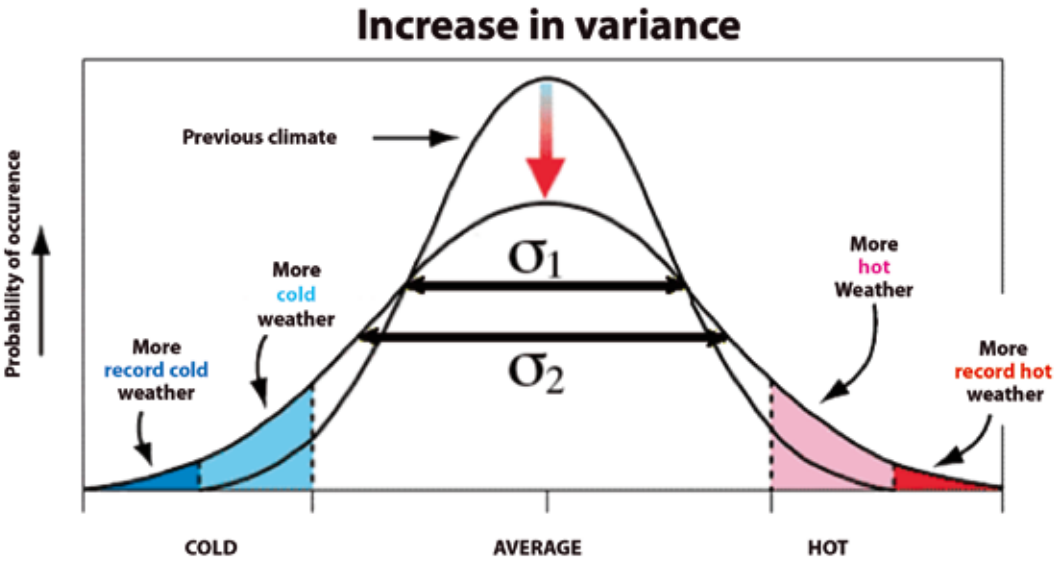
Reference: WG1 2007, Uppala et al. (2005), Kalnay et al. (1996)

1 Probabilities



MODIFIED AND BASED ON CONTRIBUTION OF WORKING GROUP I TO THE THIRD ASSESSMENT REPORT OF THE INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE, CLIMATE CHANGE 2001: THE SCIENTIFIC BASIS. FIGURE 2.32B. CAMBRIDGE UNIVERSITY PRESS.

FIGURE 1 Schematic showing the effect on extreme temperatures when the variance increases for a normal distribution of temperature.

1.1 Probability distribution

Variance and standard deviation are indicators of the dispersion of a perfectly symmetrical (Gaussian) bell curve. The standard deviation indicates how, on average, the values of the variable are grouped around the mean. A small standard deviation means that the values are not widely distributed around the mean (s1: the distribution is smaller and the probability of occurrence at the mean value approaches to 100%). Conversely, a large standard deviation signifies that the values are widely spread around the mean (s2: the distribution is wider and the probability of occurrence at the mean value becomes smaller).

The standard deviation is the square root of the variance; the latter is larger in the case of series 2 (see FIGURE 1).

In reality, for any given variable sample, the availability of data values is limited. The empirical probability distribution function built from these values is only an approximation of the probability distribution function.

1.2 Quantiles and percentiles

For a given set of values, it is also common to sort values (temperature for example) in ascending order, and to build an empirical cumulative probability distribution function. Q-quantiles are points taken at regular intervals from the cumulative function in order to divide data into q equal-

sized data subsets. The quantiles are the data values marking regular intervals of this cumulative distribution. If the division is into 100 subsets, the quantiles are called percentiles. For example, 5% of the values are above the 95th percentile and 5% of the values are below the 5th percentile.

The IPCC AR4 report refers frequently to percentiles. For example, TABLE 1 of paragraph 2 of the main sheets called "A1B" (for Europe and China) shows the future changes in temperature and precipitation percentiles.

The 50th percentile, known as the median, divides the set of values into two subsets of same size. This value is different from the mean value. The mean approaches the 50th percentile when the distribution is symmetrical, or Gaussian (in other words, with a bell curve like that in FIGURE 1).

The 5th and 95th percentiles represent the distribution tails or "extreme" values (the respective blue and red areas in FIGURE 1).

1.3 Normalized climatic indices

This comment intends to facilitate the interpretation of the figures from the paper of Tebaldi et al. (2006) which appears in the main sheets on temperature and precipitation in Europe and China (known as the "A1B sheets"). They present the change in several climatic indices (heat waves, hot nights, frost days, dry days and extremes of precipitation) for the periods 2080-2099 and 1980-1999. These indices have been calculated from data sets of 9 climatic models. In order to compare data sets with various standard deviations, Tebaldi et al. use normalized values of indices (normalisation related to the standard deviation) rather than the raw data.

For example, let us consider the change in the number of heat wave days in two models. We suppose that the number of heat wave days is equal for both datasets, but their standard deviation is different: series 1 has a smaller one than series 2. Though the number of heat wave days is equal in both series, the normalized indices value of series 1 is higher than series 2. This normalization thus allows a weighting of the index, taking the accuracy of each series into account.

1.4 Student's t-test

A statistical hypothesis test is a method of rejecting or accepting a statistical hypothesis (statistical decisions) using experimental data. The student's t-test is used to compare a sample mean with a hypothesized true mean.

1.5 Return periods

A return period is a statistical estimate of the average interval of time between two events of a certain intensity or size. This concept introduces a certain probability and is commonly used to characterise extreme events.

2 The IPCC "graded" terminology

The following terms have been used to indicate the likelihood of an event or a result. These terms indicate the consistency between studies, the number of studies and expert opinions.

(« *virtualy certain* » en anglais)	probabilité > 99 %
(« *extremely likely* » en anglais)	probabilité > 95 %
(« *very likely* » en anglais)	probabilité > 90 %
(« *likely* » en anglais)	probabilité > 66 %
(« *more likely than not* » en anglais)	probabilité > 50 %
(« *unlikely* » en anglais)	probabilité < 33 %
(« *very unlikely* » en anglais)	probabilité < 10 %
(« *extremely unlikely* » en anglais)	probabilité < 5 %

3 The NCEP/NCAR data and reanalysis centres

Meteorological centres produce analyses of meteorological data. These analyses result from a combination of simulated and observed fields. The meteorological observations are incorporated into the simulated fields in order to better reproduce the state of the atmosphere at any given moment. Nevertheless the fields obtained are not consistent in time because of the evolution of assimilation techniques and of the observations systems. A reanalysis thus involves to reprocessing the all observation fields for a given period, typically from 40 to 50 years, using the same assimilation technique. Different data reanalysis centres exist. For example, the ECMWF (European Centre for Medium-Range Weather Forecasting Uppala et al., 2005) reanalyses, are frequently used by European groups. The NCEP and NCAR reanalyses (Kalnay et al., 1996) are American reanalyses produced by the National Centers of Environmental Prediction and the National Center for Atmospheric Research.