

Introduction to making a weather-forecast

Starting next week we will make a weather forecast on Wednesday for the next Sunday (4 days ahead)

You will do this in teams of 4-5 people.

It will be a competition.

You will compete, not only with your fellow students teams, but also with KNMI, Weerplaza, “persistence”, “climate” and the “teachers”.

Success of weather forecasting

3 days ahead seems “solved”; 5 days ahead is still a challenge

Nature, 3 September 2015:

Very optimistic
portrayal of success
of the weather
forecast !

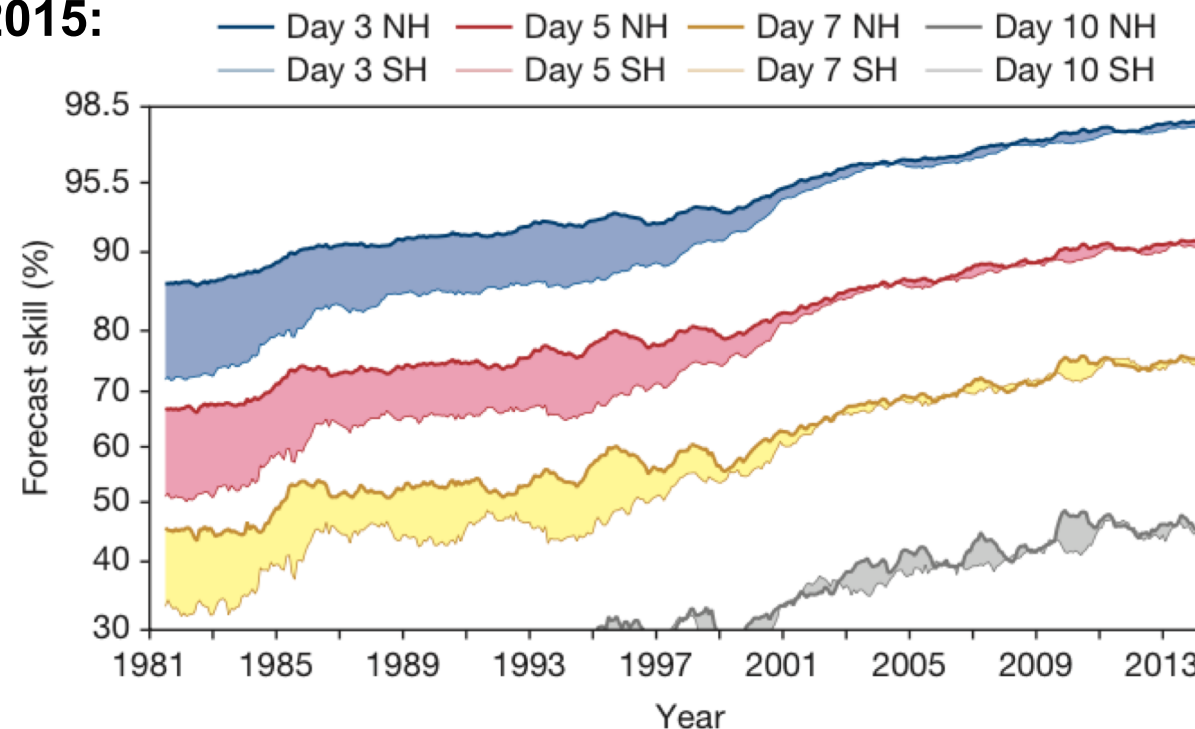


Figure 1 | A measure of forecast skill at three-, five-, seven- and ten-day ranges, computed over the extra-tropical northern and southern hemispheres. Forecast skill is the correlation between the forecasts and the verifying analysis of the height of the 500-hPa level, expressed as the anomaly with respect to the climatological height. Values greater than 60% indicate useful forecasts, while those greater than 80% represent a high degree of accuracy. The convergence of the curves for Northern Hemisphere (NH) and Southern Hemisphere (SH) after 1999 indicates the breakthrough in exploiting satellite data through the use of variational data¹⁰⁰.

You are going to make a “categorical forecast”

Statement that can either be **true** (1 point) or **false** (0 points)

A categorical temperature (maximum/minimum) forecast states that the temperature will either be “**normal**”, “**above normal**” or “**below normal**”.

The normal temperature is deduced from **observations** over the past 30 years for the particular day of the forecast.

Observations, retrieved from knmi.nl

Daily observations at De Bilt
of 13/9, 1985-2014 (30 years).

Diagram illustrating the input variables for the model:

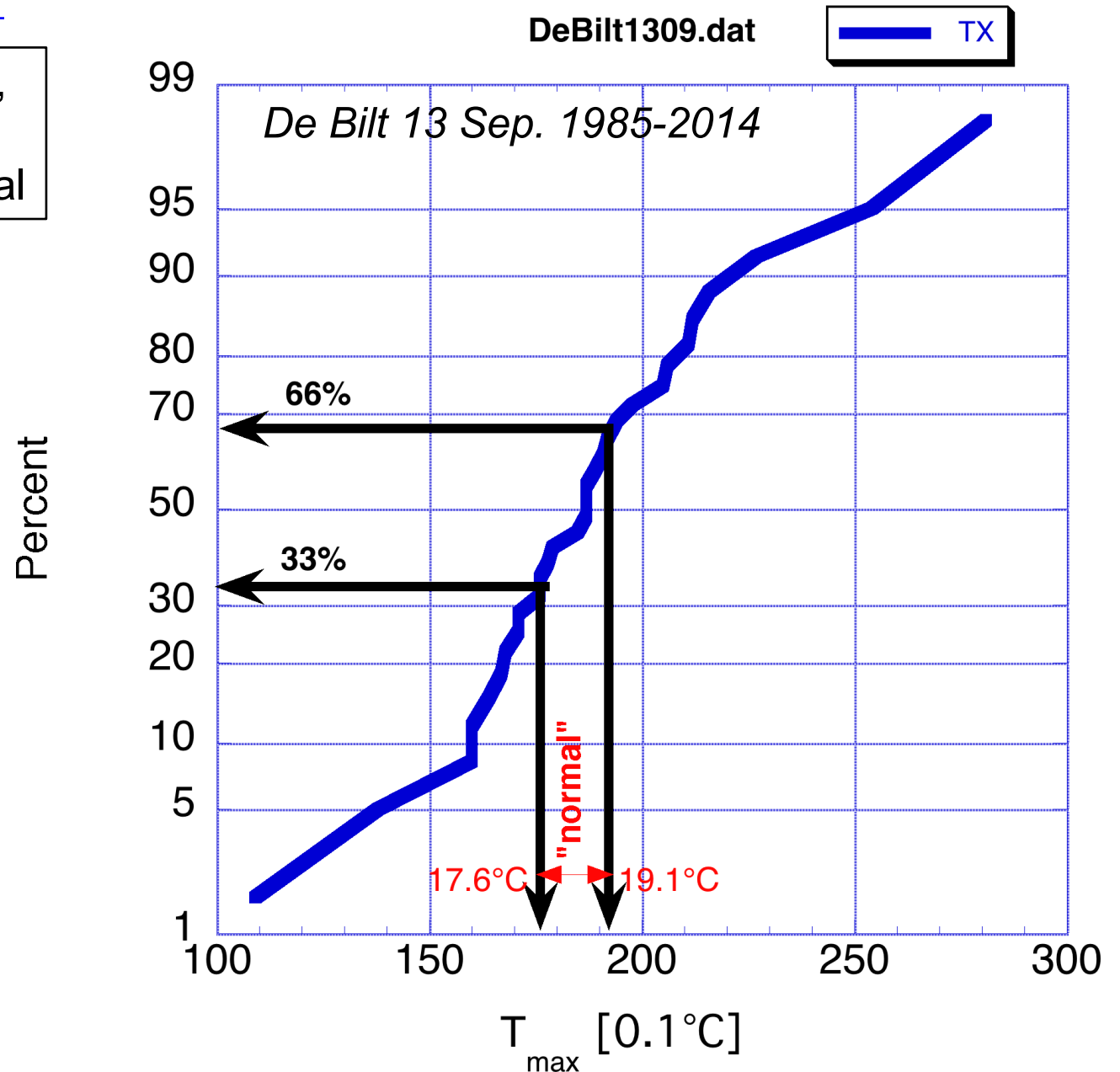
- Temperature:** $T_{\min}$ and $T_{\max}$ (units: $[0.1^{\circ}\text{C}]$)
- Solar Radiation:** sun (units: $[\%]$)
- Precipitation:** rain (units: $[0.1\text{ mm}]$)

STN,YYYYMMDD,DDVEC,FHVEC,FG,FHX,FHXH,FHN,FHNH,FXX,FXXH,TG,TN,TNH,TX,TXH,T10N,T10NH,SQ,SP,Q,DR,RH,
260,19850913,322,15,26,57,10,5,5,103,10,141,71,24,176,11,26,23,18,652,4,1,
260,19860913,101,26,31,46,10,15,3,87,11,114,46,6,167,14,33,75,58,1478,62,60,
260,19870913,228,15,21,57,14,5,3,118,15,158,140,3,187,15,137,0,0,227,93,79,
260,19880913,296,31,36,51,1,15,23,123,1,128,103,7,168,15,99,59,46,1258,36,77,
260,19890913,220,15,26,41,14,15,2,87,16,156,144,24,179,12,139,0,0,382,77,136,
260,19900913,38,10,10,21,12,0,5,51,14,129,66,4,192,13,31,54,42,1173,0,0,
260,19910913,115,5,10,36,10,0,4,67,10,126,28,5,216,13,-14,117,91,1741,0,0,
260,19920913,198,21,26,51,23,10,3,98,23,137,83,5,187,13,50,11,9,896,40,22,
260,19930913,97,51,51,67,3,21,20,113,4,124,93,22,138,1,68,0,0,143,120,148,
260,19940913,218,36,41,67,1,5,19,113,1,144,111,24,189,13,79,24,72,56,1297,0,0,
260,19950913,41,10,15,31,15,5,5,57,15,142,93,4,198,14,65,6,20,16,830,0,-1,
260,19960913,332,44,45,70,13,30,1,140,14,125,86,2,160,14,70,6,38,30,1009,8,2,
260,19970913,254,39,43,70,12,20,1,160,19,120,80,24,160,11,48,24,91,71,1376,12,17,
260,19980913,210,18,32,50,12,10,2,100,14,83,48,24,109,14,20,24,0,0,268,73,118,
260,19990913,341,24,26,50,14,10,1,100,14,189,141,5,254,13,112,6,81,63,1377,0,0,
260,20000913,280,22,25,40,11,0,22,100,13,160,104,24,206,13,78,24,58,45,1250,0,0,
260,20010913,233,33,35,70,2,10,20,150,3,128,86,22,176,13,59,24,39,30,980,55,102,
260,20020913,32,25,25,40,11,10,3,80,12,165,95,4,227,14,63,6,117,91,1720,0,0,
260,20030913,37,11,13,30,15,0,1,60,12,122,45,5,205,14,9,6,106,82,1704,0,0,
260,20040913,200,48,48,70,12,40,2,140,12,151,120,23,194,12,110,24,18,14,619,30,32,
260,20050913,192,5,10,20,13,0,1,50,14,154,105,6,212,16,66,6,98,76,1554,0,0,
260,20060913,157,21,25,40,8,10,18,80,14,215,170,2,281,15,130,6,92,71,1521,0,0,
260,20070913,121,7,11,20,9,0,1,40,11,126,72,5,185,12,56,24,23,18,992,0,0,
260,20080913,50,24,28,50,15,0,5,80,14,129,87,24,164,16,66,24,31,24,724,18,9,
260,20090913,346,33,34,60,14,10,3,130,14,146,88,1,171,16,77,6,30,23,825,2,3,
260,20100913,237,31,33,60,22,10,4,120,24,149,78,5,191,14,46,6,53,41,1202,4,1,
260,20110913,240,55,56,70,10,30,23,150,22,155,116,23,187,15,105,24,74,57,1400,20,53,
260,20120913,311,6,22,40,23,10,1,80,24,132,83,3,171,14,57,6,25,19,931,12,4,
260,20130913,199,26,28,50,13,10,1,90,13,145,90,6,178,11,67,6,29,23,856,37,15,
260,20140913,27,34,35,50,9,20,21,100,9,170,140,1,211,12,130,24,74,57,1342,0,0,

“Normal” from cumulative distribution

Forecast will be “categorical”,
e.g. temperature: normal,
below normal or above normal

*Probability [in %] that the
maximum temperature on 13
September will have a value
less than or equal to the value
plotted on the horizontal axis.*



Categorical forecast

The categorical forecast of maximum temperature in De Bilt on 13 september consists of a statement : “normal (17.6-19.1°C)”, “below normal (<17.6°C)” or “above normal (>19.1°C)”.

Each category has 1/3 chance of occurrence (according to the climatology of the past 30 years) and therefore of being a correct forecast.

If the forecast is correct, this is counted as a “**hit**”, $a=a+1$.

If the forecast is incorrect, this is counted as a “**miss**”, $b=b+1$.

The ratio of correct forecasts to total number of forecasts, which is equal to $a/(a+b)$, should of course be larger than 1/3.

Data used to reconstruct climate past 30 years

<https://www.daggegevens.knmi.nl/klimatologie/daggegevens>

20 November ± 1 day, 1992-2021

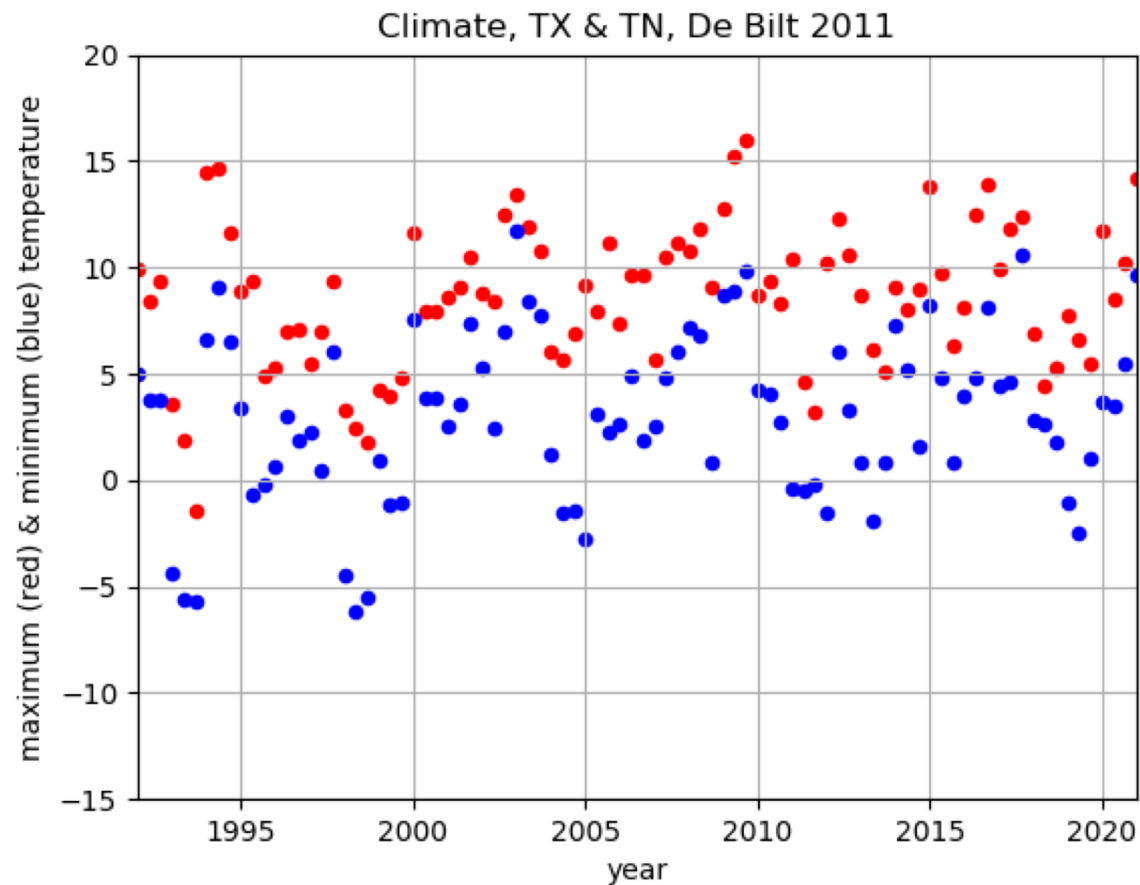
# STN	LON(east)	LAT(north)	ALT(m)	NAME
# 260	5.180	52.100	1.90	De Bilt
# TN	: Minimum temperatuur (in 0.1 graden Celsius) / Minimum temperature (in 0.1 degrees Celsius)			
# TX	: Maximum temperatuur (in 0.1 graden Celsius) / Maximum temperature (in 0.1 degrees Celsius)			
# SP	: Percentage van de langst mogelijke zonneshijnduur / Percentage of maximum potential sunshine duration			
# RH	: Etmaalsom van de neerslag (in 0.1 mm) (-1 voor <0.05 mm) / Daily precipitation amount (in 0.1 mm) (-1 for <0.05 mm)			
# STN,YYYYMMDD,	TN,	TX,	SP,	RH
260,19921119,	50,	99,	14,	13
260,19921120,	38,	84,	55,	-1
260,19921121,	38,	94,	39,	49
260,19931119,	-44,	36,	83,	0
260,19931120,	-56,	19,	2,	-1
260,19931121,	-57,	-14,	0,	-1
260,19941119,	66,	145,	0,	12
260,19941120,	91,	147,	3,	29
260,19941121,	65,	116,	4,	0
260,19951119,	34,	89,	12,	164
260,19951120,	-7,	94,	87,	0
260,19951121,	-2,	49,	84,	0
260,19961119,	6,	53,	25,	57
260,19961120,	30,	70,	0,	68
260,19961121,	19,	71,	27,	74
260,19971119,	23,	55,	76,	0
260,19971120,	5,	70,	0,	38
260,19971121,	60,	94,	8,	3
260,19981119,	-45,	33,	37,	0
260,19981120,	-62,	24,	78,	0
260,19981121,	-55,	18,	88,	0

30-year Climate in scatter plots

Range of "normal values" Sunday 20 November ± 1 day 1992-2021

TX: 7.5-10.2 °C

TN: 1.7-4.8 °C



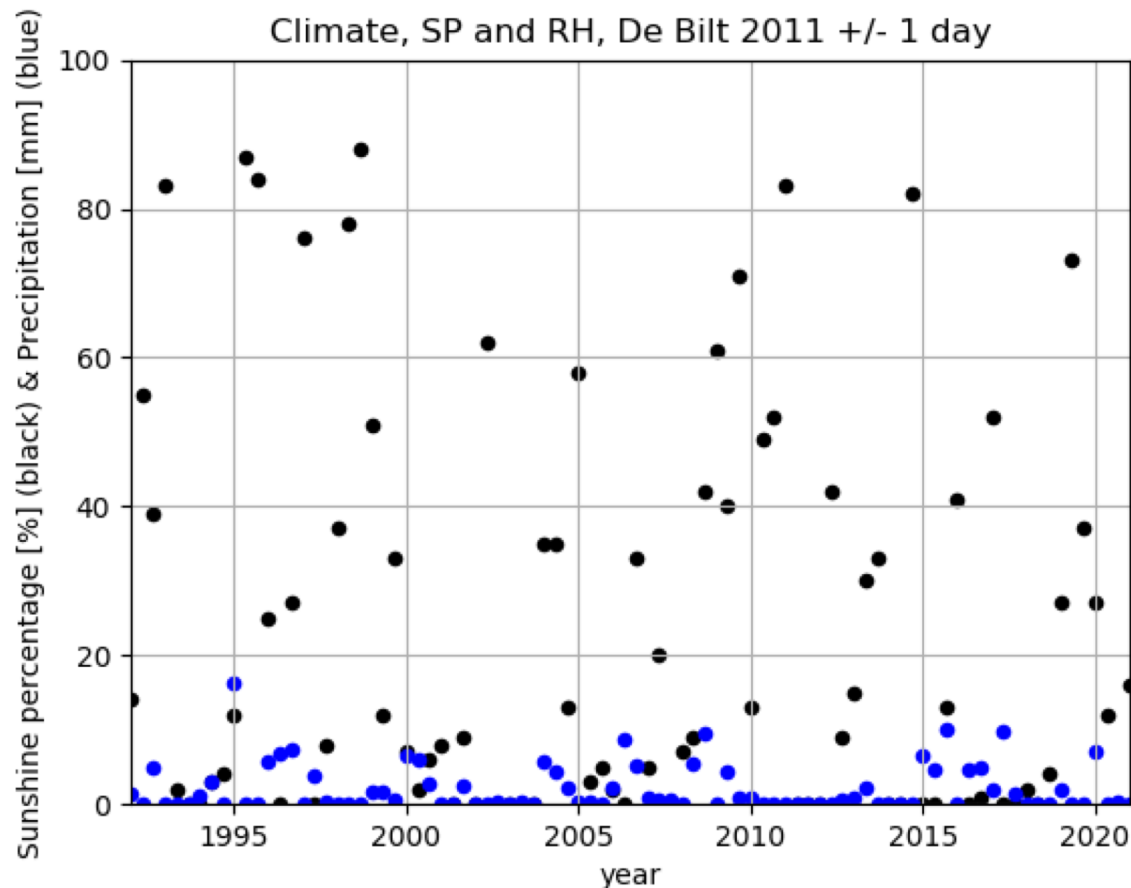
30-year Climate in scatter plots

Range of "normal values" Sunday 20 November ± 1 day 1992-2021

SP: 2.5-33.0 %

RH: 0.025-1.95 mm

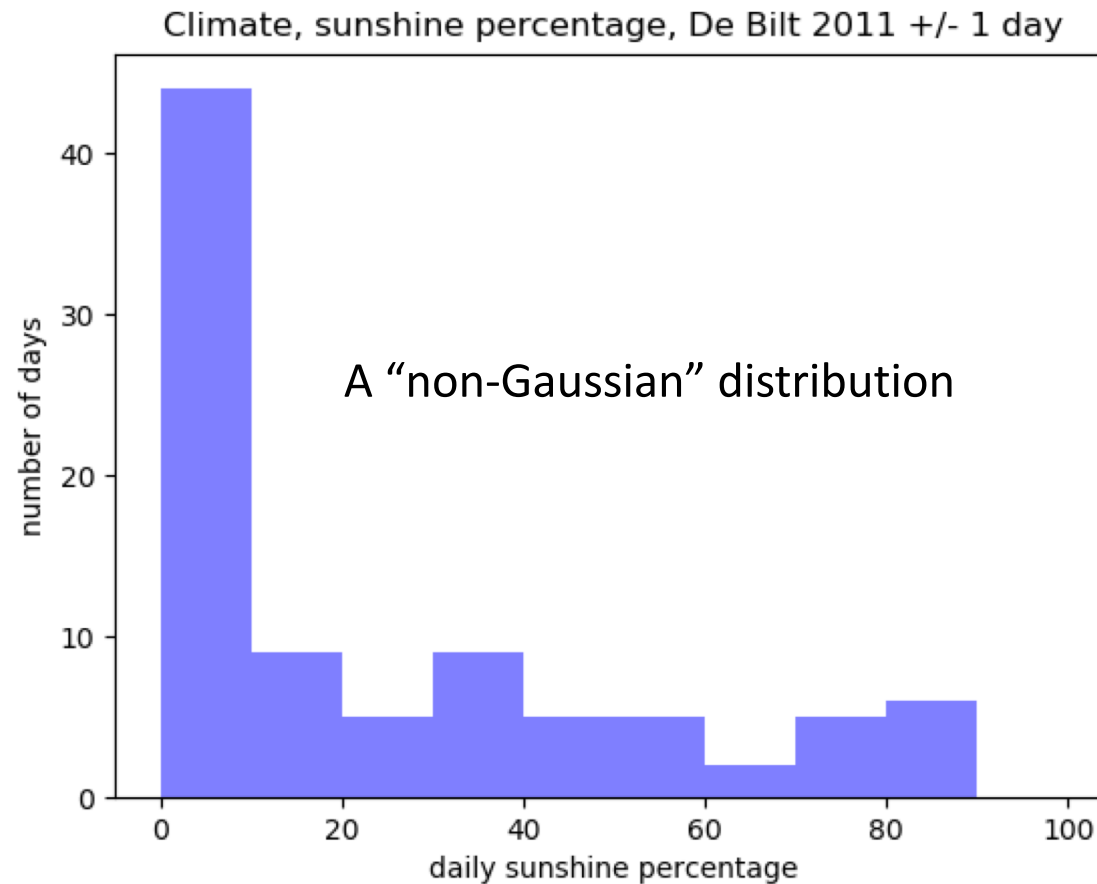
Average precipitation 2.1 mm !



30-year Climate in a histogram

Range of "normal values" Sunday 20 November ± 1 day 1992-2021

SP: 2.5 33.0 %

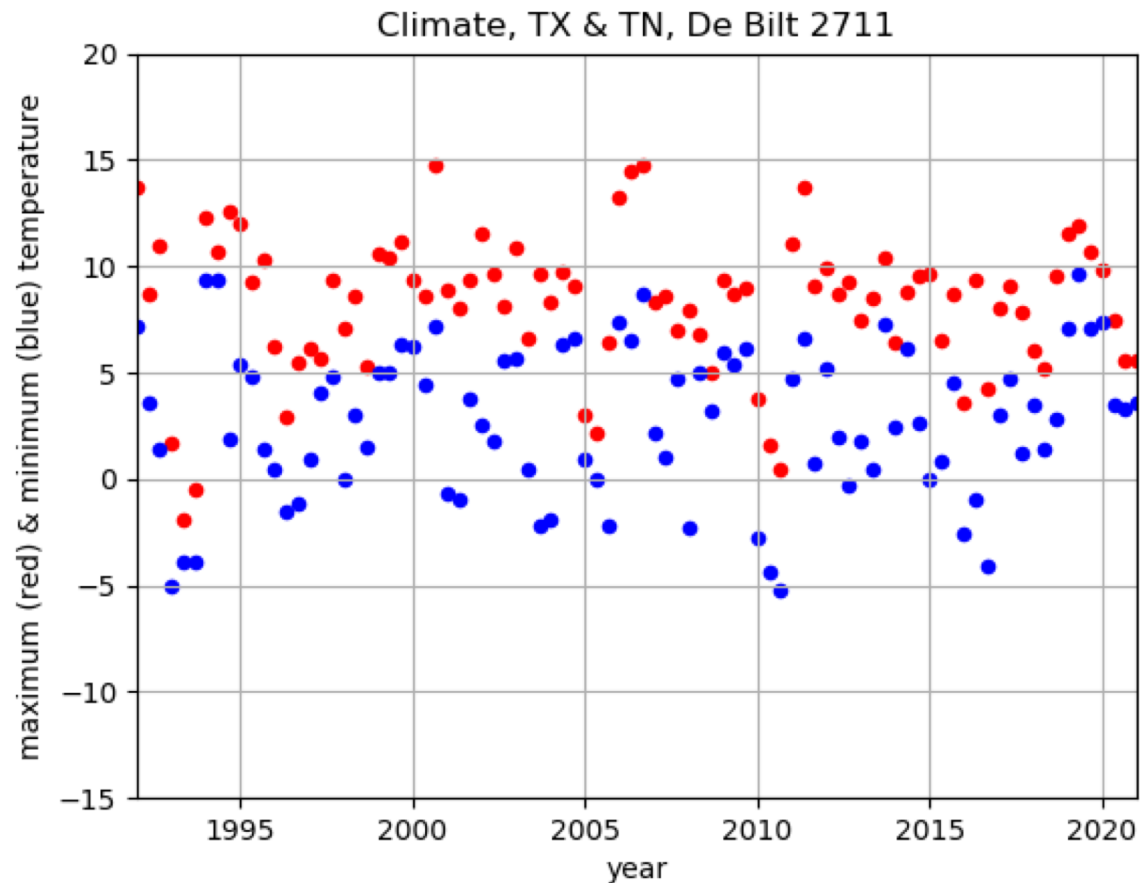


30-year Climate in scatter plots

Range of "normal values" Sunday 27 November ± 1 day 1992-2021

TX: 7.3-9.4 °C

TN: 1.1-4.7 °C



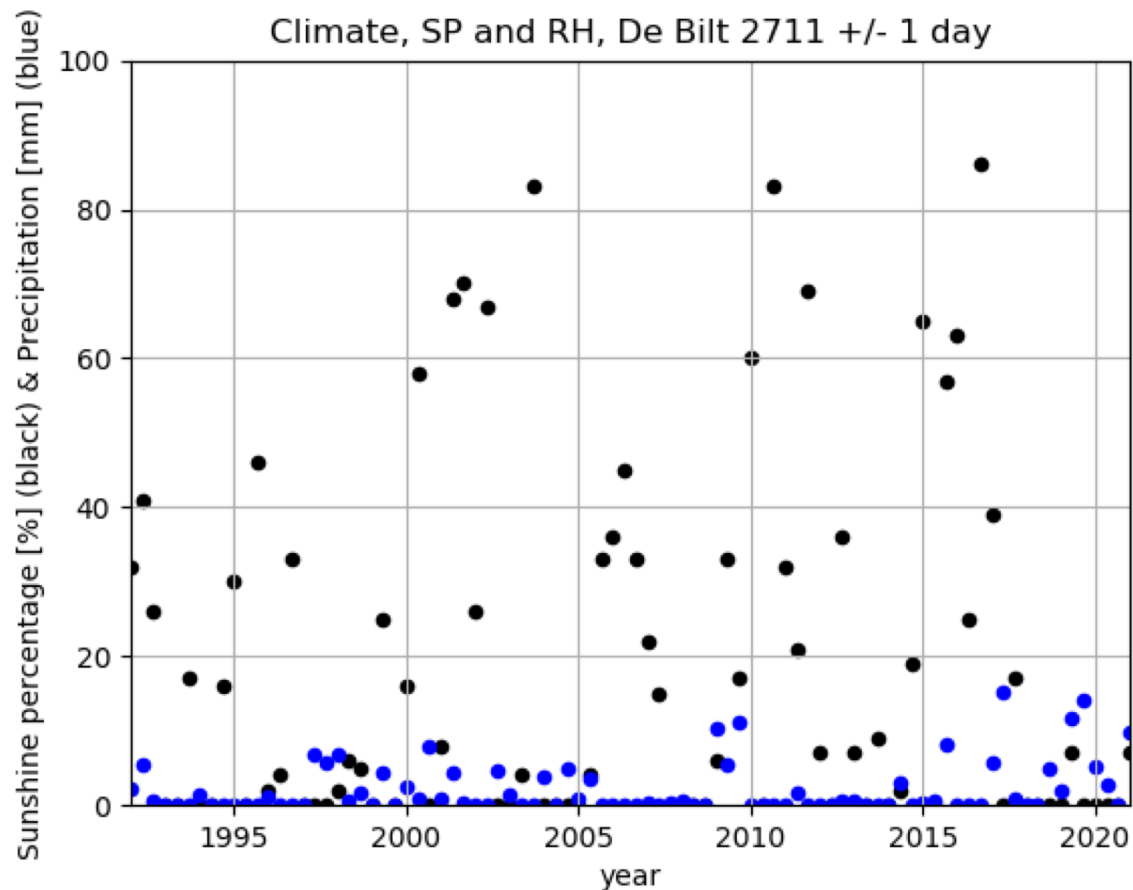
30-year Climate in scatter plots

Range of "normal values" Sunday 27 November ± 1 day 1992-2021

SP: 0.0-21.5 %

RH: 0.025-1.6 mm

Average precipitation 2.3 mm !



30-year Climate in a histogram

Range of "normal values" Sunday 27 November ± 1 day 1992-2021

SP: 0.0-21.5 %



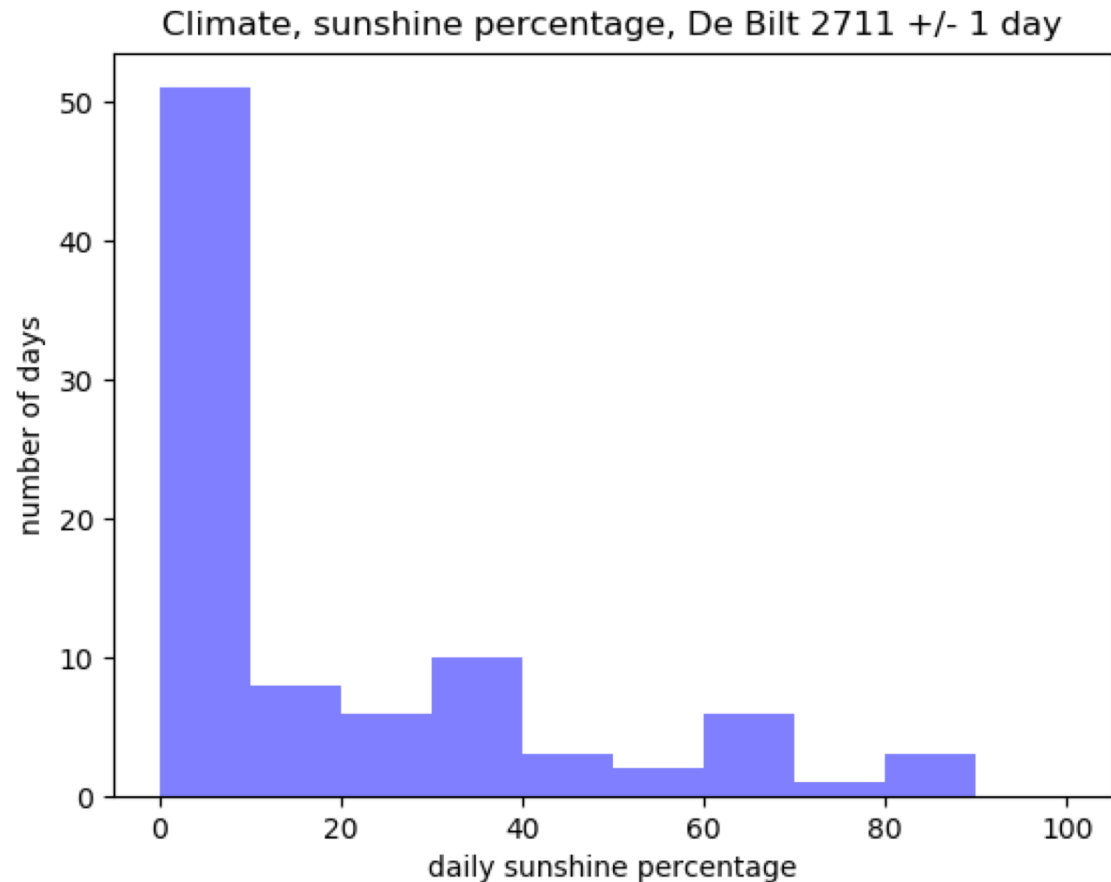
What is "below normal"?

Agreement is as follows:

Below normal (-): no sunshine

Normal (0): 1-21.5%

Above normal (+): 21.6-100%



Forecast for 20 November 2022

Normal:

TX: 7.5-10.2 °C

TN: 1.7-4.8 °C

SP: 2.5-33.0 %

RH: 0.025-1.95 mm

KNMI:

Zo

20-11-2022



Max.

5/9°

take the average (+7)

minus

Min.

-5/-1°

take the average (-3)

minus

Neerslag

3/6mm

take the average (4.5)

plus

Neerslagkans

90%

Zonneschijn

10%

zero

Windkracht

▲ Z 4

Weerplaza:

6

Zondag

20-11



☀ 10%

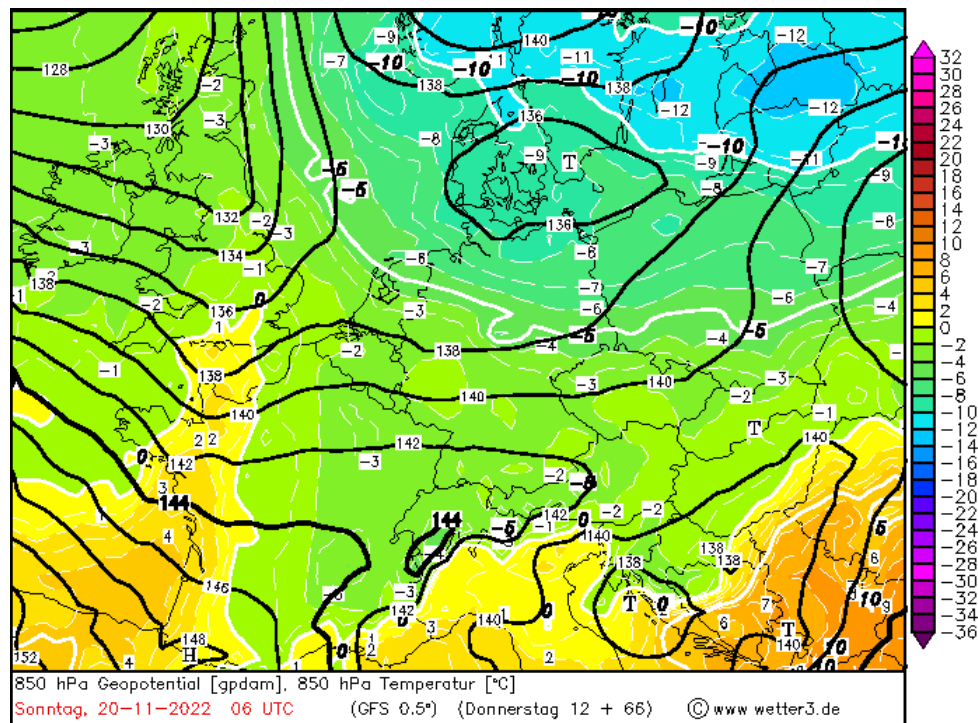
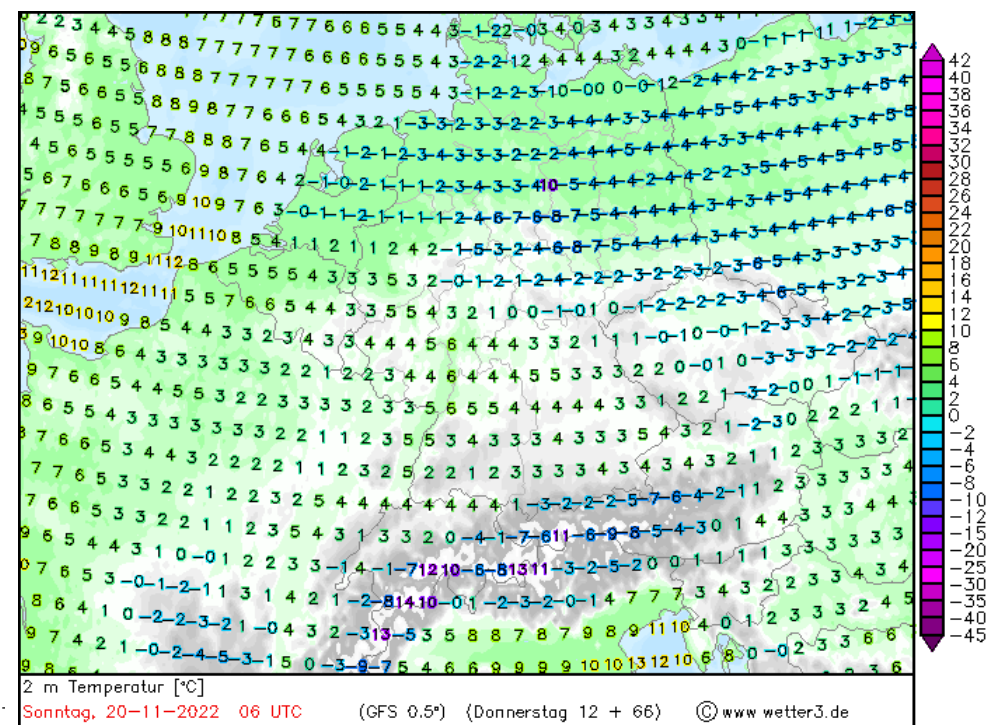
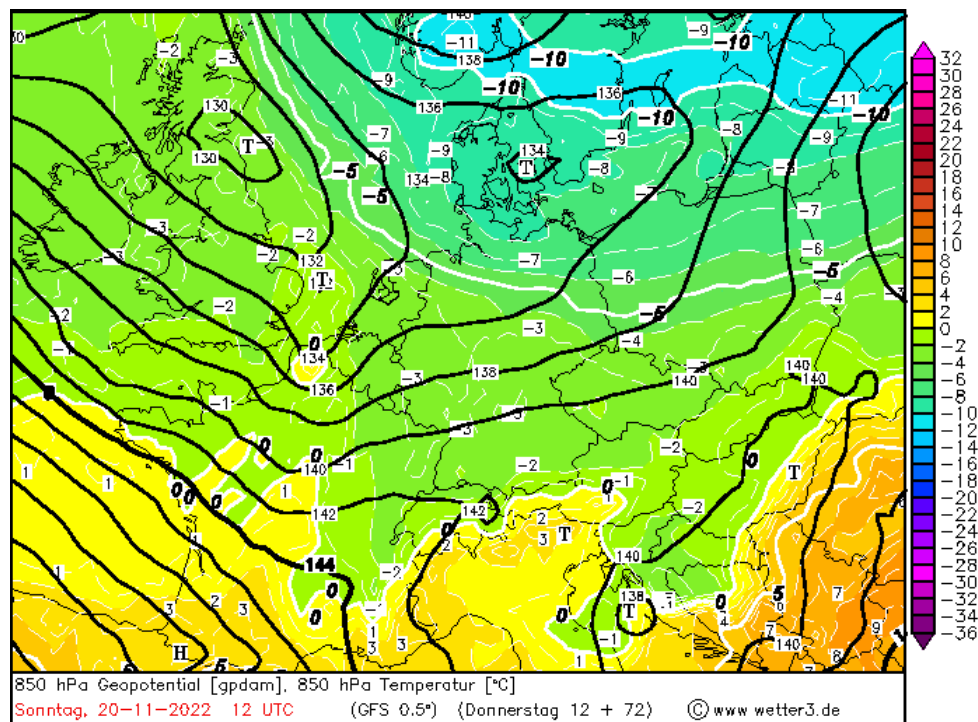
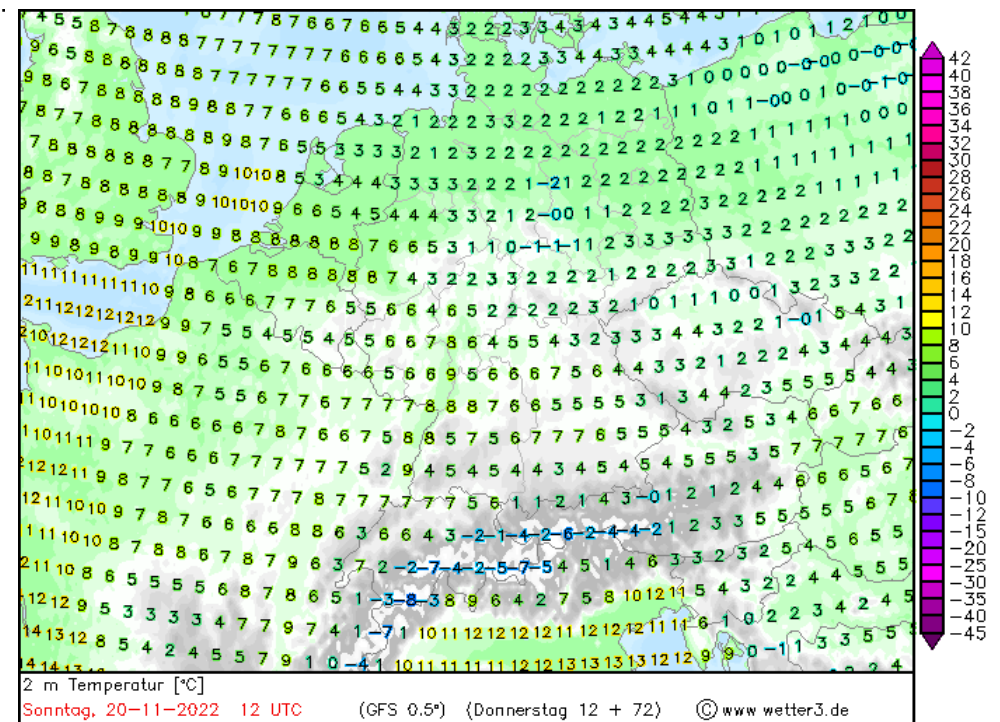
UV 0

6°C

-4°C

💧 95%

8 mm



List of forecast dates

27/11
4/12
11/12
18/12
25/12
15/1
22/1

Make teams of 4 to 5 people; give your team a name; make a collage of pictures of the team members, indicating who is who, and send in the collage as a pdf-file to a.j.vandelden@uu.nl before end of Monday 21 November.

page 12, lecture notes

PROBLEM 0.2 Making a weather forecast

Every week (on Wednesday) during the course we shall make a weather forecast for the following Sunday. The weather parameters of interest are daily maximum/minimum temperature, sunshine duration, as a percentage of the maximum possible, and precipitation. In order to simplify the evaluation of this forecast, we make the forecast categorical. A categorical forecast consists of a flat statement that can be either true (“hit”) or false (“miss”). For example a precipitation forecast, which states that there will be no rain during one day, can either be true or false. Likewise, a temperature forecast that states that the temperature will be normal can be true or false, if we define “normal” and “abnormal”. Here, we define a “normal temperature” as lying in the range $T_{\text{mean}} - \Delta T < T < T_{\text{mean}} + \Delta T$, where T_{mean} is the **mean** temperature on that particular day over the past, say, 30 years, while ΔT is equal to some fraction of the **standard deviation**, σ , of these temperature recordings. An “abnormal temperature” would then correspond to $T < T_{\text{mean}} - \Delta T$ or $T > T_{\text{mean}} + \Delta T$. By introducing the standard deviation, we assume that the probability density distribution of temperature recordings over the past 30 years is symmetric about the mean and is Gaussian. This is usually not a good approximation to the true probability distribution. It is better to deduce the range of “normal temperatures” from the **cumulative distribution function**, $P(T \leq x)$, which describes the probability that the temperature, T , is lower than, or equal to x . We divide the distribution of observed values of T into three “bins” (below normal, normal and above normal), where “normal” represents the range of values $T_1 < T < T_2$ such that $P(T \leq T_2) - P(T \leq T_1) = 1/3$, while $P(T < T_1) = 1/3$, so that, obviously $P(T > T_2) = 1/3$.