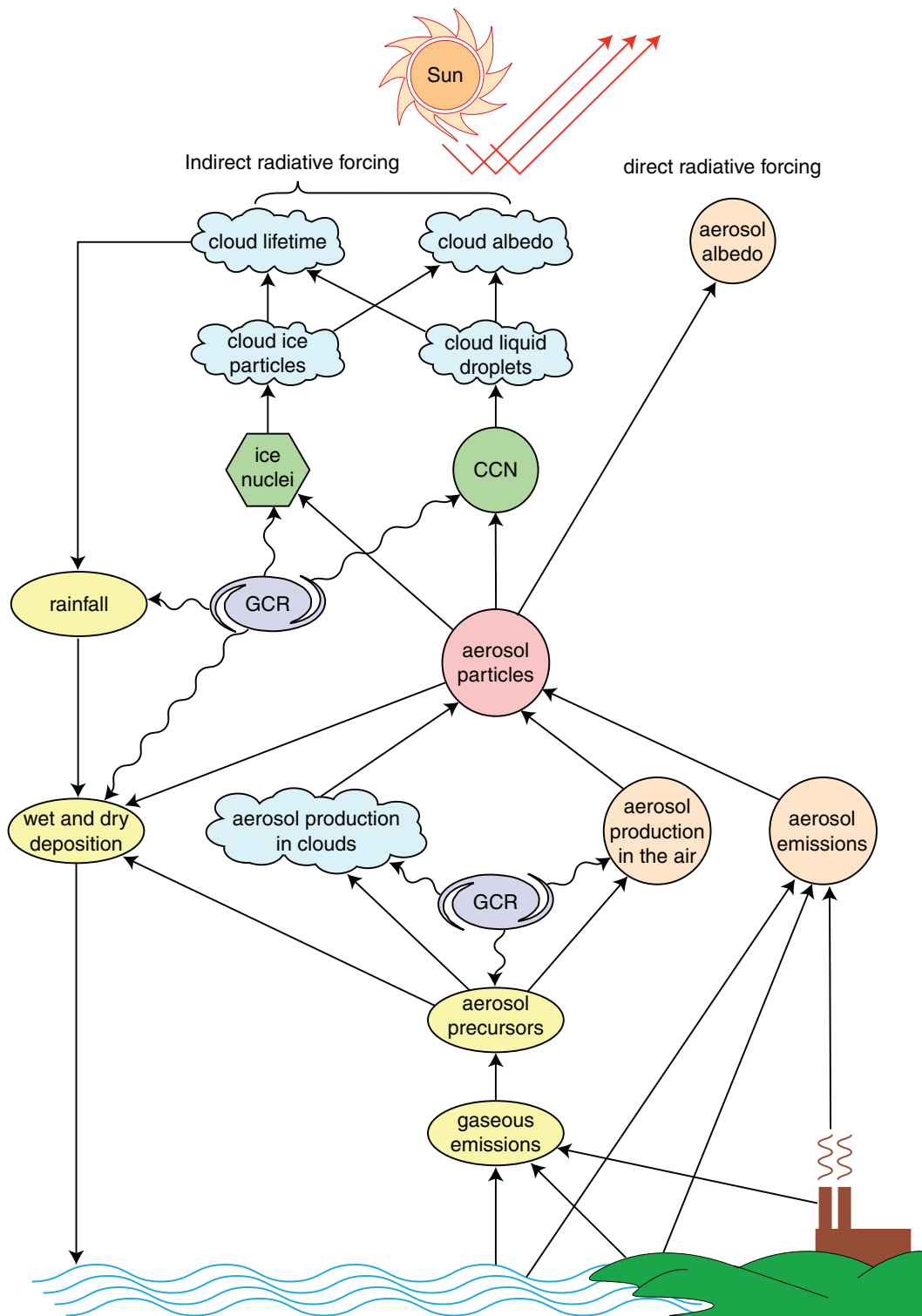


Fig. 34: Aerosol production and loss in the atmosphere, and the effects on clouds and climate of changes in the aerosol number concentration. The processes that may be affected by galactic cosmic rays (GCRs) are indicated by wavy arrows.

in the number concentration of CCN and, consequently, to changes in cloud reflectivity and lifetime (§4.4.2). A small subset (about 10^{-6}) of CCN constitute ice nuclei and these have a strong influence on cloud radiative properties and lifetime (§4.5.1).

Aerosols are removed from the atmosphere by wet or dry sedimentation (Fig. 34). The residence time depends on their composition, size and geographical location. The lifetime for wet removal is about 8 days in the lower troposphere and about 3 weeks in the middle to upper troposphere. This applies to the accumulation-mode aerosols, which constitute the predominant CCN type. Larger aerosols have a shorter residence time due to large settling velocities. Smaller aerosols are removed relatively rapidly by coagulation, which transforms them into fewer, larger aerosols. A consequence of the short lifetime of tropospheric aerosols is a large spatial variation of their composition, size and number concentrations.

4.4.4 GCR-aerosol interactions

If increase of the GCR flux were to translate into increases in CCN number concentration then this would extend cloud lifetimes, consistent with the satellite data (Fig. 2). There are several processes in the production and loss of aerosols that may be affected by GCRs, as indicated in Figs. 33 and 34.

An important source of new aerosol particles in the atmosphere is the nucleation of ultrafine condensation nuclei (UCNs) from precursor vapours, of both natural and anthropogenic origin. Despite intensive research over several decades, the origin of the ubiquitous background of ultrafine aerosols in the troposphere has not yet been determined. Moreover even the fundamental mechanism that leads to new particle formation remains poorly understood. Understanding these processes is crucial to determining the contributions of both natural and anthropogenic aerosol effects on radiative forcing of the climate. It has been suggested that ionisation from GCRs may play a key role in the formation of new aerosol particles [81]–[88].

An important precursor vapour for UCN and CCN is sulphuric acid. However the classical theory of binary $\text{H}_2\text{SO}_4\text{--H}_2\text{O}$ homogeneous nucleation fails to explain observations of new ultrafine particle formation in clean regions of the lower atmosphere, such as occurs over oceans and in pristine continental air [89]–[92]. Typically the nucleation rates predicted by classical theory are far lower (by as much as 10 orders of magnitude) than the experimentally-observed rates. Recent modelling work [93]–[95] demonstrates that thermodynamically-stable charged clusters, caused by vapours condensing onto ions, can form at much lower ambient vapour concentrations and grow significantly faster than neutral clusters.

The steps involved in the creation of CCN from condensable vapours (in this case, sulphuric acid) are shown in Fig. 35. Molecular $\text{H}_2\text{SO}_4\text{--H}_2\text{O}$ clusters form and evaporate continually by kinetic motion. Under suitable conditions some clusters will reach the critical size of about 1–2 nm diameter. Once the critical size is reached, continued growth of the cluster becomes preferential thermodynamically. The nucleation of aerosols in the atmosphere involves several competing processes which include molecular clustering, evaporation, scavenging of condensable vapour by pre-existing aerosols, and sedimentation by rainfall. In this environment, electrically charged embryos have a competitive advantage over neutral embryos. Charged clusters provide additional electrostatic attractive forces with polar molecules, allowing critical embryos to form with fewer molecules than for neutral clusters. Therefore ions can greatly enhance the rate of formation of new particles in regions where the concentration of condensable vapours is too low for stable neutral clusters to form at an appreciable rate, such as frequently occurs in the marine boundary layer. The key parameters controlling the rate of new particle formation are the concentrations of condensable vapours, the GCR ionisation rate, and the surface area of pre-existing aerosols.

Once formed, the UCN continue to grow. The main growth process up to diameters of about 10 nm is molecular condensation; for larger sizes the main growth mechanism is coagulation of existing CN. During the growth of CN to CCN, other vapours in the atmosphere such as ammonia, nitric acid and low volatility organic compounds are known to be important. The growth rate is expected to be enhanced by the presence of ions, becoming less significant with increasing size of the aerosol particle.

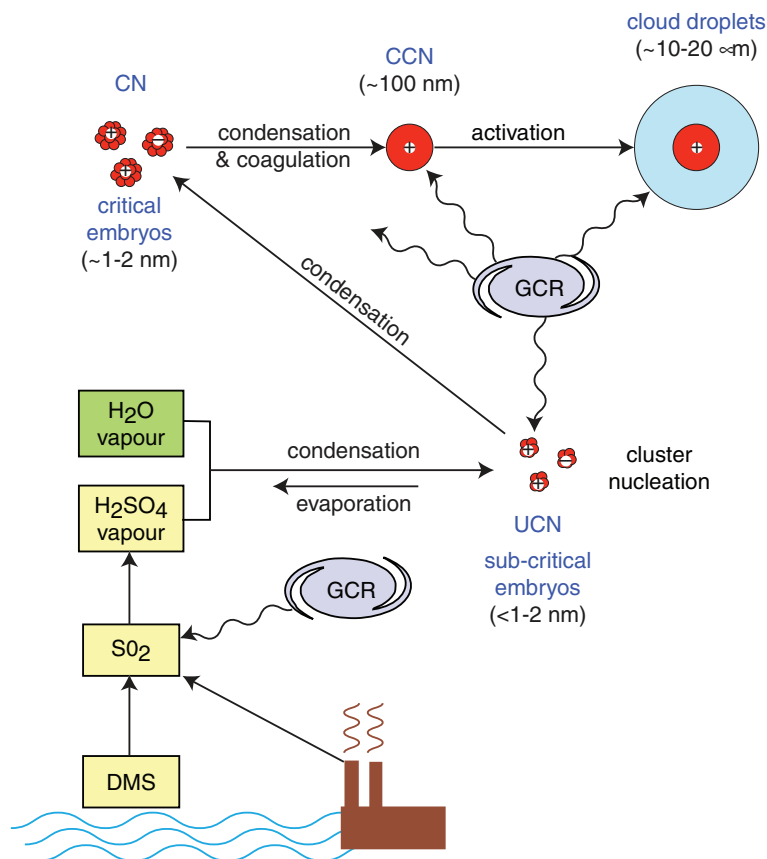


Fig. 35: The nucleation of ultrafine condensation nuclei (UCN) from trace sulphuric acid vapour, followed by aerosol growth into condensation nuclei (CN) and cloud condensation nuclei (CCN), which can activate into cloud droplets. The precursor of sulphuric acid is SO_2 , produced anthropogenically or, in remote marine environments, predominantly from dimethyl sulphide (DMS) from plankton. The processes that may be affected by galactic cosmic rays (GCRs) are indicated by wavy arrows. Charged aerosols are expected to have an enhanced growth rate and reduced evaporation relative to neutral aerosols. GCRs may also affect the activation of CCN in droplets.

This enhancement is largest when the two colliding particles have opposite sign (+−), but there is also an increased rate between one charged and one neutral particle (+0 or −0) due to image charge attractions, when compared with two neutral particles (00). Finally the activation of CCN into cloud droplets may also be influenced by charge. However the effect is expected to be small for the typical electric charges found on aerosol particles under fair-weather conditions.

There is a continual interchange between charged and neutral particles as small ions diffuse onto existing CN and CCN, either neutralising or charging them in the process. This implies that ions may potentially affect the production rate of a large fraction of the CCN produced by gas-to-particle conversion, regardless of whether or not the original UCN were produced via ion-mediated processes.

Aerosol particles and trace vapours are continually being scavenged from the atmosphere by rainfall. Following a precipitation event, the air is left with a relatively low aerosol particle number concentration and a low aerosol surface area. It is under these conditions that the nucleation of fresh UCN is most likely to occur, provided sufficient concentrations of precursor vapours are present. Furthermore, for clean environments, this indicates that the *rate* at which new particles are produced and grow into CCN can strongly affect the lifetime and radiative properties of the clouds in these regions, since there is rarely sufficient time for a large CCN population to form.

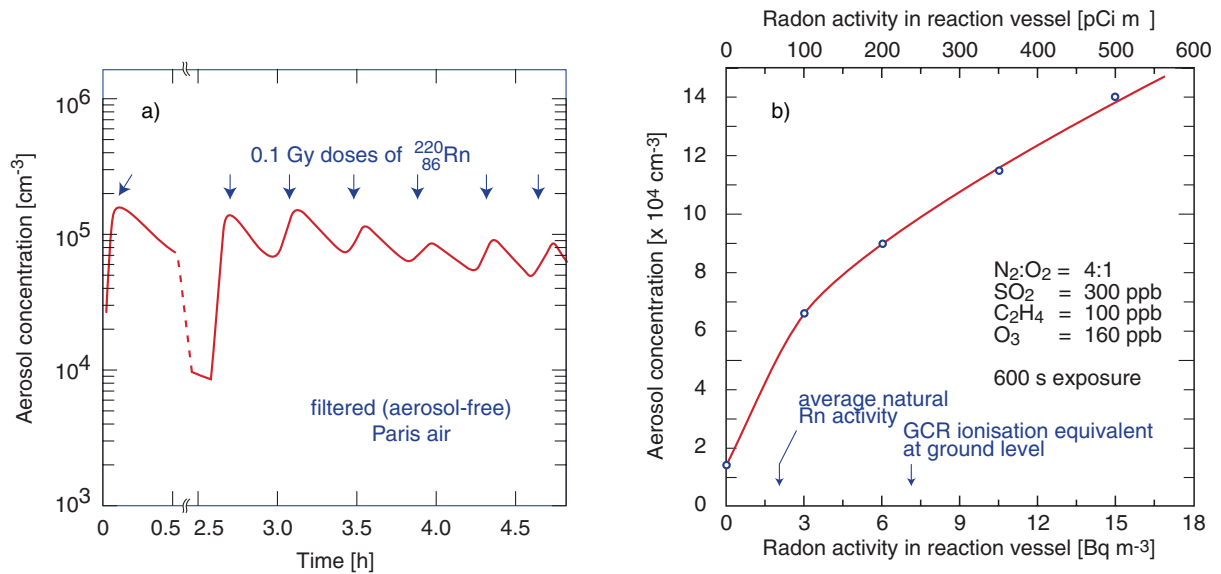


Fig. 36: Experimental evidence for enhanced nucleation of aerosols from trace gases caused by ionising radiation (α particles from radon). The measurements involve a) filtered Paris air with high irradiation doses [97] and b) artificial air with high trace gas concentrations but at naturally-occurring radiation doses [98].

4.4.5 Experimental knowledge of ion-aerosol interactions

There are only sparse experimental data on the effect of ions in the atmosphere on new particle formation—and none, to our knowledge, on the effect of ions on particle growth from CN to CCN, or on the activation of CCN into cloud droplets. Observations have been made of nucleation bursts of CN in the atmosphere that cannot be explained by classical theories. For example, Hörrak *et al.* [96] reported the spontaneous formation of bursts of intermediate size ions in urban air, which they suggest may be due to ion-induced nucleation. Also, Clarke *et al.* [89] observed formation of new ultrafine particles in the tropical marine (Pacific) boundary layer⁷ that could not be explained by classical binary (H₂SO₄-H₂O) homogeneous nucleation theory at the measured low ambient concentrations of sulphuric acid (1–5·10⁷ molecules cm⁻³).

However a recent study by Yu and Turco [93, 95] based on an ion-mediated model *is* able to reproduce the observations of Clarke *et al.*. Their model indicates that the nucleation rate of fresh CN in the marine boundary layer is generally limited by the available ion production rate from GCRs. In contrast, the nucleation rate in the upper atmosphere is generally limited by the trace vapour concentration since the temperatures are lower and the trace vapour saturation ratios correspondingly higher, and so binary homogeneous nucleation can occur at an appreciable rate. This provides a possible reason why the solar modulation signal only appears in clouds below about 3 km.

Direct experimental evidence that ions are involved in the nucleation of new particles under atmospheric conditions is lacking. However positive effects with ions have been seen. Two examples are shown in Figs. 36. Bricard *et al.* [97] observed new particle production in filtered (aerosol-free) Paris air exposed to very high radiation doses ($3 \cdot 10^8$ Bq m⁻³ × 300 s). On the other hand, Vohra *et al.* [98] carried out experiments with radon at naturally-occurring ionisation levels of 3–15 Bq m⁻³ and observed new particle production proportional to ionisation rate, but they used artificial air containing high concentrations of trace gases (300 ppb SO₂, 100 ppb C₂H₄ and 160 ppb O₃).

⁷The *boundary layer* is the layer of the atmosphere within about 1 km of the Earth's surface, within which air is subject to turbulence, friction effects and surface heating. The region extending from about 1 km to the tropopause is known as the *free troposphere*.

4.5 Ice particles

4.5.1 Overview

The second class of processes by which GCRs may affect clouds concerns ice particles. The formation of ice in clouds is important for several reasons:

1. In mixed-phase clouds, ice particles grow rapidly at the expense of liquid water droplets. When the water vapour supersaturation relative to liquid water is 0%, the supersaturation relative to ice is much higher, as large as 50% (Fig. 37). Since supersaturation is the ‘driving force’ that determines growth rate, the ice particles grow rapidly, reducing the ambient supersaturation and causing the liquid water droplets to evaporate. The ice particles rapidly grow to a size where they sediment and ‘rain-out’ the clouds below.⁸
2. The freezing of supercooled water releases latent heat, which affects cloud dynamics.
3. Ice particles modify the radiative properties of clouds, both by increasing sedimentation and by changing the reflectivity (particles that are larger and crystalline). As an indication of the importance of ice particles, the IPCC finds that uncertainties in the fraction of frozen water in clouds results in differences of up to 17 Wm^{-2} in their estimates of globally-averaged cloud forcing [49].

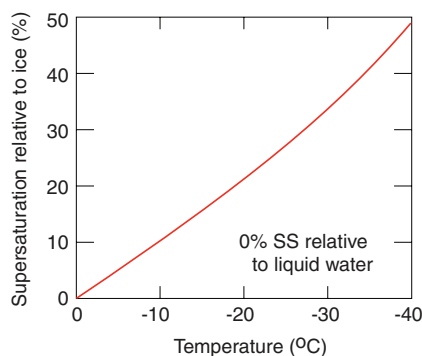


Fig. 37: The supersaturation relative to ice for water vapour that is in equilibrium with liquid water at the temperature indicated on the x axis. In the temperature range from 0°C to -40°C liquid water can exist in a supercooled state in the absence of an ice nucleus. Below -40°C , water freezes homogeneously, i.e. without need of a distinct ice nucleus.

Once a cloud extends to altitudes where the temperature is below 0°C , ice crystals may form. Two phase transitions can lead to ice formation (Fig. 38): (a) the direct deposition (sublimation) of water vapour to ice and (b and c) the freezing of a supercooled liquid droplet. The latter may occur either by the transformation of a supercooled liquid droplet into an ice particle (freezing nucleation) or collision of a supercooled liquid droplet with an ice nuclei (contact nucleation). The relative importance of these three freezing modes has not yet been established. The freezing may proceed either on a suitable ice nucleus (IN) (heterogeneous nucleation), as shown in Fig. 38, or else occur with pure water (homogeneous nucleation). For homogeneous nucleation to take place, a statistical fluctuation of the water molecules must occur to produce a stable, ice-like structure that can serve as an ice nucleus. For typical cloud droplet dimensions ($\sim 10 \mu\text{m}$), homogeneous nucleation occurs at about -40°C . Therefore, for clouds in the temperature range between 0°C and -40°C , ice particle nucleation is heterogeneous and requires a suitable IN.

Field measurements show that clouds contains a great deal of supercooled liquid water in this temperature range, since ice nuclei are very rare in the atmosphere. The IN number concentration is

⁸This is the principle behind cloud-seeding, in which a suitable ice nucleus material, e.g. AgI, induces the freezing of supercooled cloud droplets.

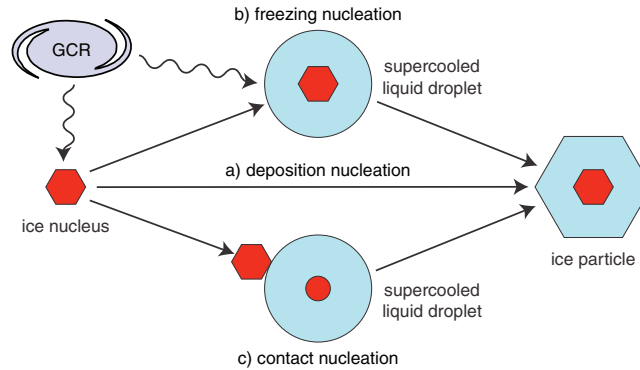


Fig. 38: Processes for ice particle formation in clouds, involving a) *deposition nucleation*: the direct sublimation of water vapour to the solid phase on an ice nucleus, b) *freezing nucleation*: condensation of a supercooled liquid droplet on a suitable ice nucleus, followed by freezing as the temperature falls, and c) *contact nucleation*: the freezing of a supercooled liquid droplet (already formed on a CCN) by external contact with an ice nucleus. GCRs may affect the efficiency of CCNs to act as ice nuclei and may directly affect freezing nucleation.

typically ~ 1 /litre at about -20°C , increasing by a factor 10 for each 4°C of additional cooling. This may be compared with a CCN number concentration of $\sim 10^6$ /litre or, said another way, only one in a million CCN constitutes a suitable IN at -20°C . Identifying such particles is a difficult task, and IN remain poorly understood. However efficient IN are generally insoluble in water and have a chemical bonding and crystallographic structure similar to ice. Examples include various insoluble salts (such as AgI), certain clay particles, various organic materials and even bacteria (which is surprising since neither of the last two have a crystalline structure). Although some IN have been identified, the number of ice particles found in clouds often exceeds the measured IN concentration by several orders of magnitude.

Part of this discrepancy can be attributed to ice multiplication in secondary processes. When a droplet freezes at temperature between about -5°C and -10°C , mechanical stresses lead to the ejection of small ice fragments, which in turn act as efficient ice nuclei. Collisions between dense graupel particles and fragile dendritic crystals also generate ice splinters. However the understanding of these processes is also poor. Recent measurements [99] suggest that freezing-thawing cycles may be an important factor in IN production. These authors found that the thermal history of droplets affects the temperatures at which they eventually freeze. In summary many observations of ice particle in clouds cannot be explained quantitatively and, in particular, there appears to be a great lack of IN to account for the observed numbers of ice particles in clouds.

4.5.2 GCR-ice particle interactions

Enhanced heterogeneous ice nucleation by electrification has been proposed by several workers. For example, Tinsley [100, 101] has proposed that cosmic rays have an important influence on cloud microphysics and climate through the following sequence of events. Cosmic rays generate ionisation in the atmosphere and determine the magnitude of the vertical conduction current, J_z (§4.6.2). This current generates highly charged droplets ($\gtrsim 100 e$) at the upper and lower boundaries of clouds due to the accumulation of space charge (§4.3.4). When these droplets evaporate they leave behind highly charged aerosols which are coated with extra sulphates and insoluble organic compounds scavenged while the droplet existed. (It has been shown that neither electric charge nor aerosol material is not lost when a droplet evaporates.) These highly charged and coated “evaporation nuclei” constitute efficient ice nuclei, either by deposition nucleation or by contact nucleation. The presence of charge enhances collisions of the evaporation nuclei with other liquid droplets by “electroscavenging” (§4.6.2), thereby generating ice particles in clouds.

If this sequence of events is correct, it would imply that increased GCR intensity leads to in-

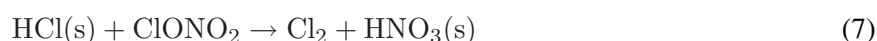
creased ice particle formation in clouds, which in turn releases latent heat, increases cyclone activity, and increases rainfall. Tinsley supports his claim with a study of the Vorticity Area Index (a measure of regional-scale cyclone motion) which he shows to decrease during Forbush decreases of the GCR flux. During Forbush events, which are caused by severe solar disturbances (CMEs), the GCR intensity is reduced by around 3–10% over a period of about 1 day, with a recovery time of a few days.

The key uncertainty in this sequence of events is whether or not charged aerosols, perhaps together with cloud processing, are more effective as IN. An enhancement due to charge is supported by very little experimental work so far [102]. Some experiments, but not all, have reported positive effects. For example, cloud chamber experiments at the University of Missouri-Rolla in 1980 [103] found an enhancement of frozen droplets at -33°C in regions where cosmic rays traversed the cloud chamber. The presence of ions was observed to raise the threshold temperature for homogeneous ice nucleation by about 2 K. The cloud-processing and evaporation aspects of Tinsley's scheme are qualitatively supported by recent studies [99] which indicate that the morphology of any crystallised solid in an aerosol strongly influences its effectiveness as an IN. In summary, there seem to be some indications that ionising radiation may affect ice nucleation, but the experimental picture is far from clear.

4.5.3 Polar stratospheric clouds

Polar stratospheric clouds (PSCs), also known as mother-of-pearl or nacreous clouds, play a key role in the process of ozone depletion in the polar regions, especially Antarctica. PSCs are clouds that form in the cold polar stratospheric winters where, despite the dryness of the stratosphere, the temperature drops low enough for condensation and freezing to occur. At temperatures above the ice frost point the particles may be either liquid solutions of nitric acid, sulphuric acid and water, or else solid nitric acid and ice in the molar ratio 1:3—so called nitric acid trihydrate (NAT). In other parts of the world the stratosphere is too warm for these clouds to form, which is one reason why the 'ozone hole' is confined to the Antarctic region.

Normally, the chlorine of anthropogenic ozone-depleting chemicals is locked up in relatively inert and stable chlorine compounds. However, during the Antarctic winter months (June to August) when the region receives no sunlight and is isolated by a wind circulation pattern known as the polar vortex, the stratosphere becomes cold enough (190–195 K) for PSCs to form. The PSCs provide a heterogeneous catalytic surface on which chlorine can be converted from inert 'reservoir' species, such as ClONO_2 and HCl , into active species:



In the presence of sunlight the Cl_2 photolyses, producing free Cl atoms which react with the ozone, thus destroying the ozone layer. Since the reaction requires sunlight, it only begins when the sunlight returns in the Antarctic spring (September to October), before the PSCs have had a chance to evaporate. The ozone hole disappears again when the Antarctic air warms up enough during late spring and summer

In reaction 7, the Cl_2 is released but the HNO_3 remains in the PSC particles. Since gaseous HNO_3 can convert active chlorine to reservoir species, this further facilitates ozone destruction. In fact, massive ozone depletion requires the abundance of gaseous HNO_3 be very low. In principle this nitric acid would be liberated when the PSCs evaporated with the return of the Sun. However, *freezing* of PSC particles allows the selective growth of a small number of particles that subsequently become large enough to sediment out of the stratosphere (§4.5.1). This process leads to *denitrification* of the polar stratosphere and to a strongly enhanced ozone loss.

Several mechanisms are recognised to be important for the formation of solid polar stratospheric cloud particles. However, these persistent and optically very thin ice clouds cannot be explained by any recognised mechanism. Since PSCs exist on a very large scale, often over several thousand kilometre regions with little spatial variability, the possible mechanisms are tightly constrained. Laboratory experiments have excluded the possibility that these solid particles form by crystallisation of the liquid

aerosols due simply to large scale cooling. An intriguing suggestion—so far unexplored—is that these clouds form by deposition nucleation of nitric acid and water directly onto cosmic ray-generated ions or ion clusters (process ‘a’ in Fig. 38). However, to date, there have been no experiments that can confirm or dispute this possibility. An understanding of the freezing mechanism of PSCs appears to be critical to a complete understanding of denitrification and ozone loss [104].

4.6 Cloud electricity

4.6.1 Overview

The third class of processes by which GCRs may affect clouds concerns the electrical nature of the atmosphere. Except for a contribution from radioactive isotopes near the land surface (Fig. 28), GCRs are responsible for generating all the fair-weather atmospheric ionization between ground level and the mid mesosphere, at about 65 km altitude (§4.3.3–4.3.4). As such, GCRs fundamentally underpin the global electrical circuit.

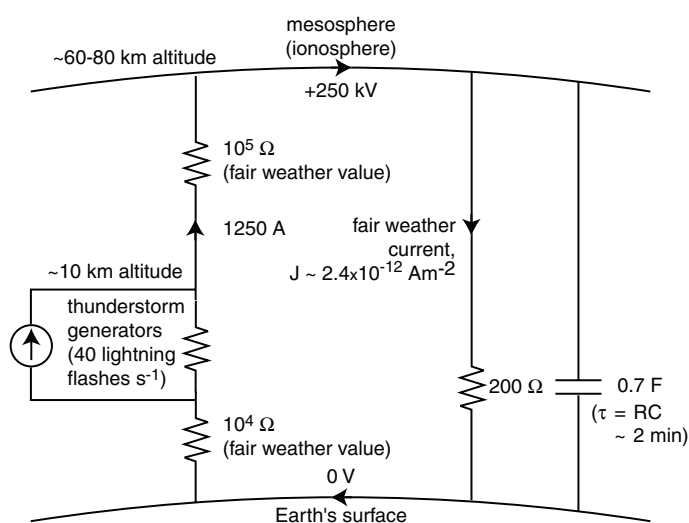


Fig. 39: Schematic of the global electrical circuit. The current generator is thunderstorms, which are predominantly located in the tropics, and the return path is the global fair-weather current flowing between the ionosphere and ground. GCRs play a central role in these processes.

The atmospheric electric circuit (Fig. 39) involves a global current of 1250 A which is sustained by thunderstorms continuously active around the tropics. The Maxwell current density below thunderstorms is the sum of several components, of which lightning contributes about half, with the remainder from electrical conduction, air convection, and precipitation. The thunderstorms carry negative charge to the ground and an equivalent positive current flows up to the ionosphere. Due to the high currents, large electric fields are generated above thunderstorms and the air is likely to break down electrically. Indeed, during the last decade, optical flashes known as sprites and elves have been detected above thunderstorms carrying the positive current up to the ionosphere. This current generator maintains the ionosphere at a relative positive potential of about 250 kV. Since the ionospheric potential drives a fair weather current of 1250 A, it represents a very powerful continuous generator of about 300 MW. The return current between the ionosphere and the Earth’s surface flows throughout the atmosphere, in regions of disturbed and undisturbed weather, and is carried by vertical drift of small ions. The average fair-weather current density, $J = 2.4 \text{ pA m}^{-2}$. Electric fields present in the atmosphere vary between fair weather values of typically 100 Vm^{-1} at the surface and about 2 Vm^{-1} at 15 km altitude (due to the higher air conductivity). In clouds, the electric fields are generally $\lesssim 500 \text{ Vm}^{-1}$ but reach about 100 kVm^{-1} in thunderstorms before a lightning discharge (Table 3).

Table 3: Typical *maximum* electric fields measured inside clouds [105]. Electric fields in all types of non-thunderstorm clouds are generally $\lesssim 0.5 \text{ kVm}^{-1}$. For comparison, the clear-air electric field is about 0.1 kVm^{-1} at ground level. The dielectric field strength of dry air at stp is 30 MVm^{-1} , which represents the breakdown field strength across plane parallel electrodes. The threshold electric fields for wet air breakdown are about $1\text{--}2 \text{ MVm}^{-1}$.

Cloud type	Maximum electric field (kVm^{-1})
stratus	1
stratocumulus	1.5
cirrostratus	1.5
altostratus	5
nimbostratus	15
thunderstorm	100

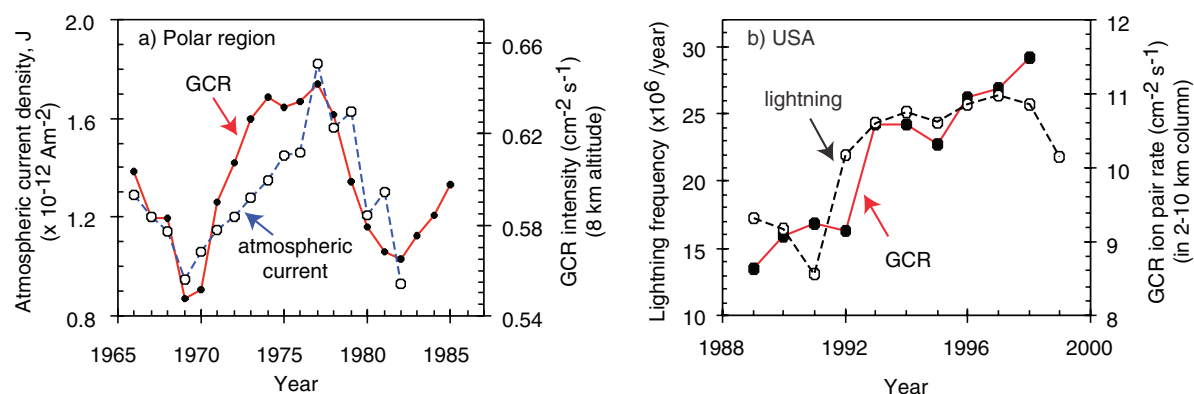


Fig. 40: a) Solar-cycle variation of the atmospheric current density, J , and the GCR intensity, both in the polar region, [107] and b) frequency of lightning recorded in the United States and change of GCR intensity for 1988–1999 [108].

4.6.2 GCR - cloud-electricity interactions

Global electrical circuit: By their effect on the ion pair concentration and, perhaps, on lightning frequency, GCRs can in principle affect the atmospheric conductivity, σ , the ionospheric potential, the atmospheric current density, J , and the atmospheric electric field, E_z . A study of the variation of atmospheric current density in the polar region over the period 1965–1985 (Fig. 40a) [107] shows evidence for a solar modulation. The increase of J around the minimum of the solar cycle is consistent with the increased conductivity of the air due to the higher GCR intensity. However, there is a second inference to be made: there must be a simultaneous increase of the current source, i.e. of lightning frequency, in order to sustain the higher net flow of fair-weather charge between the ionosphere and the ground. Some direct evidence that supports this is seen in Fig. 40b) [108]. This suggests that the efficiency of charge separation in thunder clouds may be influenced by the ionisation concentration from GCRs (a mechanism that is consistent with this picture is discussed below). However other studies of a possible relationship between solar activity and the frequency of thunderstorms or lightning have been made, and the results are not conclusive with some reporting a positive correlation, some a negative correlation and some none of any significance (see, for example ref. [106] and references therein). Therefore more experimental data are required to clarify the situation.

Since lightning is the dominant source of NO_x in the marine troposphere, if GCRs do indeed affect lightning frequency, it would imply they also modulate the production of this important reactive trace radical in the troposphere [109]. Note that this production mechanism for NO_x (and OH radicals) occurs both below thunderclouds (due to lightning) *and* above them (due to sprites and elves; § 4.6.1).

Rainfall: Stozhkov *et al.* have studied the influence of Forbush decreases of GCR intensity on precipitation in Brazil and in the former Soviet Union [55, 110]. The combined experimental data for around 70 Forbush events (Fig. 41a) indicate a decrease of rainfall of $(17 \pm 3)\%$ on the day of the Forbush decrease. These authors have also studied ground-level solar proton events (solar cosmic rays), where the cosmic ray intensity increases for a period of about a day due to high energy solar events (Fig. 41b). Here too they find some evidence for an effect, $(13 \pm 5)\%$ (53 combined events), and with a consistent sign, namely an *increase* of rainfall. In a separate study of Forbush events, Pudovkin and Veretenenko have reported [111] a correlated short-term decrease of cloudiness using data obtained visually in a narrow range of latitudes (60N–64N), which is consistent with the global satellite observations (§2).

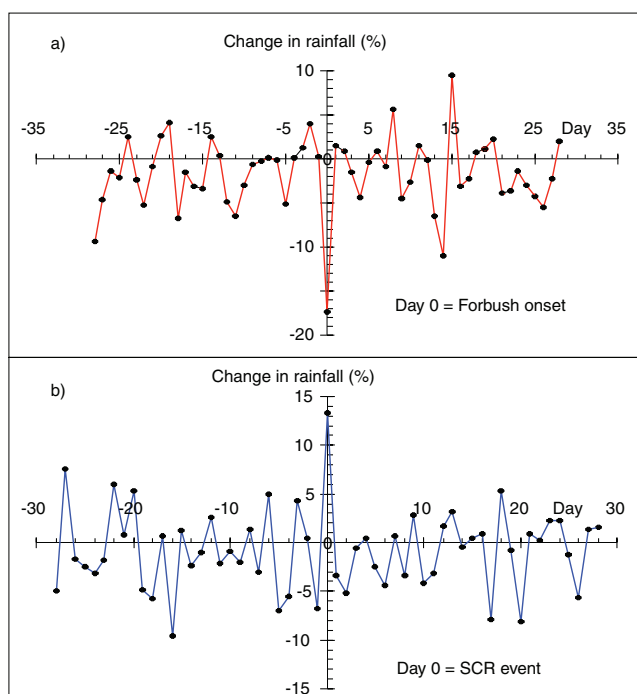


Fig. 41: Variation of rainfall recorded in the former Soviet Union and Brazil during short-term changes the cosmic ray intensity [55]. The relative change in rainfall in the days spanning a) *decreases* of the cosmic ray intensity due to Forbush events (70 combined events) and b) *increases* of the cosmic ray intensity due to ground-level solar cosmic ray events (53 combined events).

These data suggest that GCR ionisation may affect the precipitation efficiency of clouds. One candidate physical mechanism is an enhanced formation of ice particles (§4.5). Another is the increased efficiency of droplet growth by collision and coalescence, which is the dominant process by which cloud droplets grow in size from cloud droplets (about $20 \mu\text{m}$ diameter) into raindrops (about 1 mm). Cloud droplets become charged by the diffusion of small ions. Measurements indicate that the mean electronic charge on a non-thunderstorm cloud droplet is approximately represented by $q(e) = 4.1 r^2 (\mu\text{m}^2)$, where r is the droplet radius [102]. In thunderstorm conditions, the charge is about an order of magnitude larger, $q \sim 40 r^2$. The presence of charge is expected to produce large increases of coalescence efficiency due to image charge forces, which are always attractive at sufficiently close distances (see, for example, Fig. 42). This process has been termed “electroscavenging” [100, 101].

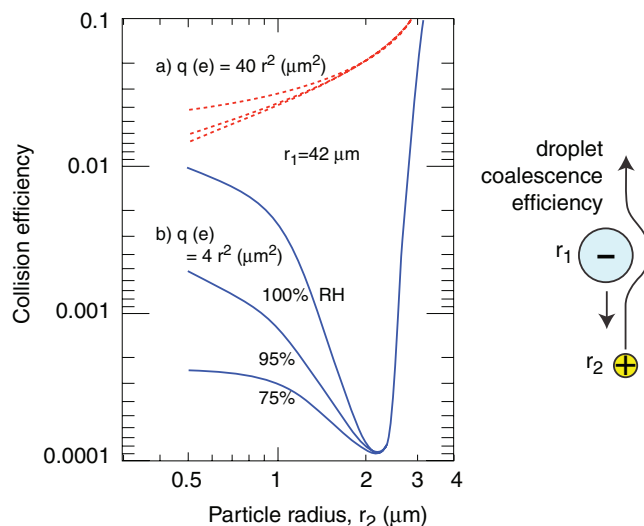


Fig. 42: The effect of charge on the collision ('sticking') efficiency for a falling large droplet (radius $42 \mu\text{m}$) and smaller droplets of various radii as indicated on the x axis [105]. The droplet charges correspond to a) thunderstorm clouds (dashed lines, and b) non-thunderstorm clouds.

Charge separation in clouds: Many attempts have been made to explain how charge is separated in electrically active clouds, but the actual physical mechanism remains unclear [105]. One hypothesis is that charge separation is due to a preferential activation of negatively charged CCN when a rising air parcel first exceeds the threshold supersaturation of water vapour [55]. A sign preference was first noticed by C.T.R. Wilson in 1899 in his original development of the cloud chamber, but this was for small ions and at very high supersaturations of several hundred per cent. The sign preference has been attributed to surface orientation of dipolar molecules, and a theoretical understanding has been proposed [112]. In the atmosphere, the ionisation created by GCRs is efficiently scavenged by aerosols, so many have a net charge of typically a few electronic units. If the sign preference noted above also occurs under natural conditions, then the first particles to activate into cloud droplets would be preferentially of negative sign. These negatively-charge droplets could then in principle grow to a sufficient size before the positively-charged CCN started to activate that the charges could become separated gravitationally. This would result in a region of net negative charge lying below a region of net positive charge, as observed in the lower region of typical thunder clouds.

Lightning trigger: As well as providing the initial source of charge, and perhaps also the mechanism for charge separation, GCRs may play a decisive role in triggering lightning. In this case, the trigger is probably a very high energy primary, near 10^{15} eV [113]. These occur at a rate of about $2 \text{ km}^{-2}\text{s}^{-1}$, and they are, of course, totally unaffected by solar modulation.

A typical lightning flash from cloud to ground is observed to occur as a series of leaders proceeding in rapid steps each of order 10 m length, with a pause of about $50 \mu\text{s}$ between each [105]. When the stepped leaders reach the ground, the main flash occurs as a return stroke of a few hundred amps current. The stepped leaders take about 1 ms to descend about 100 m whereas the return stroke takes about $1 \mu\text{s}$ (at one third the speed of light) to travel back along the full ionisation channel created by the leaders. The return stroke continues to bring negative current down from the cloud to the ground for between $10 \mu\text{s}$ and 10 ms, corresponding to peak currents of about 10 kA and an average charge of about 20 C.

A mechanism termed 'runaway breakdown' has been proposed to explain the initial streamer from which the stepped leaders propagate [113]. In the core of a very high energy GCR ($> 10^{15}$ eV), it is proposed that sufficient electron density exists for avalanche multiplication to occur in electric fields

of 100–200 kV m^{-1} , which are typical of thunderstorm clouds (Table 3) but well below the threshold fields for conventional damp air breakdown (1–2 MV m^{-1}). The avalanche polarizes the local plasma and thereby increases the electric field in a positive feedback. This raises the electron density to the point where a short streamer forms, which then triggers the first leader. A distinctive signature of the runaway breakdown mechanism is the appearance of X rays of energy around 50–100 keV starting with the first stepped leader, around 1 ms before the lightning discharge, which are caused by electron bremsstrahlung in the highly conductive plasma (long mean free path). Such X rays have recently been observed in lightning storms [114, 115].

Finally we comment that the association of an ultra high energy GCR shower with a lightning discharge may help explain the phenomena of ‘rain gushes’. The latter are the familiar dramatic increase in precipitation arriving at the ground shortly after nearby lightning [105]. It is possible that rain gushes result from a sudden increase of droplet coalescence efficiency due to a sharp rise of ionisation, as described above (Fig. 42). Various attempts have been made to explain how the ionisation generated by a lightning flash could rapidly diffuse laterally over a sufficiently large area, but none is convincing. However, since the transverse width of an ultra high energy GCR shower is about 100 m, then the combination of the initial ionisation, a triggered lightning flash, X ray emission and absorption, and some mild electrical activity over the full width of the GCR shower may together generate sufficient ionisation to cause the rain gush over a \sim 100 m-wide region.

5 THE CLOUD FACILITY

5.1 Concept

The basic concept of the CLOUD experiment (Figs. 43 and 44) is to investigate the microphysics of GCR-cloud interactions under laboratory conditions where all the experimental parameters can be precisely controlled and measured. A beam from a particle accelerator provides an artificial source of ‘cosmic rays’ that is adjustable and precisely known. The beam illuminates a 0.5 m expansion cloud chamber and a 2 m reactor chamber. The chambers can be operated at any temperature and pressure in the troposphere and stratosphere. The cloud chamber simulates the conditions inside clouds, throughout the atmosphere. The reactor chamber is important for experiments involving long beam exposures of a day or more, and provides the reacted gas/aerosol samples for analysis by the external instrumentation—mass spectrometers, particle sizers, etc. Data will be recorded under a wide variety of operating conditions, and the results compared at different beam intensities, including beam-off. In this way an unambiguous study of the effects of relativistic ionising particles on cloud microphysics can be carried out.

Such measurements are difficult to perform with cosmic rays in the atmosphere since the natural intensity variations are modest and—with the exception of relatively rare Forbush decreases—they follow the slow 11-year solar cycle. Furthermore, a laboratory experiment allows full control of the gas and aerosol mixture under test and, in addition, complete physical and chemical analysis of the products before, during and after beam exposure. This will greatly facilitate an understanding of the microphysics and chemistry of any effects that are observed. The challenges of a laboratory experiment are to duplicate the atmospheric conditions realistically and to ensure that the detector dimensions are sufficiently large that wall effects do not influence the measurements.

Surprisingly, cloud chambers have never been operated at atmospheric conditions in a particle beam. C.T.R. Wilson was inspired to develop the cloud chamber [74, 116] after observing meteorological phenomena on the mountain of Ben Nevis in 1894. He developed the cloud chamber to try to reproduce clouds in the laboratory. Although his expansion cloud chamber was crucial to the development of nuclear and particle physics in the first half of the 20th century, and earned him the 1927 Nobel Prize in Physics, Wilson remained fascinated by atmospheric phenomena throughout his life. Indeed he devoted the latter part of his research life to seeking a connection between cosmic rays and clouds, and so it is perhaps a fitting tribute that a cloud chamber be proposed for the present studies.

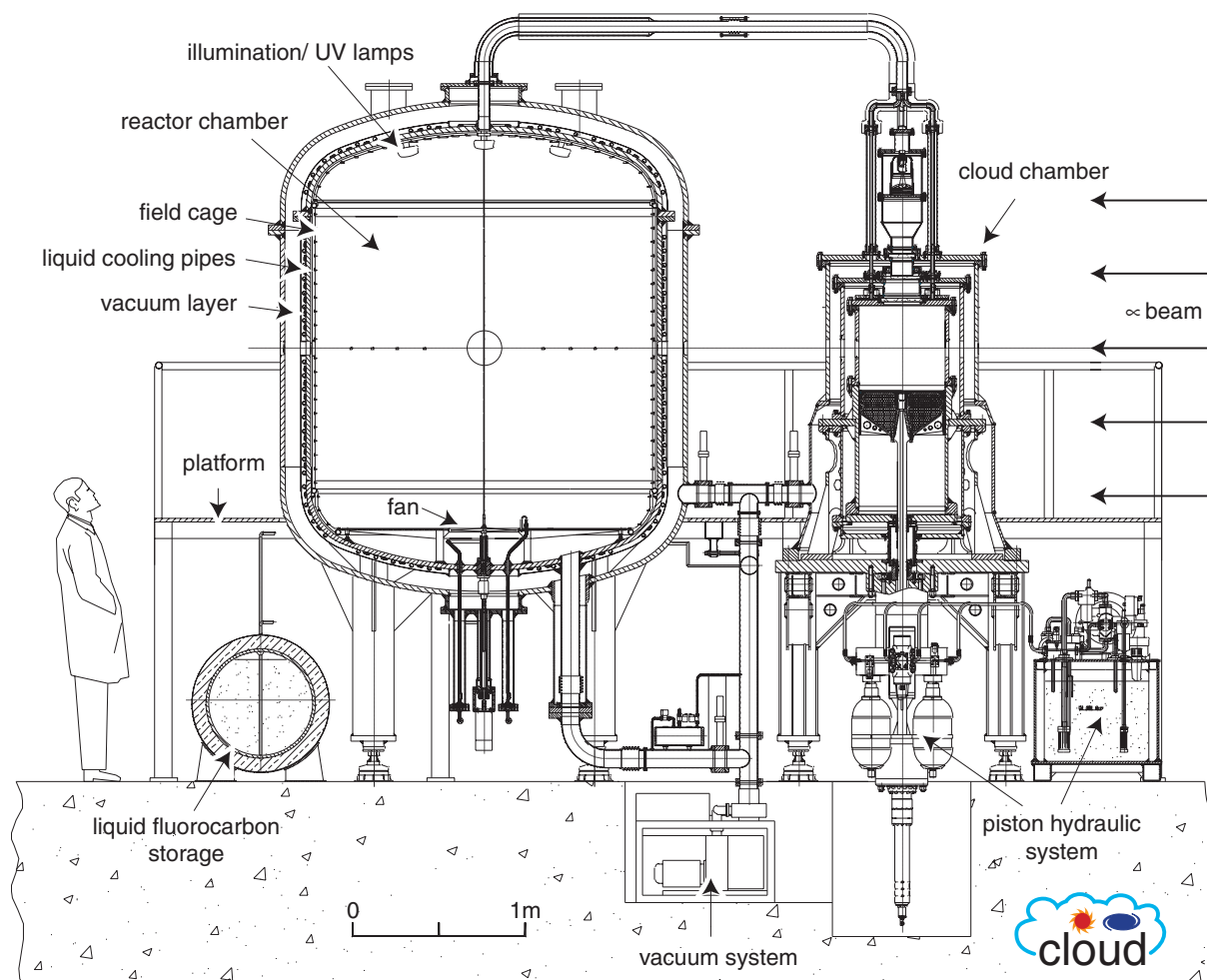


Fig. 43: Vertical section through the CLOUD facility showing the 0.5 m cloud chamber and 2 m reactor chamber. The beam counters are not shown.

5.2 Design considerations

5.2.1 Choice of expansion cloud chamber

There are several instruments capable of activating cloud condensation nuclei into droplets at the low water vapour supersaturations (few $\times 0.1\%$) found in clouds. Traditionally these are based on the diffusion cloud chamber or some variant thereof. All produce the necessary supersaturations over relatively small volumes and have inherent limitations in precision and in range of supersaturation (or, equivalently, in range of CCN size) [117]. Expansion cloud chambers, on the other hand, can in principle produce a precise and uniform supersaturation over a large volume and, moreover, maintain or dynamically adjust the supersaturation over long periods. They are therefore uniquely suited to the CLOUD experiments, many of which require long times for beam exposure, aerosol growth and droplet observation (§5.5.1). Expansion cloud chambers can also cover the full range of supersaturations between cloud conditions and those required ($\sim 500\%$) to activate small ions and embryonic aerosols in the nanometre size range. The phrase “in principle” is used above since the technical requirements are quite challenging (see below) and, to our knowledge, a cloud chamber with the performance proposed has not previously been built. However, the design requirements can be achieved with current technology, and extensive experience with cloud chambers has shown that they are high precision instruments. In particular, previous measurements by members of the CLOUD collaboration have demonstrated that the thermodynamic conditions after an

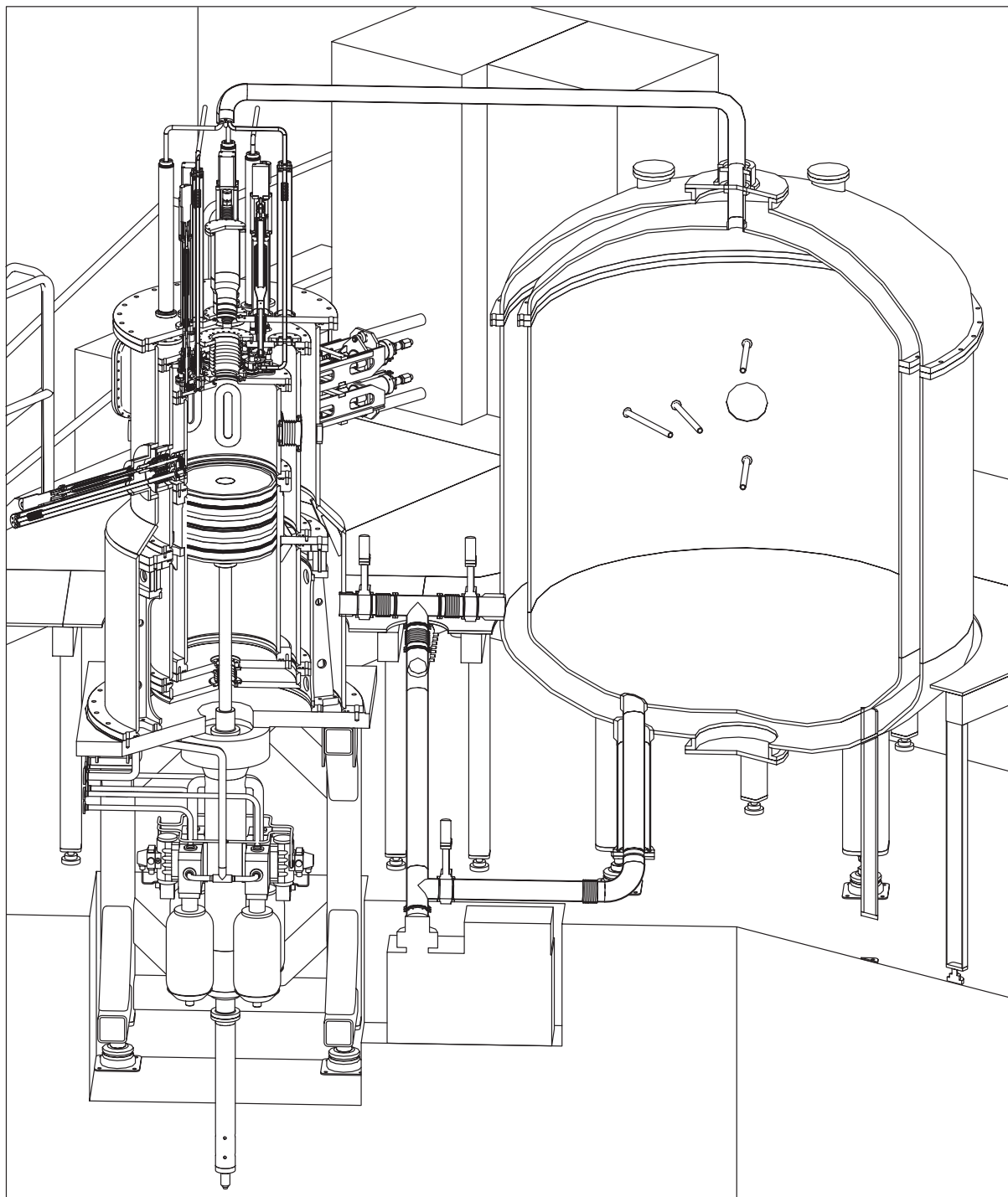


Fig. 44: Cut-away view of the CLOUD facility, showing the 0.5 m cloud chamber (left) and 2 m reactor chamber (right). The external instrumentation (mass spectrometers, ion mobility spectrometers, particle sizers, etc.) is not shown.

expansion are precisely known and reproducible provided the initial conditions are well-known and the expansion ratio (pressure change) is well-measured [118, 119].

5.2.2 Sensitive time

The requirements of long sensitive time, long droplet growth time and minimising gas/aerosol diffusion losses to the walls all argue for a cloud chamber of large dimensions. Combining these considerations with the requirements of flexible expansion control and rapid turn-round time between fills, the optimum cloud chamber size is ~ 50 cm linear dimension of the active volume. Before expansion, the walls and gas are in thermal equilibrium. Following an adiabatic expansion, the gas temperature is cooler than the walls, and the gas layer close to the walls begins to warm up, which re-compresses (and adiabatically warms) the inner gas volume. This sets the limit on the (static) *sensitive time* of the cloud chamber—the period during which no significant changes of the thermodynamic conditions occur in the central part of the chamber, where the measurements are made. For small expansion ratios, the sensitive time, $t_s \propto L^2/S^2$, where L is the linear dimension of the cloud chamber and S is the supersaturation after expansion [120]. Experience with the Vienna 25 cm cloud chamber shows that $t_s > 10$ s for $S = 0.02$ (2%). So the sensitive time of the 50 cm CLOUD chamber should be $t_s \sim 3$ min at 1% S , and ~ 1 hr at 0.2% S .

However, these are *static* sensitive times. In the CLOUD experiment there will be a *dynamical* adjustment of the piston displacement to compensate for the warming effect of the walls. In this way the sensitive time will be greatly extended. Moreover, by over-compensating for the warming effect (by means of a higher rate of slow expansion) a dynamical increase of supersaturation can be generated to simulate a selected adiabatic lapse rate characteristic of a rising air parcel.

These sensitive times may be compared with the droplet growth time, $t \propto r^2/S$, where r is the droplet radius. Measurements with the Vienna cloud chamber indicate a 1 s growth time to 1 μm radius at $S = 0.02$. This indicates growth times at $S = 0.002$ of about 10 s for droplets of 1 μm radius, and about 17 min for 10 μm radius. Therefore the sensitive time of the cloud chamber should allow measurements of droplet growth at low supersaturations up to sizes typically found in clouds. The limit is eventually set by sedimentation. The terminal velocity of a water droplet at stp, v_o ($\mu\text{m s}^{-1}$) = $130 r^2$ (μm^2). So, for example, a droplet of 10 μm diameter has a terminal velocity of 3 mm s^{-1} .

5.2.3 Losses of ions, trace gases and aerosols to the walls

The finite size of the cloud chamber will result in the loss of charged particles, trace gas molecules and aerosols to the chamber walls by diffusion. When an aerosol touches a wall, it attaches by van der Waals forces and is lost. For the purposes of the Monte Carlo study shown in Fig. 45 we have also assumed that a trace gas molecule or small ion is lost (or, equivalently, the ion is neutralised) if it collides with a wall. The results of the simulation indicate relatively small losses for aerosols, due to their low mobility, but significant diffusion of ions and trace gas molecules to the walls. The small ions are readily replaced by the ionising beam (§5.4), but losses of trace vapours due to wall adhesion or to aerosol growth must be compensated by a small inflow of makeup gas during long-exposure experiments.

The experience of the AIDA aerosol facility [121, 122] at Forschungszentrum Karlsruhe is instructive concerning the wall losses of trace gases. AIDA is a large cylindrical vessel of internal dimensions 4 m (diameter) \times 7 m (height) with an inner ceramic lining. The temperature is controlled to 0.5 K precision by cold air circulating around the outside of the vessel. It is equipped with an internal fan to ensure homogeneous mixing. Operation of the fan is found to have only a small effect on aerosol lifetimes. The measured $1/e$ lifetimes in AIDA for trace O_3 and NO_2 gases at room temperature are 70 h and 370 h, respectively. Since the lifetimes should scale as the linear dimension of the vessel (i.e. the ratio of its volume to surface area), this implies the 0.5 m CLOUD chamber should have corresponding lifetimes of about 7 h and 40 h, respectively (and four times longer for the 2 m CLOUD reactor chamber; §5.3.1).

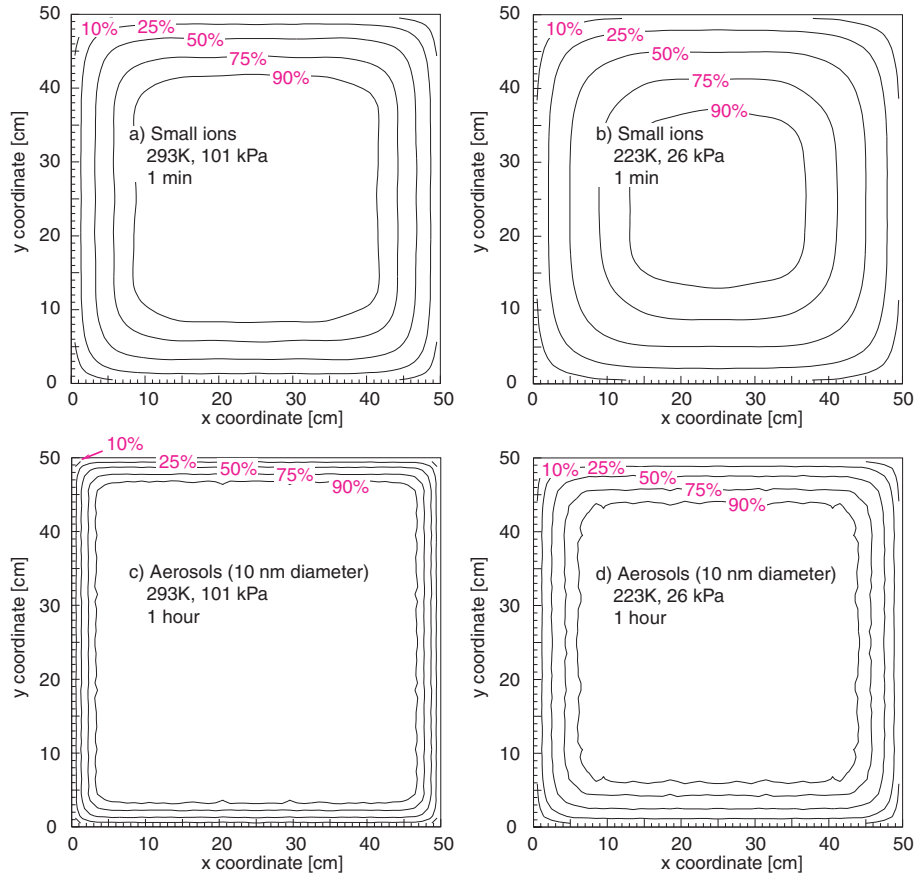


Fig. 45: Wall losses of particles in the cloud chamber due to thermal diffusion. The upper plots show the number density of small ions after a time $t = 1$ minute at a) 293 K and 101 kPa (standard conditions) and b) 223 K and 26 kPa (10 km altitude). The lower plots show the number density of 10 nm-diameter aerosols after a time of 1 hour at c) 293 K and 101 kPa and d) 223 K and 26 kPa. A particle is assumed to be lost if it touches one of the walls (which are located at the boundaries of the plots). The initial charged particle distributions were generated uniformly in x and y in the range $0 < x, y < 50$ cm. The contours indicate the fraction of the original number density of particles remaining after the indicated times. The losses can be estimated at other times by scaling the distance between a contour and its nearby wall as \sqrt{t} .

5.2.4 Water vapour supersaturation

The cloud chamber is required to operate over a wide range of water vapour supersaturations, S , from below zero (unsaturated) up to about 700%. Of particular importance is the need to provide precise simulation of cloud conditions, where the supersaturation ranges from below 0.1% up to about 1%. This corresponds to a broad activation range of aerosols, namely radii from about $1 \mu\text{m}$ (at $S \sim 0.1\%$) down to about 50 nm ($S \sim 1\%$). In order to probe the CCN distribution with sufficient resolution, the cloud chamber needs to achieve a S precision after expansion of better than 0.1% in the range $0 < S < 1\%$. In most experiments, the initial relative humidity will be established at 100% by a film of water on the upper surface of the expansion piston.

The highest supersaturations will allow activation and measurement of ultrafine condensation nuclei at sizes corresponding to small ions (~ 0.2 nm). To measure individual small ions, C.T.R. Wilson operated his expansion chambers at water vapour supersaturations between about 550% and 700%, corresponding to volume expansion ratios between 1.31 and 1.36 [74]. At an expansion ratio of 1.37 he found a dense cloud appeared, which presumably corresponds to activation of uncharged transient molecular clusters.

5.2.5 Piston displacement precision and expansion time

The required piston displacement ranges from about 10 μm ($S \sim 0.01\%$) up to 180 mm ($S \sim 700\%$). The two schemes under study for the piston actuator, namely a linear motor [123] and a hydraulic/servo-valve system (§5.3.1), can both achieve extremely fine control of the piston position. The linear motor system can reach sub- μm precision, with sufficiently precise position sensors. Experience with the hydraulic system of the 2 m Big European Bubble Chamber (BEBC) at CERN showed that its piston displacement could be reproduced to a precision of about 5 μm [124]. With these performance figures, the cloud chamber volume expansion ratio can be set precisely and extremely fine control of the supersaturation can be achieved.

The piston actuator is designed to provide a flexible choice of expansion and re-compression cycles, with piston movements reproducible to very high precision. For example, a single rapid expansion is required in experiments where the expansion pulse needs to be short compared with the droplet growth time, so that the start time is the same for all activated aerosols, independent of their size. For a large, 30%, volume expansion in CLOUD, the fastest expansion time is below 200 ms. The expansion time is less for smaller volume expansions. For comparison, the expansion time for the 25 cm Vienna chamber is 20-30 ms and, for the 38 cm Missouri-Rolla chamber, it is 200 ms.

However, other experiments require different expansion/compression cycles. One useful technique is to use a brief expansion pulse to nucleate particles for a known, short time, followed by a mild re-compression to stop nucleation but allow droplet growth to continue. Another example is to execute a continuous slow expansion to match the adiabatic lapse rate characteristic of natural cloud development, as mentioned above. This will allow studies of the dynamics of the activation of different CCN, including, for example, the effects of depletion of water vapour by the early growth of the larger CCN, which prevents activation of the smaller CCN. A continuous, progressive expansion will also allow measurement of the CCN spectrum as a function of supersaturation. During long-growth experiments such as these, the piston cycle will be adjusted to compensate for wall heating. Downdrafts (decreases of supersaturation) and evaporation/condensation cycles can also be generated by more complex expansion/compression cycles. These will be important for studies of the effects of cloud processing and ice nucleus processing.

5.2.6 Temperature and pressure range and stability

The cloud chamber is required to operate over the full range of temperatures and pressures encountered by clouds in the troposphere and stratosphere, namely $185 \text{ K} < T < 315 \text{ K}$ and $0 < P < 101 \text{ kPa}$. The maximum pressure of the chamber is actually about 150 kPa, in order to allow measurements to be made at 1 atm following a large expansion.

The requirement of a precision of better than 0.1% absolute on the water vapour supersaturation places demanding requirements on the temperature stability and uniformity of active volume of the cloud chamber. Taking a design value for the supersaturation error of one half this value, i.e. 0.05%, implies the need for a temperature stability $\Delta T = 0.01 \text{ K}$ and a pressure stability $\Delta P/P = 1.3 \cdot 10^{-4}$ or, equivalently, a volume stability $\Delta V/V = 0.9 \cdot 10^{-4}$. These stabilities and uniformities inside the active volume are required during droplet activation so that the thermodynamic conditions are well known. At other times, when the chamber is not in an active expansion cycle, the stability requirements are more relaxed. Note also that these are *stability* requirements; the absolute precisions of the temperature and pressure are less demanding.

5.2.7 Field cage

The cloud chamber and reactor chamber are equipped with a field cage to provide an electric field within the active volumes. The field cage is important for several reasons. In its simplest application, a modest electric field of about 1 kV m^{-1} will clear small ions from the cloud chamber in about 2 s. This is well

below typical charge attachment times onto aerosols and will allow control measurements to be made at effectively zero ionisation (below 1% of ground-level ionisation). A second application is to select the *sign* of the ions or charged aerosols before droplet activation, in order to allow investigation of sign-dependent effects. This can be achieved by setting the electrode voltages to produce an electric potential ‘valley’ for the selected sign at the mid-point of the chamber (and a ‘ridge’ for the opposite sign). The practical field will more resemble an inverted saddle (or saddle) and so there will still be some loss of the selected sign towards the walls at the mid-point of the chamber. However, a strong sign-selection can be achieved. A third application of the field cage is to establish realistic electric fields for the cloud electricity experiments (§4.6). The latter set the desired maximum electric field to about 10 kVm^{-1} , which covers all cloud conditions except lightning (Table 3).

Finally the field cage may provide useful information on the *magnitude* of the droplet charge *in situ*. This involves a Millikan-type measurement of the terminal velocity of a droplet vs. electric field. It is limited to droplets that are visible by the CCD cameras (i.e. above about $1 \mu\text{m}$ diameter) and have sufficient charge to be suspended, or almost suspended. For example, in an electric field of 10 kVm^{-1} , the charge required to suspend a droplet is $q_s(e) = 24r^3 (\mu\text{m}^3)$, where r is the droplet radius. This indicates minimum charges of $3e$ are required to suspend a $1 \mu\text{m}$ diameter particle and $24e$ for a $2 \mu\text{m}$ diameter particle. In dipolar small-ion clouds, the typical droplet charges from diffusion will be a few electronic charges, so this technique appears to be useful only for particles in a narrow diameter range of a few- μm 's. It may also be possible to devise useful ways to estimate aerosol mobility (charge/size) *in situ* by drifting the aerosols for a known time before droplet activation.

The electric field is provided by a field cage attached to the walls of the cloud and reactor chambers. A similar design has been successfully used in the large TPC tracking detector of ALEPH at LEP. The ALEPH TPC involves a (half) cylinder of 2.2 m length and 3.6 m diameter, with an electric field of 11.5 kVm^{-1} that is provided by field cage electrodes on the surface of the cylinder, with none placed in the active (half) volume.

5.2.8 Chamber cleaning

CLOUD experiments will frequently involve extremely low concentrations of trace vapours (around 1 pptv, equivalent to concentrations of 10^7 molecules cm^{-3} or less) and aerosols ($\ll 1 \text{ cm}^{-3}$). Therefore trace impurities must be eliminated at well below this level. Fortunately, provided suitable care is taken, the cloud chamber is *self cleaning* [119]. By repeated and progressively deeper expansions, impurities are activated and can be sedimented out of the active volume (the same can be achieved in the reactor chamber by vacuum pumping). Impurity concentrations of <0.01 condensation nuclei cm^{-3} can be reached, where a condensation nucleus refers to a molecular cluster of *any* size. However, to reach this performance, the chamber must be carefully cleaned between each experiment. In most cases the cleaning of aerosols and gas molecules attached to the walls can be carried out by vacuum baking. This involves first emptying the liquid fluorocarbon coolant and then vacuum pumping of the active volume while heating the chamber by means of heater cables wrapped around the inner chamber body. Periodically, it will be necessary to flush and clean the inside of the chamber with liquid solvents. This can be achieved in two ways. The first makes use of a central hole through the piston drive rod and piston. This can be used both to flush liquids into the chamber and also to drain them. The second involves a more thorough cleaning (or repair) by disassembly of the cloud chamber body. The latter is designed to be able to be warmed up, disassembled, repaired, reassembled and cooled down again in 24 hours. During such an operation, the lower cloud chamber assembly, including the piston and actuator system, is left intact.

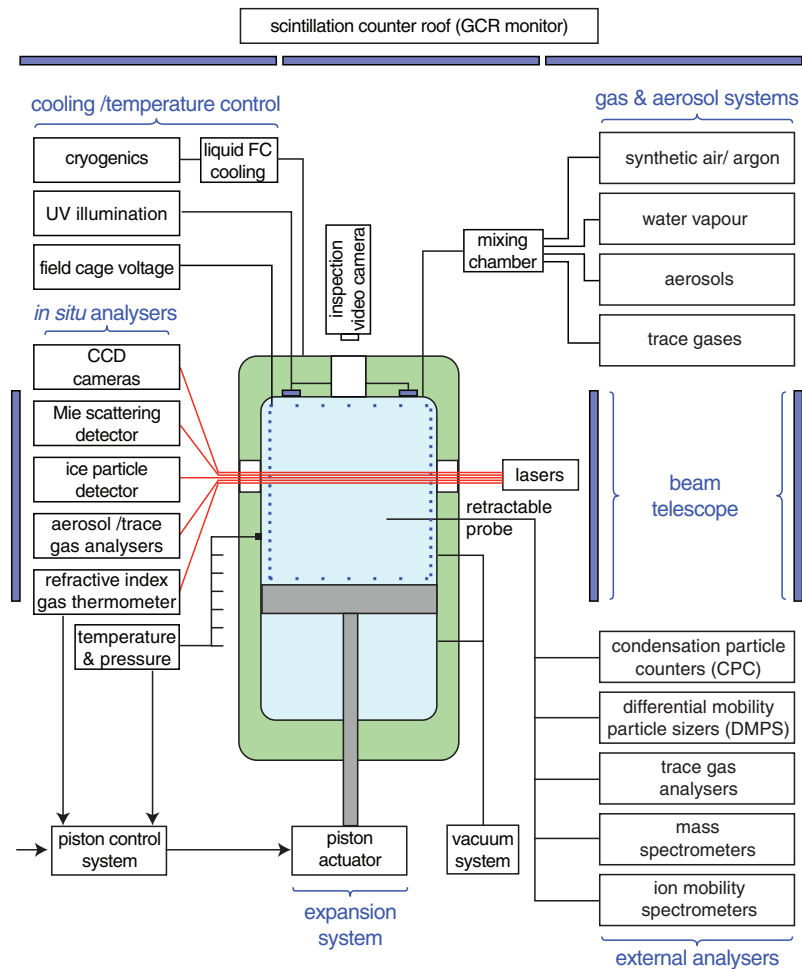


Fig. 46: Schematic diagram of the subsystems and instrumentation for CLOUD. Normally the gas/aerosol samples for the external analysers are drawn from the reactor chamber rather than the cloud chamber as indicated the figure.

5.3 Experimental design

5.3.1 Cloud chamber and reactor chamber

The CLOUD detector is shown in Figs. 43 and 44, and the subsystems and instrumentation are shown schematically in Fig. 46. The main components are a 0.5 m cloud chamber and a 2 m reactor chamber. The purpose of the reactor chamber is twofold: i) it provides the samples of reacted gas/aerosols for analysis by the external detectors, and ii) it is required for long growth-time experiments lasting several days, since it has about a factor four longer particle lifetimes than the cloud chamber, due to reduced wall losses. For these experiments the reactor chamber can also provide samples of reacted gas/aerosols for analysis in the cloud chamber since it has a much larger volume (by a factor 64). The chambers are exposed to a muon beam of large transverse dimensions (about $2 \times 2 \text{ m}^2$) which provides a simultaneous equal ionisation of both volumes.

The chambers are constructed from aluminium, with black teflon lining the inner surfaces. The teflon is made partly conductive in order to prevent any charge accumulation. Both chambers can be operated at pressures between a vacuum and 1.5 atm. Precision pressure gauges and Pt thermometers are set into the chamber walls. A teflon-coated field cage is suspended inside each chamber, isolated from the walls by standoff insulators. The field cages generate vertical electric fields inside the chambers of up to 10 kVm^{-1} with a flexible choice of field profiles.

The piston expansion system for the cloud chamber comprises the main piston, drive rod and

actuator. Two options are under study for the actuator system. One option is a hydraulic system using a servo valve, as shown in Figs. 43 and 44 and described in the CLOUD proposal [3]. It is based on the same design as was used to control the 2 m-diameter piston of BEBC [124]. The second option is a linear electric servo motor with rare-earth permanent magnet arrays attached to the drive rod and liquid-cooled iron-core electromagnetic coils fixed to the support structure [123]. Both options provide very precise displacement of the piston, and flexible electronic control of any desired expansion/compression cycle via digital signal processors. The most rapid expansions can be made under 200 ms for a 35% volume expansion. Precision sensors measuring temperature, pressure and piston displacement provide feedback to control the piston actuator. The piston is a stiff and lightweight sandwich assembly. A critical component is the piston seal. The present design foresees a *Bellofram* rolling diaphragm [125], with a permanent vacuum below the piston [123]. A bias piston at the base of the drive rod coupled to a large pressurised ballast tank compensates most of the excess downward force on the piston (2.0 tonnes at 1 atm). A small depression in the top surface of the piston allows for a thin pool of water (or ice) to establish 100% relative humidity.

The liquid cooling and temperature control system involves a closed circuit system, insulated throughout by vacuum. The liquid fluorocarbon coolant flows through a jacket surrounding the cloud chamber and reactor chamber and maintains the inner walls and piston at a precisely-controlled temperature. The gas inside the cloud chamber is allowed to reach thermal equilibrium with the walls before taking any measurements. Heater cables are also wrapped around the chamber vessels to allow cleaning by vacuum bakeout. The reactor chamber is operated at the same temperature and pressure as the cloud chamber. However, since no droplet activation is involved, a relatively modest temperature stability of a few $\times 0.1$ K is adequate.

The gas and aerosol supply systems involve four components: carrier gas, water vapour, aerosols and trace gases. The carrier gas is either pure artificial air (80% N₂, 20% O₂) or argon. Water vapour, aerosols and trace gases are mixed into this stream at the desired levels. The water vapour content in the chamber will be set by two techniques: either a) a liquid or ice film on the top of the piston or b) vapour introduced from an external humidifier. Aerosol particles will be generated with standard techniques such as nebulizers. Care must be taken to minimise transmission losses between the aerosol generators and the chamber volumes (and between the sampling probes and the external analysers). In general, sub-micron particles are quite efficiently transported in small tubes at standard carrier-gas flow rates. The main mechanisms for transport losses are diffusion onto the tube walls, gravitational settling, and inertial losses at sharp bends. Charged particles have additional losses due to electrostatic attraction. Diffusion losses are only significant for particles smaller than approximately 20 nm. Gravitational losses are usually only significant for particles exceeding about 0.5 μm diameter. All these losses, and probe sampling efficiencies, are well known [126] and can be accounted for.

Both chambers are equipped with gas/aerosol inlet and outlet pipes, as well as ports connecting to the vacuum system and connecting each chamber to the other. The inner vessels are fitted with sampling probes to extract gas and aerosols for the external analysers. The tip of each sampling probe can be independently adjusted to any radial position inside the vessel (these are visible inside the reactor chamber in Fig. 44). A small fan is installed inside the reactor vessel to provide the option of slow stirring of the gas filling to assist, where necessary, homogeneous mixing throughout the large volume. A special inlet pipe located near the fan provides fresh trace gas to replace losses to the walls or to aerosol growth. As well as windows for optical readout, the chambers are equipped with inspection and illumination windows. Internal UV lamps are also provided for experiments involving gas-phase photochemistry.

5.3.2 Instrumentation

The CLOUD instrumentation is shown schematically in Fig. 46. The *in situ* analysis of the cloud chamber comprises the following systems: a) a constant angle Mie scattering (CAMS) detector [127], b) a stereo pair of CCD cameras, c) an ice particle detector which measures backscattering of a polarised laser beam

[122], d) gas and aerosol analysers, e) a refractive index gas thermometer based on a laser interferometer [128], and f) precision temperature and pressure monitors set into the inner wall.

During activation, the cloud droplets are primarily analysed by the CAMS and CCD systems. These are complementary but, nevertheless, have a broad region of overlap where they can provide mutual cross-checks. The CAMS system can measure very high droplet number densities ($\sim 10\text{--}10^7\text{ cm}^{-3}$) whereas the CCD cameras operate best in a lower range ($\sim 0.1\text{--}10^5\text{ cm}^{-3}$). The CAMS system provides a high-resolution measurement of mean droplet radii vs. time, whereas the CCD cameras provide a measurement of droplet size in coarser time intervals, using pulse height information. Finally, the CAMS detector integrates over all illuminated droplets whereas the CCD cameras reconstruct the 3-dimensional spatial positions of individual droplets, and track their movements. This is important for identifying ice nuclei and for measuring droplet drift and sedimentation trajectories.

The illumination system for the CAMS and CCD systems comprises: a) a laser for illumination of a narrow region for the CAMS detector (and, in parallel, for the CCD cameras) and b) a xenon flash tube mounted at the top window, for the CCD cameras. A video camera is also mounted at the top window to provide a visual inspection of the chamber volume and piston surface. All windows are made of optical quality quartz (since it is UV-transparent and has a high thermal conductivity, similar to stainless steel).

As well as *in situ* analysis, samples of gas are drawn from the chambers via retractable probes and directed to an array of external instruments to analyse the chemical and physical characteristics of the aerosols, trace gases and ions. The instruments include condensation particle counters, differential mobility particle sizers, trace gas analysers, mass spectrometers and ion mobility spectrometers. The beam intensity and profile is measured by a plastic scintillation counter telescope. The counters cover the full 2 m width of the beam and include a fine-grained array of 8×8 counters of size $25 \times 25\text{ cm}^2$ to measure the transverse profile. Finally, a roof of plastic scintillation counters monitors the GCR exposure, which will be significant for measurements taken at the lowest beam exposures. More details of the instrumentation for CLOUD can be found in ref. [3].

5.4 Particle beam

5.4.1 Why a particle beam?

The basic experimental requirement of CLOUD is to duplicate atmospheric and cosmic ray conditions in the laboratory. Essentially throughout the troposphere, charged cosmic rays are minimum ionising particles (§4.3). These are mostly relativistic protons in the upper troposphere (above about 7 km altitude), and muons in the lower troposphere (Fig. 26).

The requirements of the ionisation source for CLOUD are as follows:

- Deposition of a precisely known quantity of ionisation within the cloud and reactor chambers.
- Uniform ionisation over a large volume (about $2 \times 2 \times 2\text{ m}^3$).
- An ionisation density (dE/dx) that is characteristic of minimum ionising particles.
- Easily adjustable in intensity over the required range of $1\text{--}10\times$ the natural cosmic ray intensities found in the troposphere—an intensity range of about 1000.
- Ability to traverse the walls and liquid cooling layers of the cloud chamber and reactor chamber. This sets a *minimum* energy for a particle beam of about 1 GeV/c, taking multiple Coulomb scattering also into account.
- Known timing. This is necessary for the study of fast processes and also for ice nucleation studies to distinguish between deposition nucleation and freezing nucleation.

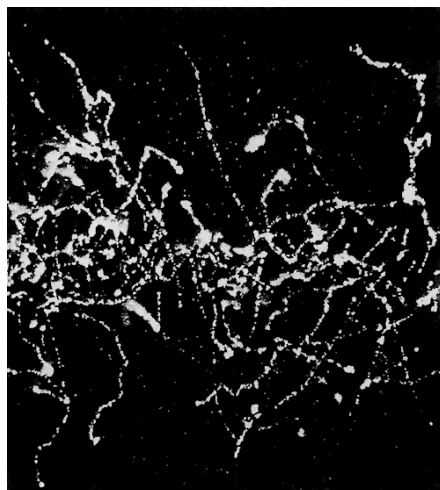


Fig. 47: An image of X ray interactions in a cloud chamber recorded by C.T.R. Wilson in 1912 [74]. The beam travels horizontally and has a diameter of about 2mm. The horizontal field of view is about 14 mm. Corrected to one atmosphere pressure, Wilson counted about 200 ion pairs per cm for the less highly ionising regions of these tracks, and over 2000 ion pairs per cm in the highly ionising regions near the ends of the tracks.

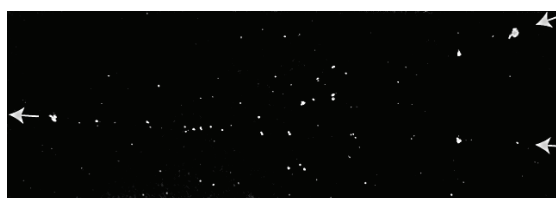


Fig. 48: An image of minimum ionising particles recorded by C.T.R. Wilson in 1912 with the same cloud chamber and operating conditions as Fig. 47 [74]. Two crossing tracks are seen, indicated by the arrows. The horizontal field of view is about 18 mm (the same scale as Fig. 47). Corrected to one atmosphere pressure, Wilson counted about 32 ion pairs per cm for these tracks, including the δ rays (the tightly packed groups appearing as bright spots).

Other sources of ionising radiation include ultraviolet (UV) radiation, radioactive sources and X ray sources. UV radiation is excluded since it induces photochemical reactions among the trace gases. In the case of radioactive sources, α emitters are excluded by their high ionisation density and short range, and β emitters are excluded since they cannot be placed inside the cloud chamber, and their range is insufficient to penetrate the chamber walls. Gamma radioactive sources and energetic X rays can penetrate the chamber walls but their ionisation deposition is dominated by stopping electrons, which have much higher dE/dx than minimum ionising particles. This can be seen visually by comparing Figs. 47 and 48, which show some remarkable cloud chamber images of X rays and minimum ionising particles, respectively, recorded by C.T.R. Wilson in 1912. In summary, a particle accelerator beam is the only source of ionising radiation capable of realistically duplicating cosmic rays in the laboratory.

In fact some little-known laboratory studies of ion-induced effects on aerosol formation have been carried out since the 1960's using traditional ionisation sources (e.g. refs. [97, 98, 129]). Members of the CLOUD collaboration have also studied these processes more recently using X rays and α particles from ^{241}Am sources [130]. Although some useful results have been obtained, these studies have generally been unable to characterise the aerosol processes adequately. The main limitations have been a) lack of control of the ionisation and ionisation rate (dE/dx) at near-atmospheric intensities, and b) non-uniformities of deposited ionisation. It has been shown [131] that local variations of ion density (such as from an α source) give rise to non-linear aerosol charging effects, which will directly affect ion-induced aerosol processes. This makes it difficult to relate results obtained with such sources to the real atmosphere.

Table 4: Summary of the approximate minimum and maximum time-averaged beam intensities for CLOUD. The GCR intensity at ground level is about $0.02 \text{ cm}^{-2}\text{s}^{-1}$, which is included in column 3 of the table. A transverse beam size of $200 \times 200 \text{ cm}^2$ is assumed. I_{GCR} signifies the natural GCR intensity at the indicated altitude.

Beam intensity		Beam + GCR intensity ($\text{cm}^{-2}\text{s}^{-1}$)	Clearing field	Simulated GCR conditions	
(s^{-1})	($\text{cm}^{-2}\text{s}^{-1}$)			intensity	altitude (km)
0	0	0.02	on	$0.01 \times I_{GCR}$	0
0	0	0.02	off	$1 \times I_{GCR}$	0
10^3	0.02	0.04	"	$2 \times I_{GCR}$	0
10^4	0.2	0.22	"	$10 \times I_{GCR}$	0
0	0	0.02	"	$0.01 \times I_{GCR}$	20
10^5	2	2	"	$1 \times I_{GCR}$	20
10^6	20	20	"	$10 \times I_{GCR}$	20

5.4.2 Beam requirements

The beam requirements for CLOUD are as follows:

- A muon beam; the ideal particles are μ 's since they replicate the GCR's in the lower troposphere, and do not interact in the detector material, as would π 's or p's (e's are excluded due to showering).
- A mean energy of several GeV, to avoid large multiple Coulomb scattering. The energy spread should be known (to calculate the ionisation energy deposition, dE/dx), but it can be broad.
- An adjustable, time-averaged intensity from about 10^3 s^{-1} to 10^6 s^{-1} (and also zero beam).
- Precise (few per cent) knowledge of the beam intensity (since the solar modulation corresponds to a rather small change of GCR intensity of only $\sim 10\%$).
- A large transverse size, $200 \times 200 \text{ cm}^2$.
- Continual operation throughout the year. A typical single experimental study will require about 4 weeks beam-time, and many experiments are foreseen (§5.5).

The desired beam intensity is between about $1\times$ and $10\times$ the natural GCR flux at any altitude. This will allow measurements of the dependence of any observed effects on ionisation rate and ion pair concentration at the natural ionisation rates in the atmosphere. The highest beam intensities will help to amplify and expose effects before they are measured at natural ionisation levels. Beam-off data will be also be recorded under conditions with the chamber clearing fields on and off, respectively, corresponding to $0.01\times$ and $1\times$ the natural GCR ion-pair concentrations at ground level.

CLOUD will measure processes over the full range of tropospheric and stratospheric conditions. At ground level, the average GCR intensity is about $0.02 \text{ cm}^{-2}\text{s}^{-1}$, whereas at 15–20 km altitude it is about a factor 100 larger, varying between about 0.8 and $2.3 \text{ cm}^{-2}\text{s}^{-1}$ depending on geomagnetic latitude (Fig. 27a). The maximum required *time-averaged* beam intensity is therefore about $10 \times 2 = 20 \text{ cm}^{-2}\text{s}^{-1}$. The beam is spread over a large transverse area of $200 \times 200 \text{ cm}^2$ in order to duplicate the quasi-uniform GCR irradiation, over the fiducial volume. The time-averaged maximum beam intensity is then $20 \times 4 \cdot 10^4 \simeq 10^6 \text{ s}^{-1}$. The *minimum* beam intensity (apart from beam-off) is about $1\times$ the natural GCR radiation at ground level. This is a factor 1000 below the maximum required intensity (a factor 100 for the atmospheric attenuation and a factor 10 for $1\times$ the GCR intensity rather than $10\times$), i.e. a time-averaged intensity of $\simeq 10^3 \text{ s}^{-1}$. These beam estimates are summarised in Table 4.

5.5 Experimental programme overview

5.5.1 Aerosol experiments

Gas-to-particle conversion: The aim is to measure the effect of ionising particle radiation on the formation rate of ultrafine condensation nuclei (UCN) in the few-nm size range from trace precursor vapours. Such trace gases include, in particular, H_2SO_4 , HNO_3 , NH_3 and certain volatile organic compounds, all in the presence of H_2O vapour. The basic parameter to be measured is the UCN nucleation rate, J ($\text{cm}^{-3}\text{s}^{-1}$), as a function of the primary experimental variables: trace vapour concentration, relative humidity, temperature, background aerosol concentration and ion-pair production rate. The cloud chamber can measure J over a wide range from about $3 \cdot 10^{-5} \text{ cm}^{-3}\text{s}^{-1}$ to $10^7 \text{ cm}^{-3}\text{s}^{-1}$. Theoretical studies [84, 85] indicate that large differences are expected in the ion-induced nucleation rate between positive and negative charges, by factors of up to 100 or more. These effects will be studied either with a ‘saddle’ electric field. Typical formation times (i.e. beam exposures) for these experiments of up to about one hour are expected.

Growth of CN into CCN: The purpose of these experiments is to measure the effect of ionising particle radiation on the rate of growth of CN into CCN (i.e. from ~ 5 nm diameter to ~ 100 nm) in the presence of condensable vapours that are known to be important in the atmosphere, such as sulphuric acid, water, ammonia and organic compounds. Both ‘dry’ and aqueous-phase growth will be studied. The latter will involve expansion and compression cycles of the cloud chamber to activate droplets and then evaporate them. The basic parameter to be measured is the CCN concentration, $CCN(S)$. A progressive expansion will be used to provide a gradually-increasing supersaturation so that the CCN concentration can be measured as a function of S . The CCN are the subset of the CN population that activate into droplets at a given water vapour supersaturation, S . The fraction of the CN that constitute CCN depends not only on their size but also on their chemical composition and other properties. The evolution of the CN size spectra during beam exposure will also be measured. A typical single run will last up to about 2 days.

Activation of CCN into cloud droplets: The goals of these experiments are to see how the presence of electrical charge affects the critical supersaturations required to activate CCN, and also how it affects the number of cloud droplets that appear. Well-defined aerosol sizes will be produced with standard aerosol generation techniques. Experiments will be performed to examine the activation of these aerosols into cloud droplets. The charge distribution on the aerosols will be measured *in situ* using the field cage, and externally with the ion mobility spectrometer. A wide range of CCN and ambient conditions (background aerosol, trace gases and thermodynamic conditions) will be explored. The simplest systems will involve typical hygroscopic aerosols found in the atmosphere, e.g. NaCl , $(\text{NH}_4)_2\text{SO}_4$ and H_2SO_4 , and also partially soluble aerosols such as CaSO_4 . Other measurements will be made on hydrophobic carbonaceous particles, with and without the presence of organic surfactants. The aims here are to investigate the effect of ionisation on the reduction of the critical supersaturations and on the increases of droplet number associated with the effects of surface charges [132] and surface tension suppression [54]. Finally, these measurements will be repeated in the presence of additional, highly-soluble, trace gases, which are expected to have an extremely low critical supersaturation, even to the point of activating into cloud droplets at below 100% relative humidity [133].

Dynamics of CCN activation and kinetic limitations on cloud droplets: The purpose of these experiments is to study the number of droplets that activate from a given CCN population under various dynamical supersaturation conditions. The calculation of the number of cloud droplets that will activate from a given CCN concentration is generally calculated under the assumption the classical Köhler theory (§4.4.1). However this assumes an equilibrium situation whereas, in real clouds, aerosol growth before activation may be kinetically limited [134]. The result can be large errors in the calculated number of

cloud droplets. Since the radiative forcing of clouds is very sensitive to droplet number (§4.4.2), it is important to be able to determine the concentration of cloud droplets to a few per cent. These experiments will first involve the preparation of a known, polydisperse, CCN population, with various CCN concentrations and compositions. The number of droplets that activate from these different CCN populations will then be investigated in the cloud chamber under different realistic adiabatic lapse rates, simulated by the appropriate progressive expansion cycles, both with and without beam ionisation.

Production of NO and OH and their influence on aerosols: These experiments will first quantify the poorly-known production rates of a) nitric oxide (NO) and b) hydroxyl radicals (OH) by cosmic rays. Both are important reactive chemicals in the atmosphere and, in certain regions, production by GCRs may be a significant, or perhaps the dominant, mechanism. Estimates made over twenty years ago [68, 69] suggest about 1–2 OH radicals and 1.5 NO molecules are produced per ion pair. These production rates would imply maximum mixing ratios of around 1 pptv per day. The production of NO will be measured in pure artificial air and the production of OH will be measured in argon and water vapour. In a second step, the effect of GCR-produced NO and OH on aerosol nucleation, growth and activation will be studied by comparing measurements taken with two different carrier gases: artificial air and argon.

Production of aerosol precursors from trace gases: The aim of these experiments is to quantify the influence of ionising radiation on various rate-limited chemical reactions that are important for atmospheric production of aerosol precursors. For example, sulphuric acid vapour is among the most important aerosol precursor gases for CCN production. It is mainly produced by oxidation of SO₂, which in turn may be emitted directly or else as a product of DMS oxidation (Fig. 35). However the oxidation of SO₂ in the atmosphere is slow since it proceeds by interaction with the rare hydroxyl radical. The lifetime of SO₂ by the OH reaction is about 1 week at typical OH concentrations (10⁵ – 10⁶ molecules cm⁻³). In contrast, the lifetime of SO₂ by dry and wet deposition is about a day. Any process that appreciably increases the rate of SO₂ oxidation will therefore increase the H₂SO₄ concentration. Since GCRs produce OH radicals, they can in principle influence this reaction. There is in fact some experimental evidence that ionising particles may indeed enhance the rate of H₂SO₄ formation [82]. These experiments will involve exposing various atmospheric concentrations of trace gases under different thermodynamic conditions to a range of beam fluxes, and then analysing the reaction products.

Effect of highly ionising particles: Although the majority of cosmic rays in the lower troposphere (below about 6 km) are muons, a few per cent are protons. These undergo nuclear reactions, generating highly ionising nuclear fragments. The fraction of strongly interacting particles (p, n, π^{\pm} and heavy ions) increases at higher altitudes. It is therefore important to investigate the effect of highly ionising fragments on the nucleation of aerosols (and perhaps on other processes summarised here) [135]. These experiments will involve the introduction of small amounts of radon gas into the chamber gas mixtures. Radon isotopes are α emitters, e.g. ²²²Rn decays to ²²⁰Po and a 5.5 MeV α , with a 3.8 d half life. The α particle range in air is about 5 cm at stp, corresponding to an ionisation density, dE/dx , that is about a factor 500 larger than it is for muons (minimum ionising particles). The data taken with Rn irradiation will be compared with beam data to study the effects of high ionisation densities.

5.5.2 Ice particle experiments

Ice particle formation by deposition nucleation: The aim is to study the influence of ionisation on the ability of aerosols to act as ice nuclei (IN) on which water vapour can be directly sublimated to the solid phase (Fig. 38a). The cloud chamber will be filled with selected aerosols and exposed to the beam. The beam is then turned off and expansions made to a range of final temperatures between 0°C and -40°C.

The efficiency of these aerosols to act directly as IN will be measured as a function of temperature and aerosol charge (using the electric field created by the field cage). The presence of an ice crystal in a cloud of drops is readily identified by the CCD cameras since an ice crystal scatters much more light than a water drop [103]. The ice fraction will also be measured with a backscattered polarised laser beam.

Ice particle formation by freezing nucleation and contact nucleation: These experiments will investigate the influence of ionising radiation on the formation of ice particles via the intermediate stage of supercooled liquid droplets. The expansion chamber will be used to create a supercooled cloud by expansion and growth of droplets at temperatures between 0°C and -40°C. The temperature of the droplets can be controlled by the expansion ratio and the initial temperature of the chamber before expansion. After forming the droplets, they are exposed to the beam and the freezing rate is measured. There are several ways to disentangle freezing nucleation (Fig. 38b) from contact nucleation (Fig. 38c). The time development will be different for the two processes (contact nucleation is slower since it is controlled by diffusion of low-mobility aerosols and by sedimentation) and, in addition, CCN that do not act as IN can be included to provide the condensation nucleus for the supercooled droplets. It is important to observe the freezing process over an extended time. This can be achieved by dynamically adjusting the expansion piston so that the supercooled droplets neither evaporate nor grow so large that they are lost by sedimentation.

Secondary production of ice particles: These experiments will study the effect of charge on the production of secondary ice particles by shattering and splintering of freezing droplets. Mixed-phase clouds will be created at various temperatures and the particles grown to sizes where sedimentation and collisions take place. The secondary production of ice particles by contact freezing will be studied. These measurements will use the CCD cameras to follow the evolution of individual particles and their interactions with other particles.

Efficiency of highly-charged evaporation nuclei as IN: These experiments will study the Tinsley mechanism (§4.5.2). This will first involve forming cloud droplets at an appropriate temperature in the range from 0°C to -40°C, in the presence of selected initial aerosols and trace vapours, including both soluble and insoluble (organic) compounds. The droplets will be charged to relatively high values, $\sim 100e$, using the beam and electric field. The droplets will then be evaporated by an adiabatic compression, and the cycle repeated. In this way the effectiveness as IN will be investigated of highly charged aerosols coated with cloud-processed material.

Effect of freezing-thawing cycles on IN efficiency: These experiments will study the effect of freezing-thawing cycles on the efficiency with which various aerosols act as IN. After filling the cloud chamber with the aerosols under study, a series of expansion/compression cycles will be used to activate droplets, allow some to form ice particles, and then evaporate them and repeat the cycle.

Growth of ice particles in mixed-phase clouds: The aims of these experiments are to measure the dynamics and ice-particle growth rates in clouds formed of co-existing liquid and ice particles. Clouds at various temperatures between 0°C and -40°C will be generated containing liquid particles and a few ice particles. The latter will be created using efficient IN such as AgI, in the appropriate low ratios relative to the CCN ($10^{-4} - 10^{-6}$).

Reflectivities of ice and liquid clouds: The aim here is to measure optical reflectivities (albedo) of ice clouds and liquid clouds as functions of water content and particle number concentrations.

Freezing mechanism of polar stratospheric clouds: The aim of these experiments is to investigate the deposition freezing nucleation of HNO_3 and water vapours onto ion clusters, forming nitric acid hydrates. Particles composed of such hydrates are the principal component of the polar stratospheric clouds that initiate the destruction of ozone. The operating temperatures are typical polar stratospheric values of between 190 and 200 K. Nitric acid and water vapour will be present in the chamber at partial pressures representative of the stratosphere (10^{-4} Pa for HNO_3 vapour and $5 \cdot 10^{-2}$ Pa for H_2O vapour, corresponding to about 10 ppb and 5 ppm, respectively). At these pressures and temperatures the nitric acid hydrates become supersaturated and can condense as crystals provided a suitable ice nucleus is present. Sulphuric acid vapour will be included in the air mixture to represent the species most likely to contribute to the initial ion cluster formation.

5.5.3 *Cloud electricity experiments*

Effect of charge and electric field on the coalescence efficiency of droplets: The aim of these experiments is to study the effect of charge and electric field on the efficiency with which cloud droplets coalesce on colliding. Laboratory studies [136] have indicated that raindrops (of around 0.5 mm diameter) are about a factor 100 more efficient at collecting aerosol when they are charged rather than neutral, and theoretical studies support this picture [105]. Since collisional growth is the main mechanism by which cloud droplets grow into raindrops, this is an important process for determining the lifetime of a cloud. These experiments will involve generation of polydisperse droplet distribution in the cloud chamber (from a broad initial CCN distribution). The coalescence efficiency will be measured for individual droplets using the CCD camera system, as a function of droplet size, charge and electric field. In addition, special runs will be taken with high ionisation depositions to simulate the ion densities that may occur in thunderstorm conditions, where the coalescence efficiencies may be greatly enhanced.

Charge separation mechanism in clouds: The aim of these experiments is to investigate whether there are any differences in critical supersaturation between positively and negatively charged CCN. Theoretical studies [112] suggest that negative particles may activate more readily than positive particles. If this occurs under natural cloud conditions then it could lead to charge separation, the precursor for lightning. In these experiments the beam will first ionise the cloud chamber volume and the charge allowed to be scavenged by a predetermined CCN population. The beam will be turned off and the clearing field used to separate the positive and negative charged aerosols. A progressive piston expansion will gradually raise the supersaturation. The activation of the positively-charged CCN region can then be compared with the negatively-charged CCN region. After activation, the charges on the aerosols can be measured by further drifting in an electric field set with the field cage. The activations will be measured as functions of the CCN type, beam intensity, aerosol charge, electric field and adiabatic lapse rate (rate of rise of supersaturation).

Lightning trigger mechanism: The aim here is to investigate the proposed triggering mechanism for lightning due to high energy primary cosmic rays [113]. The charge amplification in the cores of these showers will be studied under typical cloud conditions (i.e. cloud droplets and electric fields). A single pulse of about $2 \cdot 10^4$ beam particles over a 2×2 m² area simulates the peak electron shower density in the core of 10^{15} eV primary GCR. Higher particle densities can easily be generated. The field cage generates electric fields of up to 10 kV m⁻¹, which will simulate the fields in all but lightning clouds. Higher fields typical of lightning breakdown conditions (100 kV m⁻¹) can be generated using specially designed electrodes inserted into the chamber via the sampling probes. For these experiments an X ray detector will be used to observe possible X ray emissions from avalanche processes.

Effect of electrical discharge on cloud droplet formation: The aim of these experiments is to study the effects of electrical discharges on cloud droplet number and on droplet coalescence efficiency, and especially to quantify the effects of NO_x production from electrical discharges. Since GCRs may influence the rate of lightning, they may in turn influence this important source of NO_x in the atmosphere. It is interesting to note that cloud chamber experiments in Edinburgh in the 1960's observed backgrounds due to 'hypersensitive condensation nuclei' [137]. These activated into droplets at very low water vapour supersaturations—of order 1%—and were attributed to trace amounts of NO_2 vapour created by electrostatic discharges. As before, the discharges will be generated by specially designed electrodes inserted into the cloud chamber via the sampling probes.

5.5.4 Data interpretation and cloud modelling

The experimental results obtained with CLOUD will be evaluated with aerosol and cloud simulations. The basic physics aims of these simulations are as follows:

1. To incorporate the microphysics of ion-mediated processes into the aerosol and cloud models.
2. To examine the sensitivity of clouds under atmospheric conditions to variations in the GCR intensity, in the presence of other sources of natural variability.

There will be a close feedback between the simulations and the experimental observations to confirm a) that the simulations closely reproduce the experimental data and b) that the underlying microphysics is understood. The simulations will also eventually become important in guiding the direction of the experimental programme.

As well as a detector simulation, this work will require a complete simulation of the microphysical processes under study. Already existing within the CLOUD collaboration are detailed models of aerosol nucleation, growth and activation, and also two sophisticated cloud models—for cumulus and marine stratus clouds, respectively. Ion-mediated microphysics will be incorporated into these models based on the experimental data from CLOUD. If item 2 above reveals discernible GCR effects on single clouds, suitable parametrisations will be prepared for the climate modelling community to explore the influence of GCR effects on the global climate.

6 CONCLUSIONS

The Sun is a variable star. The sunspot record over the last 400 years reveals both a quasi-periodic solar cycle of about 11 years and also longer-term changes, including a *grand minimum* lasting about 70 years during which the sunspots all but disappeared. This coincided with the most pronounced of several prolonged cold spells between 1450 and 1890 which are collectively known as the Little Ice Age. Light radio-isotope records in ice cores, tree rings and other archives extend the measurements of solar variability back over the last 250 kyr. They reveal numerous occasions when the Sun waxed or waned over centennial periods.

The Earth's climate is far from stable. Natural forcings have caused large climate changes in the past. The most important are the periodic shifts from glacial epochs lasting about 100 kyr to warm inter-glacial epochs lasting about 10 kyr. These transitions involve global average temperature changes of about 5°C , and polar changes of $10\text{--}15^\circ\text{C}$. They coincide with the Milankovitch orbital variations of the Earth around the Sun, but amplification mechanisms are needed to account for the magnitude of the observed climate changes. However, in addition to these large climate shifts, recent palaeoclimatic studies have shown that significant climate variations also occurred within the Holocene and the last ice age on centennial and millennial timescales. There is increasing evidence that many of these are triggered by solar forcing. It appears that the Little Ice Age was but one of around 10 solar-induced cold spells during the last 10,000 years.

However, despite the extensive evidence, solar variability remains a controversial agent of climate change since no causal mechanism has been established to link the two phenomena. Estimates of the variation of the solar irradiance appear to be too small to account for the observed climate variability. Until recently there were no indications where else to look. However, recent satellite data have provided a new clue: low cloud cover may be influenced by the Sun—either as a result of changes in the electromagnetic (UV) radiation or in the galactic cosmic rays, whose flux is modulated by the solar wind. This could provide an effective initial step by which an energetically-weak solar signal is amplified into a significant climate forcing. Subsequently, in response to the cloud radiative and hydrological changes, further amplification processes may occur such as shifts of the global thermohaline system, which can exert major global climatic effects.

The cosmic ray hypothesis for solar-climate variability rests on the microphysical interactions of cosmic rays and clouds, which are experimentally poorly-known. The CLOUD experiment proposes a major investigation of ion-aerosol-cloud microphysics under controlled laboratory conditions using a beam from a particle accelerator, which provides a precisely adjustable and measurable artificial source of cosmic rays. The heart of the experiment is a precision cloud chamber that recreates cloud conditions throughout the atmosphere. CLOUD is designed as a flexible, general purpose detector, and a broad range of measurements of cloud microphysics is planned in the areas of aerosols, ice particles and atmospheric electricity. Unique in the world, this facility would open up an essentially new line of atmospheric research. Its primary task is to pursue the question of how cosmic rays may influence clouds. If clouds respond to the solar variations that modulate the cosmic rays reaching the Earth, there are consequences for the evaluation of climate change, since the 20th century warming coincided with a large increase of solar activity. To settle the issue, one way or the other, is therefore of great importance.

Finally, since this paper begins with a quote, it is perhaps fitting to end with another [138]:

The physics of cloud formation appears in every aspect of the climate variation, including direct heating by sunlight. I would hope that with the important issues that need to be decided in connection with climate variation and global warming, an all out effort can be launched to understand this vital aspect of atmospheric science. There is a lot of laboratory work to be done under carefully controlled conditions, e.g. the CERN [CLOUD] experiment, on the nucleation of aerosols, ice crystals, and water drops, as well as the related field work to study the application of the laboratory results.

...

Our ability to handle the scientific challenge of climate change will be a subject of future historical research and writing. The eyes of future generations will be upon us. There has never been an individual scientific problem that will have so much impact as the global warming inquiry on the long term well being of the human race.

*Eugene N. Parker, University of Chicago
Conference Summary, The Solar Cycle and Terrestrial Climate
Santa Cruz de Tenerife, September 2000*

Acknowledgements

It is a pleasure to thank Jürg Beer, Gordon Bowden, Nigel Calder, Maurice Jacob, Lew Keller, Ralph Nelson, David Ritson, Arnold Wolfendale and my CLOUD collaborators for many interesting and enjoyable discussions.

References

- [1] H. Svensmark and E. Friis-Christensen, Variation in cosmic ray flux and global cloud coverage—a missing link in solar-climate relationships, *Journal of Atmospheric and Solar-Terrestrial Physics* **59** (1997) 1225.

- [2] N. Marsh and H. Svensmark, Low cloud properties influenced by solar activity, <http://arXiv.org/abs/physics/000507>, *Phys. Rev. Lett.* **85** (2000) 5004–5007.
- [3] CLOUD collaboration (B. Fastrup *et al.*), A study of the link between cosmic rays and clouds with a cloud chamber at the CERN PS, http://www.cern.ch/cloud/iaci_workshop/cloud.html, CERN proposal SPSC/P317, SPSC 2000-021 (2000).
- [4] CLOUD collaboration (B. Fastrup *et al.*), Addendum to the CLOUD proposal, http://www.cern.ch/cloud/iaci_workshop/cloud.html, CERN SPSC/P317 Add.1, SPSC 2000-030 (2000).
- [5] CLOUD collaboration (B. Fastrup *et al.*), CLOUD: an Atmospheric Research Facility at CERN, http://www.cern.ch/cloud/iaci_workshop/cloud.html, CERN SPSC/P317 Add.2, SPSC 2000-041 (2000).
- [6] J.M. Wilcox, Solar activity and the weather, *J. Atm. & Sol.-Terr. Phys.* **37** (1975) 237–256.
- [7] W. Herschel, Observations tending to investigate the nature of the Sun, in order to find out the causes or symptoms of its variable emission of light and heat, *Philosophical Transactions of the Royal Society*, London (1801) 265.
- [8] J.A. Eddy, The Maunder minimum, *Science* **192** (1976) 1189.
- [9] C. Fröhlich and J. Lean, The Sun’s total irradiance: Cycles and trends in the past two decades and associated climate change uncertainties, *Geophys. Res. Lett.* **25** (1998) 4377–4380.
- [10] J. Lean, J. Beer and R.S. Bradley, Reconstruction of solar irradiance since 1610: Implications for climate change, *Geophys. Res. Lett.* **22** (1995) 3195–3198.
- [11] M. Lockwood, R. Stamper and M.N. Wild, A doubling of the Sun’s coronal magnetic field during the past 100 years, *Nature* **399** (1999) 437.
- [12] S.C. Kernthaler, R. Toumi and J.D. Haigh, Some doubts concerning a link between cosmic ray fluxes and global cloudiness, *Geophys. Res. Lett.* **26** (1999) 863–865.
- [13] T.S. Jorgensen and A.W. Hansen, *Journal of Atmospheric and Solar-Terrestrial Physics* **62** (1997) 73–77. A reply is in the adjacent paper: H. Svensmark and E. Friis-Christensen, *Journal of Atmospheric and Solar-Terrestrial Physics* **62** (1997) 79–80.
- [14] J.E. Kristjansson and J. Kristiansen, Is there a cosmic ray signal in recent variations in global cloudiness and cloud radiative forcing?, *J. Geophys. Res.* **105** (2000) 11851–11863.
- [15] W.B. Rossow, A.W. Walker, D.E. Beuschel, and M.D. Roiter, *International Satellite Cloud Climatology Project (ISCCP): documentation of new cloud datasets*, WMO/TD **737**, World Meteorological Organization, Geneva (1996).
- [16] M. Lockwood, Changes in galactic cosmic ray flux, solar irradiance and climate over the past 150 years, in *Proc. of the Workshop on Ion-Aerosol-Cloud Interactions*, ed. J. Kirkby, CERN, Geneva, CERN 2001-007 (2001).
- [17] D.V. Hoyt and K.H. Schatten, *The Role of the Sun in Climate Change*, Oxford University Press, Oxford (1997).
- [18] C.E.P. Brooks, Variations in the levels of the central African Lakes Victoria and Albert, *Geophysical Memoirs, London*, **2** (1923) 327–334.

- [19] G.W. Lockwood and D.T. Thompson, Photometric variability of Neptune, 1972–2000, in press (2001).
- [20] N. Marsh and H. Svensmark, GCR and ENSO trends in ISCCP-D2 low cloud properties, submitted to *J. Geophys. Res.* (2001); also http://www.dsri.dk/~ndm/CLOUD_UPDATE/UPDATE.html.
- [21] D.L. Hartmann, Radiative effects of clouds on Earth's climate, in *Aerosol-Cloud-Climate Interactions*, International Geophysics Series, Vol. 54, ed. P.V. Hobbs, Academic Press Inc., San Diego (1993) 151.
- [22] J. Beer, Ice core data on climate and cosmic ray changes, in *Proc. of the Workshop on Ion-Aerosol-Cloud Interactions*, ed. J. Kirkby, CERN, Geneva, CERN 2001-007 (2001).
- [23] J.R. Herman and R.A. Goldberg, *Sun, Weather, and Climate*, NASA SP 426 (1978).
- [24] K.C. Taylor *et al.*, The Holocene-Younger Dryas transition recorded at Summit, Greenland, *Science* **278** (1997) 825.
- [25] W.S. Broecker and G.H. Denton, What drives glacial cycles?, *Scientific American* **262** (1990) 43–50.
- [26] B. Stenni *et al.*, An ocean reversal during the last deglaciation, *Science* **293** (2001) 2074–2077.
- [27] J.P. Sachs *et al.*, Glacial surface temperatures of the Southeast Atlantic Ocean, *Science* **293** (2001) 2077–2079.
- [28] R. Muscheler, J. Beer, G. Wagner and R.C. Finkel, Changes in deep-water formation during the Younger Dryas event inferred from ^{10}Be and ^{14}C records, *Nature* **408** (2000) 567–570.
- [29] H. Renssen, B. van geel, J. van der Plicht and M. Magny, Reduced solar activity as a trigger for the start of the Younger Dryas?, *Quaternary International* **68–71** (2000) 373–383.
- [30] G.C. Bond and R. Lotti, Iceberg discharges into the North Atlantic on millennial time scales during the last glaciation, *Science* **267** (1997) 1005–1010.
- [31] G.C. Bond *et al.*, A pervasive millennial-scale cycle in North Atlantic Holocene and glacial climates, *Science* **278** (1997) 1257–1266.
- [32] A. Ganopolski and S. Rahmstorf, Rapid changes of glacial climate simulated in a coupled climate model, *Nature* **409** (2001) 153–158.
- [33] G. Wagner, C. Laj, J. Beer *et al.*, Reconstruction of the paleo accumulation rate of central Greenland during the last 75 kyr using the cosmogenic radionuclides ^{36}Cl and ^{10}Be , and geomagnetic field intensity data, in press (2001).
- [34] G.C. Bond, B. Kromer, J. Beer, R. Muscheler, M.N. Evans, W. Showers, S. Hoffmann, R. Lotti-Bond, I. Hajdas and G. Bonani, Persistent solar influence on North Atlantic climate during the Holocene, *Science* **294** (2001) 2130–2136.
- [35] M. Magny, Lake level fluctuations in the Jura and French subalpine regions associated with ice-rafting in the North Atlantic and variations in the polar atmospheric circulation, *Quaternaire* **10** (1999) 61–64.
- [36] U. Neff *et al.*, Strong coincidence between solar variability and the monsoon in Oman between 9 and 6 kyr ago, *Nature* **411** (2001) 290–293.

- [37] B. van Geel, J. Buurman and H.T. Waterbolk, Archaeological and palaeoecological indications of an abrupt climate change in The Netherlands, and evidence for climatological teleconnections, around 2650 BP, *J. Quaternary Science* **11** (1996) 451–460.
- [38] M. Grosjean *et al.*, A late Holocene (<2600 BC) glacial advance in the south-central Andes (29°S), northern Chile, *The Holocene* **8** (1998) 473–479.
- [39] V.J. Polyak and Y. Asmerom, Late Holocene climate and cultural changes in the southwestern United States, *Science* **294** (2001) 148–151.
- [40] P.E. Damon and C.P. Sonett, Solar and terrestrial components of the atmospheric ^{14}C variation spectrum, in *The Sun in Time*, eds. C.P. Sonett, M.S. Giampapa and M.S. Matthews, University of Arizona Press, Tucson (1991) 360.
- [41] G.A. Brook *et al.*, *The Holocene* **9** (1999) 695.
- [42] R.F. Denniston *et al.*, *Quat. Res.* **53** (2000) 196.
- [43] D. Verschuren, K. Laird and B. Cumming, Rainfall and drought in equatorial East Africa during the past 1100 years, *Nature* **403** (2000) 410–414.
- [44] G. Cini Castagnoli *et al.*, The role of cloud cover on the solar illumination signal recorded by $\delta^{13}\text{C}$ of a shallow water Ionian Sea core, in *Proc. of the Workshop on Ion-Aerosol-Cloud Interactions*, ed. J. Kirkby, CERN, Geneva, CERN 2001-007 (2001).
- [45] J. Beer, S. Tobias and N. Weiss, An active Sun throughout the Maunder Minimum, *Solar Physics* **181** (1998) 237–249.
- [46] J. Beer, A. Blinov, G. Bonani, H.J. Hofmann and R.C. Finkel, Use of ^{10}Be in polar ice to trace the 11-year cycle of solar activity, *Nature* **347** (1990) 164–166.
- [47] M. Lockwood, *J. Geophys. Res.* **106** (2001) 16021–16038.
- [48] T. Landscheidt, Solar forcing of El Niño and La Niña, in *Proc. of the First Solar & Space Weather Euroconference; The Solar Cycle and Terrestrial Climate*, Santa Cruz de Tenerife, Spain, ESA SP-463 (2000) 2135–2140.
- [49] *Climate Change 2001: The Scientific Basis*, Intergovernmental Panel on Climate Change, eds. J.T. Houghton *et al.*, Cambridge University Press, UK (2001).
- [50] G.C. Reid, Influence of solar variability on global sea surface temperatures, *Nature* **329** (1987) 142–143.
- [51] D. Ritson, private communication.
- [52] W.B. White, J. Lean, D.R. Cayan and M.D. Dettinger, Response of global upper ocean temperature to changing solar irradiance, *J. Geophys. Res.* **102** C2 (1997) 3255–3266.
- [53] W.B. White, D.R. Cayan and J. Lean, Global upper ocean heat storage response to radiative forcing from changing solar irradiance and increasing greenhouse gas/aerosol concentrations, *J. Geophys. Res.* **103** C10 (1998) 21,333–21,366.
- [54] R.J. Charlson, J.H. Seinfeld, A. Nenes, M. Kulmala, A. Laaksonen and M. Christina Facchini, Reshaping the theory of cloud formation, *Science* **292** (2001) 2025–2026.

- [55] Y.I. Stozhkov, N.S. Svirzhevsky and V.S. Makhmutov, Cosmic ray measurements in the atmosphere, in *Proc. of the Workshop on Ion-Aerosol-Cloud Interactions*, ed. J. Kirkby, CERN, Geneva, CERN 2001-007 (2001).
- [56] Y.C. Unruh, S.K. Solanki and M. Fligge, Modelling solar irradiance variations: comparison with observations, including line-ratio variations, in *Solar Variability and Climate*, eds. E. Friis-Christensen *et al.*, *Space Science Reviews* **94**, Nos. 1–2 (2000) 145–152.
- [57] J.D. Haigh, The impact of solar variability on climate, *Science* **272** (1996) 981.
- [58] D. Shindell *et al.*, Solar cycle variability, ozone, and climate, *Science* **284** (1999) 305–308.
- [59] K.J.H. Phillips, *Guide to the Sun*, Cambridge University Press, UK (1992).
- [60] N.O. Weiss and S.M. Tobias, Physical causes of solar activity, in *Solar Variability and Climate*, eds. E. Friis-Christensen *et al.*, *Space Science Reviews* **94**, Nos. 1–2 (2000) 99–112.
- [61] J. Schou *et al.*, Helioseismic studies of differential rotation in the solar envelope by the Solar Oscillation Investigation using the Michelson Doppler Imager, *Astrophysical Journal* **505** (1998) 390–417.
- [62] E.N. Parker, The passage of energetic charged particles through interplanetary space, *Planet. Space Sc.* **13** (1965) 9–49.
- [63] L.J. Gleeson and W.I. Axford, Cosmic rays in the interplanetary medium, *Astrophys. Journal* **149** (1967) 1115–1118.
- [64] D.E. Groom *et al.* (Particle Data Group), The Review of Particle Physics, *The European Physical Journal* **C15** (2000) 150–156.
- [65] V.I. Ermakov, G.A. Bazilevskaya, P.E. Pokrevsky and Y.I. Stozhkov, Ion balance equation in the atmosphere, *J. Geophys. Res.* **102** (1997) 23413.
- [66] S. Tonwar (private communication, based on data from the L3+C experiment at CERN).
- [67] P.J. Crutzen and F. Arnold, Nitric acid formation in the cold antarctic atmosphere: a major cause for the springtime ‘ozone hole’, *Nature* **324** (1986) 651.
- [68] M. Nicolet, On the production of nitric oxide by cosmic rays in the mesosphere and stratosphere, *Planet. Space Sci.* **23** (1975) 637.
- [69] C.H. Jackman, J.E. Frederick and R.S. Stolarski, Production of odd nitrogen in the stratosphere and mesosphere: an intercomparison of signal strengths, *J. Geophys. Res.* **85** C12 (1980) 7495.
- [70] P.J. Crutzen, I.A.S. Isaksen and G.C. Reid, Solar proton events: stratospheric sources of nitric oxide, *Science* **189** (1975) 457.
- [71] A.A. Viggiano and F. Arnold, Ion chemistry and composition of the atmosphere, *Handbook of Atmospheric Electrodynamics*, Vol. 1, ed. H. Volland, CRC Press, Boca Raton (1998).
- [72] H. Heitmann and F. Arnold, Composition measurements of tropospheric ions, *Nature* **306** (1993) 747.
- [73] A. Krieger and F. Arnold, First composition measurements of stratospheric negative ions and inferred gaseous sulphuric acid in the winter arctic vortex: Implications for aerosols and hydroxyl radical formation, *Geophys. Res. Lett.* **21** (1994) 1259.

- [74] C.T.R. Wilson, Expansion apparatus, *Proc. Royal Society of London A* **87** (1912) 277.
- [75] J.H. Seinfeld and S.N. Pandis, *Atmospheric Chemistry and Physics*, Wiley, New York (1998).
- [76] S. Twomey, Aerosols, clouds, and radiation, *Atmos. Environ.* **25A** (1991) 2435–2442.
- [77] P.V. Hobbs, Aerosol-cloud interactions, in *Aerosol-Cloud-Climate Interactions*, International Geophysics Series, Vol. 54, ed. P.V. Hobbs, Academic Press Inc., San Diego (1993) 33–73.
- [78] S.E. Schwartz and A. Slingo, Enhanced shortwave cloud radiative forcing due to anthropogenic aerosols, in *Clouds, Chemistry and Climate*, Proceedings of NATO Advanced Study Workshop, ed. P. Crutzen and V. Ramanathan, Springer, Heidelberg (1996) 191–236.
- [79] L.F. Radke, J.A. Coakley and M.D. King, *Science* **246** (1989) 1146–1149.
- [80] D. Rosenfeld, Suppression of rain and snow by urban and industrial air pollution, *Science* **287** (2000) 1793.
- [81] F. Arnold, Multi-ion complexes in the stratosphere: implications for trace gases and aerosols, *Nature* **284** (1980) 610.
- [82] F. Raes and A. Janssens, Ion-induced aerosol formation in a H₂O-H₂SO₄ system - I. Extension of the classical theory and search for experimental evidence, *J. Aerosol Sci.* **16** (1985) 217.
- [83] F. Raes, A. Janssens and R.V. Dingenen, The role of ion-induced aerosol formation in the lower atmosphere, *J. Aerosol Sc.* **17** (1986) 466.
- [84] I. Kusaka, Z.G. Wang and J.H. Seinfeld, Ion-induced nucleation: a density functional approach, *J. Chem. Phys.* **102** (1995) 913–924.
- [85] I. Kusaka, Z.G. Wang and J.H. Seinfeld, Ion-induced nucleation. II. Polarizable multipolar molecules, *J. Chem. Phys.* **103** (1995) 8993–9009.
- [86] R.P. Turco, J.-K. Zhao and F. Yu, A new source of tropospheric aerosols: ion-ion recombination, *Geophys. Res. Lett.* **25** (1998) 635.
- [87] F. Yu and R.P. Turco, Ultrafine aerosol formation via ion-mediated nucleation, *Geophys. Res. Lett.* **27** (2000) 883.
- [88] M. Kulmala, Atmospheric aerosols: formation and growth, in *Proc. of the Workshop on Ion-Aerosol-Cloud Interactions*, ed. J. Kirkby, CERN, Geneva, CERN 2001-007 (2001).
- [89] A.D. Clarke *et al.*, Particle nucleation in the tropical boundary layer and its coupling to marine sulfur sources, *Science* **282** (1998) 89.
- [90] R.J. Weber *et al.*, Measurement of new particle formation and ultrafine particle growth at a clean continental site, *J. Geophys. Res.* **102** (1997) 4375–4385.
- [91] R.J. Weber *et al.*, New particle formation in the remote troposphere: a comparison of observations at various sites, *Geophys. Res. Lett.* **26** (1999) 307–310.
- [92] M.K. Kulmala, L. Pirjola and J.M. Mäkelä, Stable sulphate clusters as a source of new atmospheric particles, *Nature* **404** (2000) 66.
- [93] F. Yu and R.P. Turco, From molecular clusters to nanoparticles: The role of ambient ionisation in tropospheric aerosol formation, *J. Geophys. Res.* **106** (2001) 4797–4814.

- [94] R. D'Auria and R.P. Turco, A thermodynamic-kinetic model for ionic cluster formation, growth and nucleation, *Proc. of the Workshop on Ion-Aerosol-Cloud Interactions*, ed. J. Kirkby, CERN, Geneva, CERN 2001-007 (2001).
- [95] F. Yu, Cosmic rays, particle formation, natural variability of global cloudiness, and climate implications, in *Proc. of the Workshop on Ion-Aerosol-Cloud Interactions*, ed. J. Kirkby, CERN, Geneva, CERN 2001-007 (2001).
- [96] U. Hörrak, J. Salm and H. Tammet, Bursts of intermediate ions in atmospheric air, *J. Geophysical Research* **103** D12 (1998) 13909.
- [97] J. Bricard *et al.*, Formation and evolution of nuclei of condensation that appear in air initially free of aerosols, *J. Geophysical Research* **73** (1968) 4487.
- [98] K.G. Vohra, M.C. Subba Ramu and T.S. Muraleedharan, An experimental study of the role of radon and its daughter products in the conversion of sulphur dioxide into aerosol particles in the atmosphere, *Atmospheric Environment* **18** (1984) 1653.
- [99] B. Zuberi *et al.*, Heterogeneous freezing of aqueous particles induced by crystallized $(\text{NH}_4)_2\text{SO}_4$, ice, and letovicite, *J. Physical Chemistry A* **105** (2001) 6458–6464.
- [100] B.A. Tinsley and G.W. Dean, Apparent tropospheric response to MeV–GeV particle flux variations: a connection via the solar wind, atmospheric electricity and cloud microphysics, *J. Geophysical Research* **96** (1991) 22283.
- [101] B.A. Tinsley, Influence of solar wind on the global electric circuit, and inferred effects on cloud microphysics, temperature, and dynamics in the troposphere, in *Solar Variability and Climate*, eds. E. Friis-Christensen *et al.*, *Space Science Reviews* **94**, Nos. 1–2 (2000) 231–258.
- [102] H.R. Pruppacher and J.D. Klett, *Microphysics of Clouds and Precipitation*, 2nd ed., Kluwer Academic Publishers, The Netherlands (1997).
- [103] R.J. Anderson, R.C. Miller, J.L. Kassner and D.E. Hagen, A study of homogeneous condensation-freezing nucleation of small water droplets in an expansion cloud chamber, *J. Atmospheric Sciences* **37**, No. 11 (1980) 2508–2520.
- [104] A.E. Waibel, T. Peter, K.S. Carslaw, H. Oelhaf, G. Wetzell, P.J. Crutzen, U. Pöschl, A. Tsias, E. Reimer and H. Fischer, Arctic ozone loss due to denitrification, *Science* **283** (1999) 2064–2069.
- [105] D.R. MacGorman and W.D. Rust, *The Electrical Nature of Storms*, Oxford University Press, UK (1998).
- [106] K. Schlegel *et al.*, Thunderstorms, lightning and solar activity—Middle Europe, *J. Atmospheric and Solar-Terrestrial Physics*, **63** (2001) 1705–1713.
- [107] R.G. Roble, On solar-terrestrial relationships in the atmospheric electricity, *J. Geophys. Res.* **90** (1985) 6000.
- [108] R.E. Orville and G.R. Huffines, Lightning ground flash measurements over the contiguous United States: a ten-year summary 1989–1998, in *Proc. 11th Int. Conf. on Atmospheric Electricity*, Alabama, USA (1999) 412.
- [109] B.J. Kellett, The production of atmospheric nitric oxide by cosmic rays and solar energetic particles, in *Proc. of the Workshop on Ion-Aerosol-Cloud Interactions*, ed. J. Kirkby, CERN, Geneva, CERN 2001-007 (2001).

- [110] Yu. I. Stozhkov *et al.*, Rainfalls during great Forbush decreases, *Il Nuovo Cimento* **18C** (1995) 335.
- [111] M.I. Pudovkin and S.V. Veretenenko, Cloudiness decreases associated with Forbush decreases of galactic cosmic rays, *J. Atmos. Solar-Terrestrial Phys.* **75** (1995) 1349.
- [112] A.I. Rusanov and F.M. Kuni, Reformulation of the thermodynamic theory of nucleation on charged particles, *J. Colloid and Interface Science* **100** (1984) 264.
- [113] A.V. Gurevich, K.P. Zybin and R. Roussel-Dupre, Lightning initiation by simultaneous effect of runaway breakdown and cosmic ray showers, *Phys. Lett. A* **254** (1999) 79–87.
- [114] A.P. Chubenko *et al.*, Intensive X-ray emission bursts during thunderstorms, *Phys. Lett. A* **275** (2000) 90–100.
- [115] C.B. Moore, K.B. Eack, G.D. Aulich and W. Rison, Energetic radiation associated with lightning stepped-leaders, *Geophys. Res. Lett.* **28** No. 11 (2001) 2141–2144.
- [116] P. Galison, *Image and Logic; a Material Culture of Microphysics*, University of Chicago Press, Chicago (1997).
- [117] A. Nenes, P.Y. Chuang, R.C. Flagan and J.H. Seinfeld, A theoretical analysis of cloud condensation nucleus (CCN) instruments, *J. Geophys. Res.* **106** D4 (2001) 3449–3474.
- [118] P.E. Wagner, Experiments on nucleation processes in aerosols, in *Proc. of the Workshop on Ion-Aerosol-Cloud Interactions*, ed. J. Kirkby, CERN, Geneva, CERN 2001-007 (2001).
- [119] J.L. Schmitt, Comments on the operation of a Wilson expansion cloud chamber, in *Proc. of the Workshop on Ion-Aerosol-Cloud Interactions*, ed. J. Kirkby, CERN, Geneva, CERN 2001-007 (2001).
- [120] N.N. Das Gupta and S.K. Ghosh, A report on the Wilson cloud chamber and its application in physics, *Rev. Modern Physics* **18** (1946) 225–290.
- [121] H. Bunz *et al.*, The novel aerosol chamber facility AIDA: status and first results, in *Proc. 7th European Symposium on the Physico-Chemical Behaviour of Atmospheric Pollutants*, eds. B. Larson, B. Versino and G. Angeletti, Venice (1996) 673.
- [122] O. Möhler *et al.*, The Karlsruhe aerosol chamber facility AIDA: Technical description and first results of homogeneous and heterogeneous ice nucleation experiments, in *Proc. of the Workshop on Ion-Aerosol-Cloud Interactions*, ed. J. Kirkby, CERN, Geneva, CERN 2001-007 (2001).
- [123] G.B. Bowden, private communication.
- [124] A. Hervé and G. Bachy, Huge electrohydraulic servovalves control 2-ton cryogenic piston, *Hydraulics & Pneumatics*, Pention/IPC, Cleveland, Ohio (1977); and, The hydraulic expansion servo-system of the Big European Bubble Chamber, CERN preprint CERN/EF 77-6 (1977).
- [125] Bellofram Corporation, Burlington, MA 01803, USA, <http://www.marshbellowfram.com>.
- [126] W.C. Hinds, *Aerosol Technology: Properties, Behavior, and Measurement of Airborne Particles*, John Wiley & Sons, Inc., New York (1998).
- [127] P.E. Wagner, A constant-angle Mie scattering method (CAMS) for investigation of particle formation processes, *J. Colloid Interface Sci.* **105** (1985) 456.

- [128] F. Pavese and G. Molinar, Modern Gas-Based Temperature and Pressure Measurements, *The International Cryogenics Monograph Series*, Plenum Press, New York (1992).
- [129] T.P. Burke and J.A. Scott, The production of condensation nuclei by alpha radiation, *Proc. Royal Irish Academy* **73** A (1973) 151–158.
- [130] J.M. Mäkelä, Studies on irradiation induced aerosol particle formation processes in air containing sulphur dioxide, Ph.D. thesis, Univ. Helsinki, *Acta Polytechnica Scandinavica*, Applied Physics Series **82** (1992).
- [131] C.F. Clement and R.G. Harrison, Enhanced localised charging of radioactive aerosols, *J. Aerosol. Sci.* **31**, 3 (2000) 363–378.
- [132] A.S. Wexler and Z. Ge, Hydrophobic particles can activate at lower relative humidity than slightly hygroscopic ones: a Köhler theory incorporating surface fixed charge, *J. Geophys. Res.* **103** D6 (1998) 6083–6088.
- [133] M. Kulmala, A. Laaksonen, R.J. Charlson and P. Korhonen, Clouds without supersaturation, *Nature* **388** (1997) 336.
- [134] P.Y. Chuang, R.J. Charlson and J.H. Seinfeld, Kinetic limitations on droplet formation in clouds, *Nature* **390** (1997) 594–596.
- [135] A.D. Erlykin and A.W. Wolfendale, Cosmic rays and the Earth's atmosphere, in *Proc. of the Workshop on Ion-Aerosol-Cloud Interactions*, ed. J. Kirkby, CERN, Geneva, CERN 2001-007 (2001).
- [136] A.K. Barlow and J. Latham, A laboratory study of the scavenging of sub-micron aerosol by charged raindrops, *Quart. J. R. Met. Soc.* **109** (1983) 763–770.
- [137] G.R. Evans and A.A. Watson, Production of hypersensitive condensation nuclei, *Nature* **196** (1962) 729; G.R. Evans *et al.*, Some remarks on the origin and elimination of background cloud in expansion cloud chambers, *J. Sci. Inst.* **41** (1964) 770; G.R. Evans and A.A. Watson, Nitrogen dioxide vapour and the creation of condensation nuclei, *Proc. Royal Society of Edinburgh* **LXVII** part III-10 (1966) 136.
- [138] E.N. Parker, Summary and perspectives, in *Proc. of the First Solar & Space Weather Euroconference; The Solar Cycle and Terrestrial Climate*, Santa Cruz de Tenerife, Spain, ESA SP-463 (2000) 263–269.