Methodology document for the WHO e-atlas of disaster risk.
Volume 1. Exposure to natural hazards
Version 2.0

Wind speed hazard modelling

Dr Zine El Abidine El Morjani

Taroudant poly-disciplinary faculty
of the Ibn Zohr University of Agadir, Morocco

Last Update: January 2011

Cataloguing-in-Publication Data

Methodology document for the WHO e-atlas of disaster risk. Volume 1. Exposure to natural hazards Version 2.0: Heat wave hazard modelling

1. Disasters – Wind Speed. 2. Geographic Information Systems. 3. Risk Management.

ISBN: 978-9954-0-5395-0

Acknowledgements

The development of the methodology for the preparation of the wind speed hazard distribution map is the product of contributions by several institutions and individuals. The funding to conduct the research leading to the development of the protocol has been provided by the World Health organization (WHO). The elaboration of the models was carried out by Zine El Abidine El Morjani, Taroudant poly-disciplinary faculty of the Ibn Zohr University of Agadir, Morocco and Steeve Ebener, WHO Mediterranean Centre for Health Risk Reduction, Tunisia. The data collection and analysis, implementation of the models, and documentation of the methodology were carried out by Zine El Abidine El Morjani.

Contents

Preface

Being able to conduct geographically based risk assessment at the sub national level requires being in a position to spatially distribute all the elements reported in following conceptual formula[1]:

$$\text{Risk} \propto \frac{\text{Hazard x Vulnerability}}{\text{Capacity}}$$

This process being very much driven by the type of hazard faced by the population and/or the key infrastructures in a given country the World Health Organization has been working, since 2006 on the development and improvement of an electronic atlas which could stimulate ministries of health and other health stakeholders to improve their disaster management capacity as well as serve as the entry point for conducting sub national geographically based risk assessments.

The WHO e-atlas of disaster risk models the distribution of natural hazards and population's exposure and provides baseline data and maps needed to advocate for resources to improve disaster preparedness; aid emergency response measures; and assist in identifying, planning and prioritizing areas for mitigation activities.

The first version of the e-atlas published in 2008 covered the WHO Eastern Mediterranean Region (22 countries) and five natural hazards (flood, seismic [earthquake], landslide, heat and wind speed) and was distributed to more than 500 users.

Encouraged by this success, working in close collaboration with the WHO Regions and taking advantage of the establishment of the Vulnerability and Risk Analysis and Mapping programme (VRAM), it was decided to publish a second version of the e-atlas that would, this time, also the 46 countries forming the WHO African Region as well as 32 countries of the WHO European Region (due to limited resources, this version of the e-atlas focuses on Central Europe only).

Building on the successful collaboration established between the Taroudant polydisciplinary faculty of Ibn Zohr University, Agadir, Morocco and the VRAM, most of the models used in the first version of the e-atlas have been improved and heat replaced by heat wave, a current preoccupation of many ministries of health.

In order to allow for any other region or country to also apply the models on their own it has been decided to document not only the research behind the models but also provide users with a protocol that would allow them to generate the final hazard distributions maps. The present series of methodology document is the result of this documentation.

---

[1] Modified from: Office of the United Nations Disaster Relief Co-ordinator (UNDRO). *Mitigating natural disasters. phenomena, effects and options. A manual for policy makers and planners*. New York, United Nations, 1991.

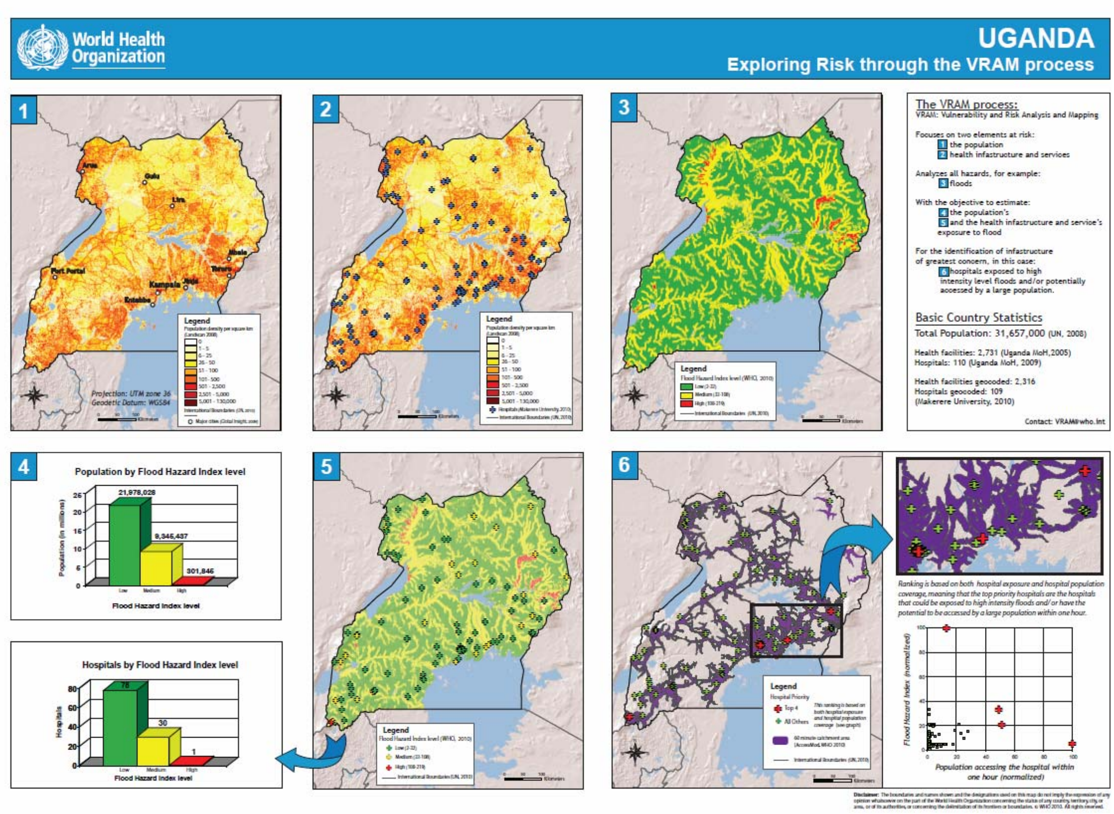
It is important to underline that the hazard distribution maps resulting from the application of these models are nevertheless only the first step of a process allowing countries to assess their risk at the sub national level.

Analysing vulnerability and capacity require a process which is difficult to be applied at the level of a region for the following reasons:

- availability of desegregated data
- incompatibility of indicators from one country to an other
- important differences in terms of health context between one country and another.

WHO has therefore been looking at having the vulnerability, capacity and therefore indirectly risk analysis, conducted on a country by country basis.

In this context, the VRAM is supporting Member States and partners to strengthen their capacity in order to conduct such analysis and have it presented in a manner such as the figure below.



The result of such analysis is then to be integrated in the country Disaster Risk Reduction (DRR) and Health Emergency Preparedness and Response Programmes (HEPRP) and serve, among other things, to build safer hospital, improve mass casualties' management and help specialized units within health Organizations (including MoH) for public health planning.

At the same time, the baseline data, information and maps collected or produced through the process can be used by health authorities and partners to take informed decisions in times of crises.

1. Introduction

This document describes the methodology and protocol developed by the Taroudant polydisciplinary faculty of Ibn Zohr University in close collaboration with the VRAM and then used to generate and document maps presenting the spatial distribution of wind speed hazard for the *WHO e-atlas of disaster risk, volume 1: exposure to natural hazards, Version 2.0.*

The methodology used to distribute wind speed hazard goes through two phases.

The first phase relies on frequency analysis techniques to estimate the annual maximum daily mean wind speed for different return periods at select weather stations in and around the three WHO Regions.

In the second phase, a stepwise regression analysis is conducted in order to define regression allowing to spatially distribute the intensity level of wind speed hazard for the e-atlas region mentioned above.

The methods and process presented in this document could be applied to other geographic areas provided that the analyses use geospatial data of similar or better quality and resolution.

2. Methodology

Spatial distribution estimates of meteorological data are increasingly important as inputs when modelling regional and global meteorological hazards. This atlas uses an objective mapping technique that employs empirical data and statistical procedures to estimate the spatial distribution of the annual maximum daily mean wind speed for a two, five and ten year return period as a measure of the wind speed hazard in the e-atlas region.

This technique goes through the application of the following steps (Figure 1).

1. Extraction of the daily mean wind speed data for the each weather station located in and around the region covered by this version of the e-atlas, from 1997 to 2008.

2. Estimation of the annual maximum daily mean wind speed for two, five and ten year return periods using Gumbel frequency analysis.

3. Identification of the relevant parameters for each climatic zone and return period, and selection of the regression model to spatialize the annual maximum daily mean wind speed using a stepwise regression analysis.

4. Interpolation of the annual maximum daily mean wind speed for each climatic zone and each return period using the selected regression models.

5. Aggregation and classification of the resulting maps in order to create the wind speed hazard distribution

Because of their respective climatological characteristics this process has been applied separately on five climatic zones (see section 2.3). The final map is created by aggregating the result obtained for each zone.

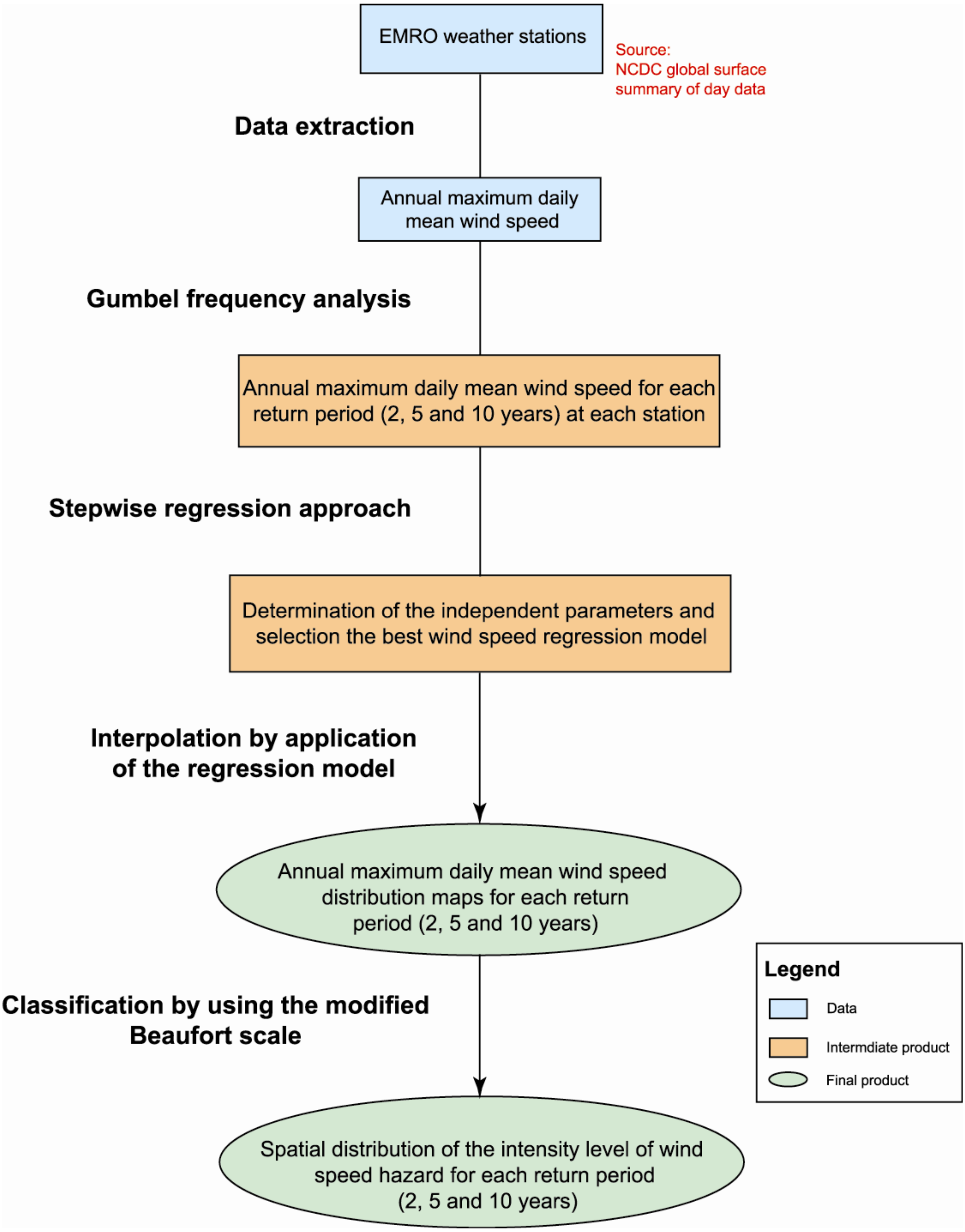


Figure 1. Methodology used for obtaining the spatial distribution of wind speed hazard over the region covered in this version of the e-atlas

2.1 Extraction and preparation of the daily wind speed data

The daily mean wind speed for the 1997-2008 period was mined for 3458 weather stations located within and around the countries of the study area from the Global Surface Summary of Day (GSSD) dataset produced by the US National Climatic Data Center (NCDC).

Accessible from the internet (http://www7.ncdc.noaa.gov/CDO/cdoselect.cmd?datasetabbv=GSOD&countryabbv=&georegionabbv= [Accessed December 15, 2010]), this dataset provides the following 18 surface meteorological elements for over 9000 stations:

- Mean temperature (.1 Fahrenheit)
- Mean dew point (.1 Fahrenheit)
- Mean sea level pressure (.1 mb)
- Mean station pressure (.1 mb)
- Mean visibility (.1 miles)
- Mean wind speed (.1 knots)
- Maximum sustained wind speed (.1 knots)
- Maximum wind gust (.1 knots)
- Minimum and Maximum temperature (.1 Fahrenheit)
- Precipitation amount (.01 inches)
- Snow depth (.1 inches)
- Indicator for occurrence of: fog, rain or drizzle, snow or ice pellets, hail, thunder and tornado/Funnel Cloud

Historical data are generally available for 1929 to the present, but the period 1973-present is the most complete.

It is important to note that no data were available for Afghanistan, Angola, Burundi, Democratic Republic of the Congo, Djibouti, Eritrea, Ethiopia, Iraq, Islamic Republic of Iran, Malawi, Rwanda and Somalia.

2.2 Estimation of the annual maximum daily mean wind speed for future predictions

Since meteorological data vary over space and time, it is necessary to analyse data from past periods and numerous locations to fully interpret past trends. Identifying historical patterns allows for the estimation of future meteorological trends. The estimation of future meteorological trends is based on the use of a probability distribution function as directed by several authors [Fuller, 1914; Foster, 1935; Gumbel, 1941; Gumbel, 1942; Kite, 1977; Stolte and Dumontier, 1977; Gerard and Karpuk, 1979; Condie and Lee, 1982; Moin and Shaw, 1985; Stedinger et al., 1992; USACE, 1993; and Jones et al., 2005]. A probability distribution function yields expected meteorological conditions over various time periods (return periods) in the future.

This approach does not require a comprehensive understanding of meteorology or meteorological phenomena but examines the relationship between the past magnitude and frequency of occurrence of the phenomena in order to identify some statistical regularity between them. In effect, the past is extrapolated into the future.

Frequently used probability distribution functions include: Gumbel, lognormal, Pearson type 3, log Pearson type 3, and the gamma distribution. Despite the extensive literature on the topic, there is no preferred distribution function for the frequency analysis of meteorological data because each function has a unique set of advantages and disadvantages. The problem is complicated by the necessity to evaluate meteorological data for return periods that exceed the length of the observed record.

Because of their respective climatologic characteristics, this process has been applied separately on five zones (see section 2.3). The final map was created by aggregating, in a seamless way, the results obtained for each zone

In our context, Gumbel's extreme value distribution function (Gumbel, 1941; Gumbel, 1954; Gumbel, 1960; Landwehr et al., 1979; Vogel, 1986; Sarma, 1999; El Morjani, 2003; Koutsoyiannis, 2004, He et al., 2006, El Morjani et al;, 2007) was the most appropriate function because it seeks to identify the temporal distribution of extreme values for the various return periods.

Additionally, this probability distribution function (equation 1) is one of the most widely used for extreme value prediction when analysing hydrological and meteorological data [Meylan and Musy, 1998]:

$$F(x) = e^{-e^{-\frac{x-a}{b}}} \qquad \text{Equation 1}$$

with: $F(x)$= cumulative distribution function
a and b = adjustment parameters; a is a location parameter and b is a scale parameter.

Replacing $\frac{x-a}{b}$ with the reduced variate u, the cumulative distribution function becomes:

$$F(x) = e^{-e^{-u}} \rightarrow u = -\ln[-\ln F(x)] = -\ln\left[-\ln\left(1-\frac{1}{T}\right)\right] \qquad \text{Equation 2}$$

With $F(X) = 1 - \frac{1}{T}$ and T = the return period.

Taking the station of Al Hoceima (Morocco) as an example (Table 1), the extreme value series for the annual maximum daily mean wind speed is fitted to a Gumbel distribution through the following steps:

1. The annual maximum value for the daily mean wind speed series is ranked in increasing order.

2. The empirical frequency is computed for each value using the Hazen formula

$$F = \frac{r - 0.5}{n} \qquad \text{Equation 3}$$

with r is the rank for each value and n is number of the years of record.

3. The Gumbel reduced variate u is calculated by applying equation 2.
4. The mean and the standard deviation of the maximum wind speed and of the Gumbel reduced variate are then calculated.
5. A graph is then generate to verify if the plotting of the annual maximum daily mean wind speed versus the Gumbel reduced variate u shows a form in conformity with the Gumbel distribution, and the results are linear. If not, this graph allows the identification of outliers, which can be removed from the dataset. In the case of the Kasba-Tadla weather station, the graph shows an adequate fit to the Gumbel distribution and the absence of any outliers (Figure 2).

Table 1. Ranking of the annual maximum daily mean wind speed, calculations of empirical frequency, Gumbel reduced variate and associated statistical parameters for the weather station of Al Hoceima (Morocco)

Year	Annual maximum daily mean wind speed (m/s)	Rank	Empirical frequency $F(x) = (r - 0.5)/n$	$u = -\ln[-\ln F(x)]$
2006	5.3	1	0.04	–1.16
2000	5.66	2	0.13	–0.73
1998	5.97	3	0.21	–0.45
2005	6.02	4	0.29	–0.21
2007	6.33	5	0.38	0.02
1999	6.48	6	0.46	0.25
2002	6.74	7	0.54	0.49
2003	6.75	8.5	0.67	0.90
1997	7.15	8.5	0.67	0.90
2008	7.41	10	0.79	1.45
2001	7.92	11	0.88	2.01
2004	8.8	12	0.96	3.16
Mean	**6.7**			**0.55**
Standard deviation	**0.99**			**1.23**
a	**6.26**			
b	**0.81**			

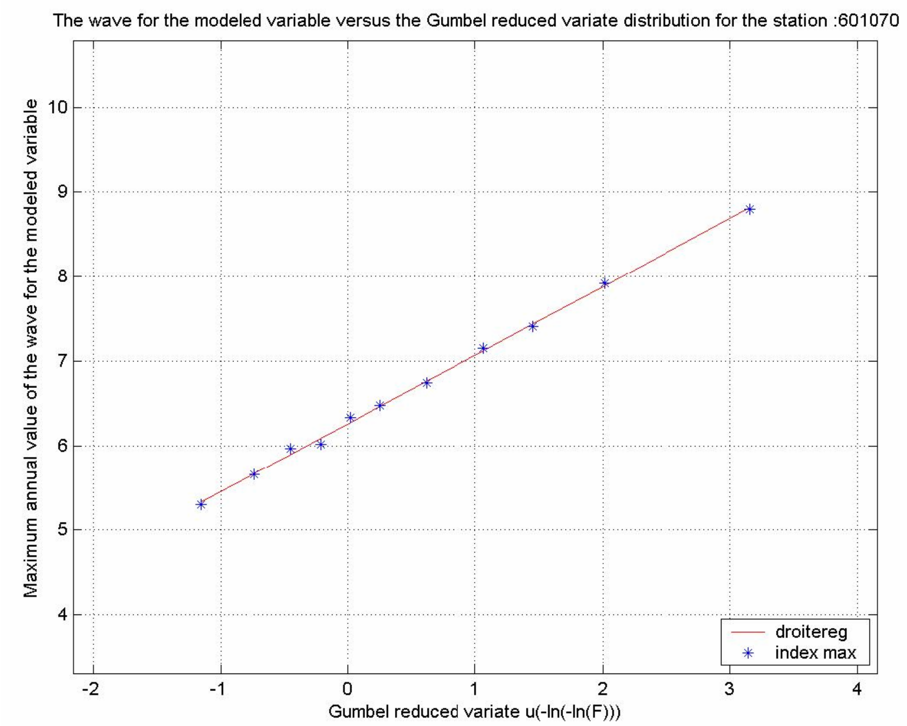


Figure 2. The annual maximum daily mean wind speed versus the Gumbel reduced variate distribution for the weather station at Al Hoceima (Morocco)

6. The statistical adjustments parameters of the distribution *a* and *b* are determined using the least-rectangles method (Table 1):

$$b = \frac{S_x}{S_u} \qquad \text{Equation 4}$$

$$a = \bar{x} - b\bar{u} \qquad \text{Equation 5}$$

with: S_x: standard deviation of the annual wind speed
S_u: Gumbel reduced standard deviation
$\bar{x}$: mean of the of the annual wind speed
$\bar{u}$: Gumbel reduced mean.

Note: the adjustment parameters will remain the same for each return period.

7. the annual maximum daily mean wind speed value for the desired return period *T* (two, five and ten years) are calculated using the following statistical model (Table 2):

$$X_T = a + bu = a + b(-\ln\left[-\ln\left(1-\frac{1}{T}\right)\right] \qquad \text{Equation 6}$$

with X_T = value of variate with a return period *T*.

The computing process is divided into the following three steps, using a five year return period as an example:

a. Calculation of the non-exceedance probability $F(x) = 1 - \frac{1}{T}$. For T = 5 years, $F(x)$ = 0.8.

b. Calculation of the Gumbel reduced variate using equation 2, u = 1.5.

c. Computation of the annual maximum daily mean wind sped using the linear relation reported in equation 6 (a and b are determined using equation 4 and equation 5: a = 6.26 and b = 0.81 → $X_5 = 6.26 + (0.81 \times 1.5) = 7.48$).

Table 1. Annual maximum daily mean wind speed for the weather station of Al Hoceima (Morocco) for two, five and ten year return periods

Return period (years)	Adjustment parameters		Annual maximum daily mean wind speed (m/s)
	a	***b***	
2	6.26	0.81	**6.56**
5	6.26	0.81	**7.48**
10	6.26	0.81	**8.08**

The process described above has been implemented in the **EatlasClimMod 1.0 application** to generate the annual maximum daily mean wind speed for different return period (two, five and ten years). This application has been developed under Matlab software, and is being converted to standalone application developed in C++.

2.3 Identification of the independent variables and selection of the regression model

Spatial interpolation is widely used for translating irregular scattered meteorological data (data collected at discrete locations [i.e. at points]) into continuous data surfaces (rasters).

The choice of interpolation method is especially important in the WHO Regions where meteorological data are sparse, and there are large value changes over short spatial distances. Additionally, the spatial density, distribution and spatial variability of sampling stations influence the choice of interpolation technique [MacEachren and Davidson, 1987].

Given a set of meteorological data, researchers are confronted with a variety of stochastic and deterministic spatial interpolation methods to estimate meteorological data values at unsampled locations:

- *deterministic* estimation methods including *inverse distance weighting* [Legates and Willmont, 1990, Eischeid et al., 1995; Lennon and Turner, 1995; Willmott and Matsuura, 1995; Collins and Bolstad, 1996; Ashraf et al., 1997; Dodson and Marks, 1997] and *spline methods* [Eckstein, 1989; Hutchinson and Gessler, 1994; Hulme et al., 1995; Lennon and Turner, 1995; Collins and Bolstad, 1996]

- *stochastic techniques* including the *kriging* and *cokriging* techniques [Matheron, 1963; Hudson and Wackernagel, 1994; Collins and Bolstad, 1996; Hammond and Yarie, 1996; Holdaway, 1996; Ashraf et al., 1997, El Morjani, 2003] and *polynomial regression* [Myers, 1990; Collins and Bolstad, 1996; Benzi et al., 1997; Chessa and Delitala, 1997; Hargy, 1997; Vogt et al., 1997; Agnew and Palutikof, 2000; El Morjani, 2003; Li et al., 2006].

For a summary description of these methods, refer to Collins and Bolstad [1996] and El Morjani [2003].

The characteristics of the data found for the e-atlas region (low spatial data density, a high spatial variability and the absence of meteorological data for many countries) resulted in implausible outputs when applying the inverse distance weighted and kriging interpolation methods, more specifically as follows.

The application of the inverse distance weighting method over a test area (Islamic Republic of Iran and Pakistan) generated specking or “bird’s eye” effects around the station locations, which was not plausible as the spatial variation for the annual maximum daily mean wind speed did not follow a regular trend.

The application of the kriging technique over the same test area produced results that were inconsistent with the original data. Whatever the model used (spherical, exponential or Gaussian) the statistical cross-validation was not able to fit the theoretical spatial semivariogram. This might be because the density of weather stations is too low and the study area too large to support the use of the kriging interpolation method.

It has therefore been necessary to find another model that produces results of good quality.

The same variables as the ones used for the spatialization of the heat wave index (see the Heat wave hazard modelling document which can be found on the WHO e-atlas of disaster risk, Volume 1. Exposure to natural hazards Version 2.0) have been used for wind speed, namely:

- *topographical parameters* including elevation (Z); the mean elevation within a 3×3 pixel window (Z9); aspect (Asp) and slope (Slp);

- *geographical factors* including relative longitude (d_X) and latitude (d_Y); and the distance to the nearest coastline (d_Coast).

With the variables identified, to which their squares (i.e. d_X^2, d_Coast^2, …) were also added, stepwise (back and forth steps) linear regression technique was used to identify their statistical significance, if any, and their relative contribution to the determination of the dependent variable (annual maximum daily mean wind speed), thereby eliminating any insignificant variables.

Because of their respective climatologic characteristics, this process has been applied separately on the following five zones:

- Zone 1: African countinent
- Zone 2: Middle East countries (Bahrain, Iraq, Israel, Jordan, Kuwait, Lebanon, Oman, Qatar, Saudi Arabia, Syrian Arab Republic, United Arab Emirates, West Bank and Gaza Strip, Yemen)
- Zone 3: Afghanistan, Islamic Republic of Iran, Pakistan
- Zone 4: Armenia, Azerbaijan, Georgia, Kazakhstan, Kyrgyzstan, Mongolia, Tajikistan, Turkmenistan and Uzbekistan
- Zone 5: Albania, Austria, Belarus, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Estonia, Greece, Hungary, Latvia, Lithuania, Moldova, Republic of, Montenegro, Poland, Romania, Serbia, Slovakia, Slovenia, former Yugoslav Republic of Macedonia, Turkey and Ukraine.

The final map was created by aggregating, in a seamless way, the results obtained for each zone.

A stepwise linear regression analysis was performed separately for each of these zones and return periods using S-Plus software. The validation of each regression was carried out using R^2 variance analysis as well as a detailed probability and residual analysis in order to identify the significant variables and therefore select the best regression model possible.

As an example, the stepwise regression analysis for a two year return period applied on the 121 weather stations located in zone 2 gives the results shown in Table 3.

Three of the seven independent variables explained most of the variation of the annual maximum daily mean wind speed for a two year return period (Wdsp_2): The three variables (elevation [Z] in metres, latitude [d_Y] and distance to the nearest coastline [d_Coast]) both expressed in kilometres (see section 3.4.1.5)) were used to derive the following regression equation for a 2 year return period:

$$\text{Wdsp\_2} = (0.002482 \times Z) + (0.006111 \times d\_Y) - (0.012497 \times d\_Coast) - (0.000001 \times Z^2) - (0.000001 \times d\_Y^2) + (0.000019 \times d\_Coast^2)$$

Equation 7

This equation shows a positive correlation between [Z], [d_Y] and the annual maximum daily mean wind speed, and an inverse correlation with [d_Coast]. This means that the annual maximum daily mean wind speed increases with elevation [Z] and latitude [d_Y] and decreases with distance to the nearest coastline [d_Coast].

Slope, longitude and aspect do not appear in the regression equation because they do not significantly explain the variation of annual maximum daily mean wind speed in zone 2.

Ninety-five percent of the variance in extreme wind speed are explained by the three variables retained in the regression equation ($R^2 = 0.95$). The model is considered valid and reliable because of the strong correlation ($R = 0.97$) and the high degree of confidence that exists in the selected variables (very small probability [F statistic]). This is clearly statistically significant, so the true slope of the regression line is probably not zero.

Table 3. Regression model for the annual maximum daily mean wind speed over zone 2 for a two year return period)

Variable	Regression coefficient	Standard error	*t* value	Probability Pr(>|*t*|)
Z	0.002482	0.001229	2.018447	0.005871
d_Y	0.006111	0.000365	16.7108	0.000000
d_Coast	–0.012497	0.004348	–2.873828	0.004831
Z^2	–0.000001	0.000000	–1.895487	0.006538
d_Y^2	–0.000001	0.000000	–9.343988	0.000000
d_$Coast^2$	0.000019	0.000007	2.751341	0.006897

Residual standard error	0.43629
Degrees of freedom	115
Multiple R^2	0.95
F statistic	132.94746
Probability (*F* statistic)	0.0000

2.4 Spatialization of the annual maximum daily mean wind speed over region covered in this version of the e-atlas

Once the regression model was found for each of the five zones and return periods, they were applied separately to each zone before aggregating the maps to form one unique layer covering all of the e-atlas region.

2.5 Classification of the annual maximum daily mean wind speed distribution maps into the final hazard maps

The final step in the methodology consisted in deriving the spatial distribution of the intensity level of wind speed hazard for each return period from the annual maximum daily mean wind speed distribution maps. This has been done by reclassifying these maps according to the five intensity levels selected for this project (very low, low, medium, high and very high).

This reclassification was done using the modified Beaufort classification for wind speed used by the US National Weather Service (http://www.srh.noaa.gov/mfl/hazards/info/beaufort.php; link checked 15 December 2010), reported in Table 4.

Table 4. Correspondence between the modified US National Weather Service classification and the five wind speed hazard intensity levels

Wind speed range (m/s)	Intensity level
<3.3	Very low
3.3–10.7	Low
10.7–17.1	Medium
17.1–24.4	High
>=24.4	Very high

3. Implementation

This section describes how the methodology presented in section 2 was implemented using the software listed in section 3.1.

The names of the files are reported in bold in the text. A “*” following the file name indicates a data layer which is described in the *Methodology and implementation process for generating the dataset,* document that can also be found on the e-atlas DVD.

3.1 Required software

The implementation of the methods and processes presented in this document requires the following software under Windows NT 4 (Service Pack 5, or 6a), Windows 2000 (Service Pack 3 or 4) and Windows XP (all versions) and any more recent version of Windows:

- ArcView 3.x with the Spatial Analyst 1.1 extension; both developed by the Environmental Systems Research Institute, Inc., for the geospatial operations.

- The following publicly available scripts and extensions which are accessible directly in the e-atlas DVD (in the tools section) have also been used :

  - Grid Analyst (GridAnalyst.avx)
  - Compiled Table Tools (Compiled_Table_Tools.avx)
  - XTools (Xtoolsmh.avx)
  - Grid and Theme Projector v.2 (grid_theme_prj.avx).

  These scripts should be saved on the computer in the C:\ESRI\AV_GIS30\ARCVIEW\EXT32 and then uploaded as extensions in ArcView before starting the process presented in the following sections.

- Matlab 6.0 (or higher), developed by MathWorks. Matlab has been used for the development of the EatlasClimMod 1.0 because software such as Excel can’t handle the large files used in the context of the present work (tables of approximately 160 000 lines) and there was a need for a platform for programming.

- The EatlasClimMod 1.0 application. This application has been developed to calculate and estimate the different climatic variables, including annual maximum daily mean wind speed for different return period, used in the context of this version of the WHO e-atlas. The codes of this application and the instruction file are all available in a zip file named EatlasClimMod.zip located in the tools section of the e-atlas DVD. This zip file contains the following functions which needs to be placed anywhere on your computer:

calcul_HeatIndex()	fct_CalculHeatIndexPerDay()	index_return()
calcul_HeatWave()	WaveModelling ()	save_liste_STNs()
delete_errorLine()	fct_MainProgram()	save_one_STN()
delete_null_line()	fct_Preprocessing()	check_succession()
fix_threshold()	Gumbel()	
fct_rank()	GUI_ HWI_Gumbel ()	

All these functions are accessed by the user through a graphic interface which has been developed under Matlab 6.0. This interface as well as the different functions will be presented in the next sections.

- A text editor such as WordPad to display the results tables.

- S-Plus 6.0, developed by Insightful Corporation, to explore and identify statistically significant parameters and their relative contribution to the spatialization of the wind speed using a stepwise multiple regression.

The minimum and recommended hardware requirements for running Matlab and the EatlasClimMod 1.0 application are as follows:

- Processor: Intel Pentium 3 and above
- 256 MB of RAM (512 MB or more is recommended)
- 600 MB of free hard drive space (1 GB is recommended)
- A colour graphics card and monitor (SVGA is recommended)

As a reference, EatlasClimMod 1.0 application has been used on the following computer configuration in the context of the WHO e-atlas project:

- processor Pentium Intel 4 (1.4 GHZ)
- 512 MB of RAM
- 2 Go of free hard drive space

It is therefore recommended, if possible to use a computer presenting these characteristics or better.

3.2 Extraction and preparation of the meteorological data

The source of the meteorological data used in this protocol is the Global Summary of Day (GSOD) dataset produced by the US National Climatic Data Center (NCDC) (see section 2.1).

The following steps describe how the daily mean wind speed has been extracted from this dataset to cover the 1997 - 2008 period:

1. Download and save the global summary of day data in ASCII format from: http://www7.ncdc.noaa.gov/CDO/cdo web site, after choosing the geographic region (Africa, Asia, Europe, Middle East) and the date range from 01/01/1997 to 31/12/2008.

2. Open each meteorological regional file in Microsoft Office Access and save them as a dBASE IV (***.dbf**) table using a specific naming convention (e.g. **africa.dbf**). These files will not open in Microsoft Office Excel because they contain more than 65 356 records. The fields names and their description can be found in Annex 1.

3. Open **africa.dbf**, and choose C-Tables Tools>Delete Multi-Fields and select all the fields to be deleted (all except: Field 1: STN, Field 3: YEARMODA and Field 14: WDSP). This operation keeps the columns that are needed for the analysis.

4. Set the missing day mean wind speed, having a dummy value of 999.9, to Null as follows:

    a. in **africa.dbf** add a field called "mnspd" in its attribute table by choosing the number type, in order to convert the WDSP into number typ
    b. select the header of the "mnspd" column and click on the Calculate button
    c. type the following formula in the Calculator window: [WDSP].AsNumber
    d. click on the Query Builder button, and type ([mnspd] = 999.9)
    e. Select Calculate from the Field menu
    f. Type the following expression Number.MakeNull and click OK.

5. Convert the mean wind speed from knots to m/s using the following steps:

    a. Open **africa.dbf** and add a new field "mnspd_m"

    b. select the header of the "mnspd_m" column, click on the Calculate button and type the following formula: ([mnspd]*0.5144) in the Calculator window.

6. Compute the annual maximum daily mean wind speed for each weather station and for each year of observation as follows:

    a. open the **africa.dbf** table and add three new fields by choosing C-Tables Tools>Create Multi-Field, naming them "stn_s", "yearmoda_s", "year", and "stn_year", where:
        - "stn_s": station number in string type

- "yearmoda_s": date in string type containing year, month and day
- "year": only year in string type
- "stn_year": concatenate station number and year

b. select the header of the "stn_s" column, click on the Calculate button and type the following formula: [stn].AsString in the Calculator window to convert STN into string type. Redo the same operation for the "yearmoda_s" header using this formula: [yearmoda].AsString

c. select the header of the "year" column, click on the Calculate button and type the following formula in the calculator window: [yearmoda_s].Left(4). This field contains only the year of observation for each station

d. select the header of the "stn_year" column, click on the Calculate button and type the following formula in the calculator window: [stn_s]+[year] to concatenate these two fields into one field

e. select the header of the "stn_year" field, from the menu, choose Field>Summarize. In the dialog box that appears:

   - under Save As, change the name of the output file from default to **Africa_max_mnspd.dbf**
   - under Field, scroll down the list and click on "mnspd_m"
   - under Summarize By, choose "maximum"
   - click OK

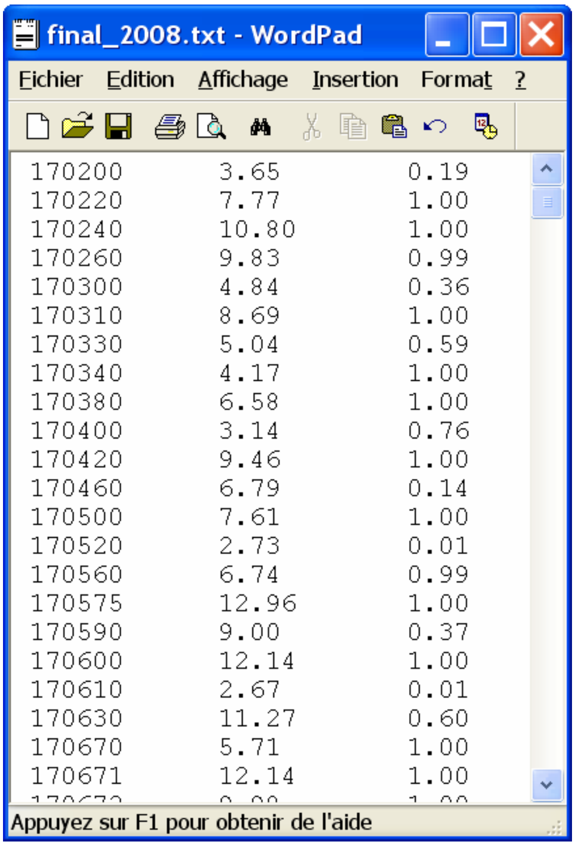
f. In this table note that the "max_mnspd _m" and "count" fields have been added automatically. Use the "count" field to create a annual frequency field "freq" by dividing by 365.

7. Generate the new files for each year of observation (1997–2008) from **Africa_max_mnspd.dbf**. These files are needed for the application of the Gumbel method (section 3.3) and are derived as follows:

   a. use the Query Builder tool and type ([year] = "1997") to select the year = 1997 and save the table under **final_year.txt** (year = 1997, 1998, … , 2008) by selecting File>Export, and Export Format = Delimited Text

   b. delete all the fields except "stn", "max_mnspd" and "freq"

   c. using a text editor such as WordPad, delete the header (first record containing "stn", "max_mnspd" and "freq" of these files) in order to have files with only numerical contents. This operation facilitates the application of the Gumbel method in Matlab. Figure 3 present one example of such file



170200	3.65	0.19
170220	7.77	1.00
170240	10.80	1.00
170260	9.83	0.99
170300	4.84	0.36
170310	8.69	1.00
170330	5.04	0.59
170340	4.17	1.00
170380	6.58	1.00
170400	3.14	0.76
170420	9.46	1.00
170460	6.79	0.14
170500	7.61	1.00
170520	2.73	0.01
170560	6.74	0.99
170575	12.96	1.00
170590	9.00	0.37
170600	12.14	1.00
170610	2.67	0.01
170630	11.27	0.60
170670	5.71	1.00
170671	12.14	1.00

Figure 3. Example of annual maximum daily mean wind speed file (final_2008.txt) (with column 1: Station, column 2: maximum mean wind speed, column 3: annual frequency)

8. Repeat step 7 for the other years of observation.

9. Put all the files created **final_year.txt** (year = 1997, 1998, … , 2008) into specific folders named Africa/data, Europe/data, etc.

10. Repeat steps 1 to 9 for the other regions.

3.3 Calculation of the annual maximum daily wind speed for different return periods

The application of the Gumbel frequency method to estimate the annual maximum daily mean wind speed for different return periods, takes place using the EatlasClimMod 1.0 application under Matlab 6.0 (or higher).

The following steps have to be followed in order to start the EatlasClimMod 1.0 application:

1. Run the MATLAB software. The window presented in Figure 4 will appear.

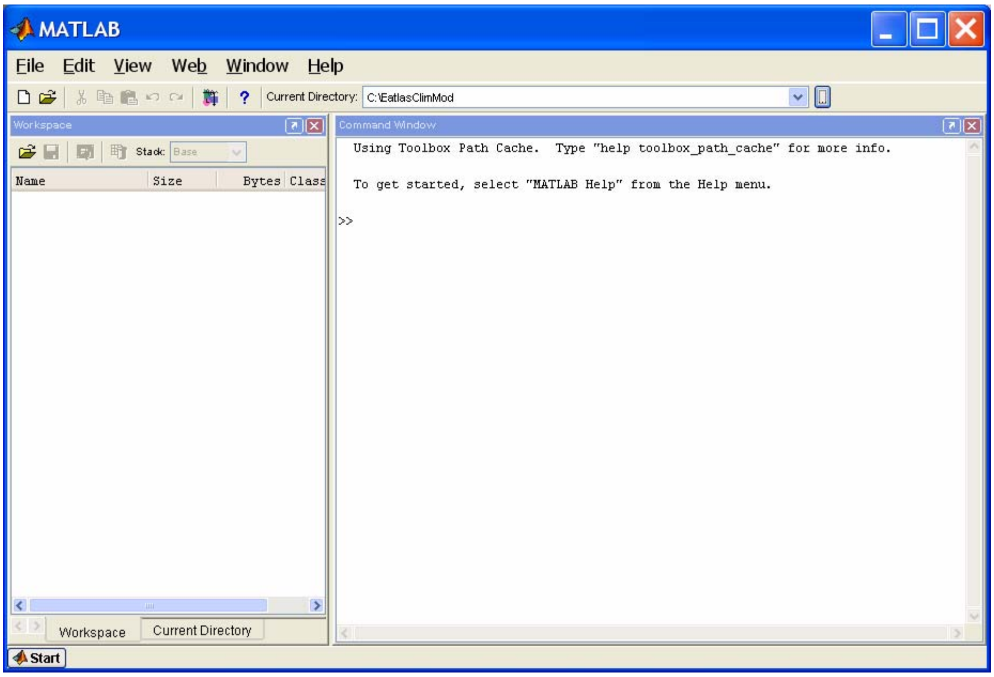


Figure 4. Matlab interface

2. Specify the path to the EatlasClimMod 1.0 application as the current directory using the browsing button  on the upper right side of the window.

3. Launch the EatlasClimMod 1.0 application, by going to the **File>Open**, select the GUI_HWI_Gumbel.m file as shown in Figure 5 and click on Open.

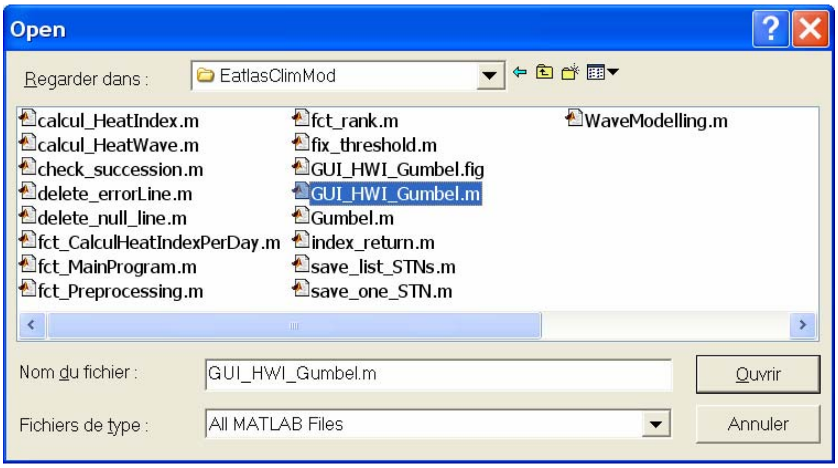


Figure 5. Open file window in MATLAB to run the EatlasClimMod 1.0 application

The EatlasClimMod main program will then be opened in Matlab as shown in Figure 6.

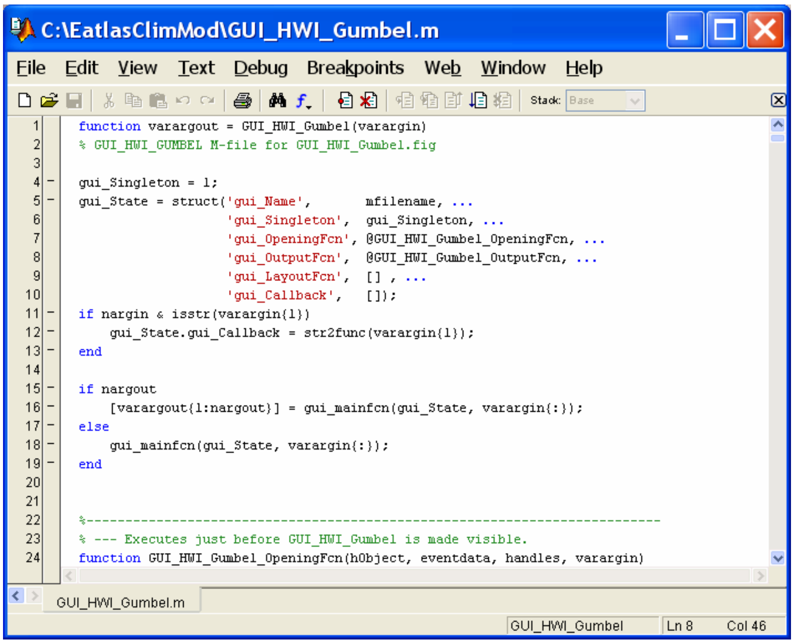


Figure 6. Window appearing in Matlab once the EatlasClimMod file has been opened

4. From there, press F5 or click on Run button  to run the **EatlasClimMod 1.0** application. The start up screen of this application will then appear as shown in Figure 7.

Figure 7. EatlasClimMod 1.0 application startup screen

This start up screen gives access to two menus: *Operation* and *About*.

The **About** menu give the user has access to the Help file or to the summary screen window (Figure 8)

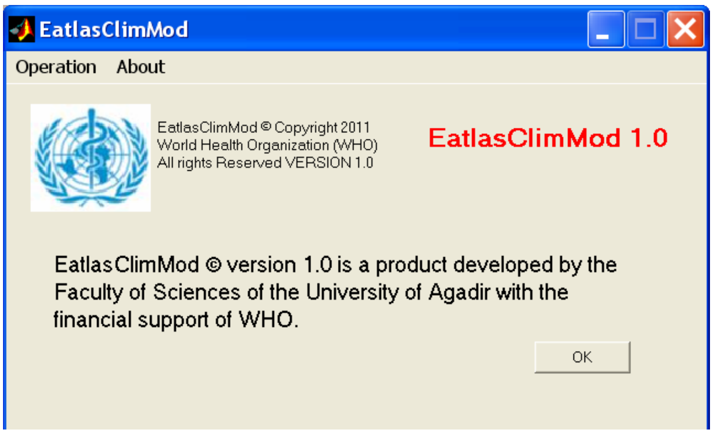


Figure 8. EatlasClimMod 1.0 © summary screen

The **Operation** menu gives access to five options:

- **Preprocessing**: used for data pre-processing;
- **Heat index**: used for the calculation of the daily heat index;
- **Wave modelling:** used to calculate the annual maximum wave for any variable over a given number of consecutive days
- **Unique stations files**: used to save the data of each weather station in a separated file;

➢ **Gumbel analysis**: used to predict the wind speed for different return periods; this options contains two sub-options:

- **All stations**: used to apply the Gumbel method on all weather stations
- **One station**: used to apply the Gumbel method on a single weather station

➢ **Exit**: used to close the application.

From the above, only the last two options (Unique stations files and Gumbel analysis**)** are needed to estimate the annual maximum daily mean wind speed for different return periods.

3.3.1 Creation of weather station specific files

Applying the Gumbel frequency method requires the creation of weather stations specific files containing the annual maximum daily mean wind speed obtained through the process presented in section 3.2.

This operation is carried out in the **EatlasClimMod 1.0** application using the following steps:

1. Click on **Operation>Unique stations files**. This leads to the window presented on Figure 9.

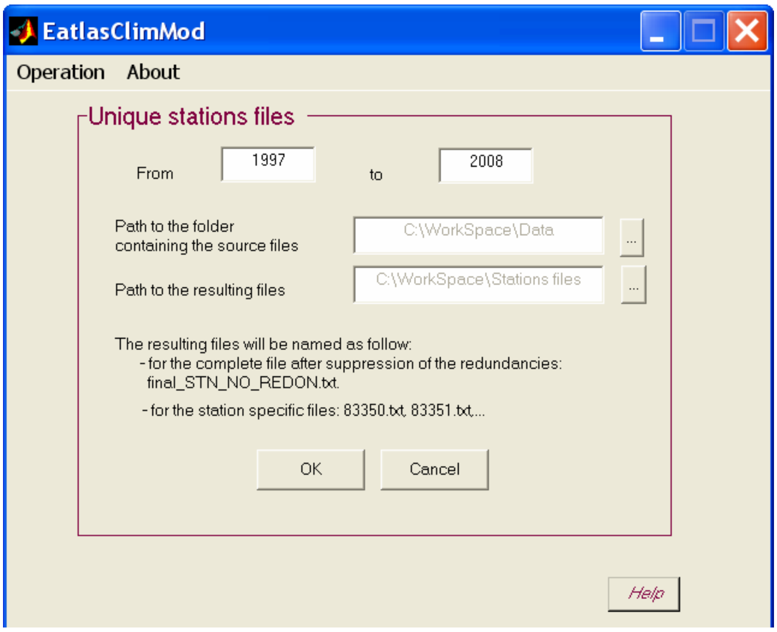


Figure 9. Window used to specify the parameters for creating the weather station specific files

In this window, the user must specify:

a. the period of consecutive years for which he wants the data to be extracted by weather station,
b. the path to the folder containing the files resulting from the calculation of the annual maximum daily mean wind speed (see section 3.2)
c. the path to the folder in which he wants the results to be saved

For b. and c. the path can be changed by clicking on the browse button

2. Once the parameters entered in this window, click on **Ok** to start the process. This will automatically run the two functions behind this process, namely: save_list_STNs() and save_one_STN().

The resulting files will be named as follow:

- ***final_STN_NO_REDON.txt*** for the file containing all the source data sorted by weather station and year and with all the redundancies deleted
- ***station_number.txt*** (for example, **400010.txt**) for the weather station specific files (see Figure 10 for an example of such file).

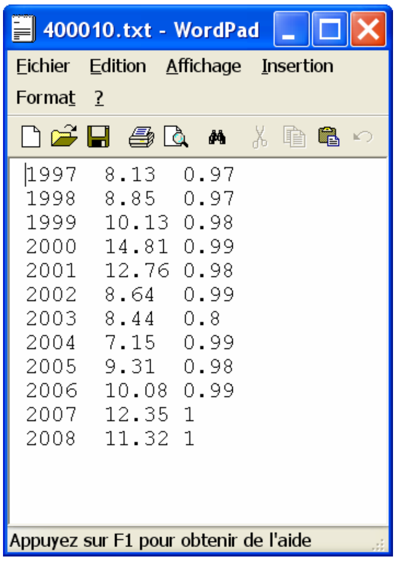


Figure 10. Example of weather station specific file with column 1: year, column 2: annual maximum daily mean wind speed, column 3: annual frequency

3.3.2 Application of the Gumbel frequency analysis

The Gumbel frequency analysis technique has been programmed and included in the **EatlasClimMod 1.0** application in order to predict the annual maximum daily mean wind speed for any given weather station and return period.

The use of this application requires the introduction of two thresholds which are used as filters to remove any weather station from the calculation in case these are not respected.

While the user can specify these thresholds manually in **EatlasClimMod 1.0**, they have been fixed as follow in the context of the WHO e-atlas. Namely, a weather station would not be taken into account if:

- the dataset for that given station does not contain a daily observation for at least 70% of the days in the year (255 days),
- the number of year of observation for that station, after applying the first filter, is lower than 8 years.

The threshold at eight years will give a good prediction of the annual maximum daily mean wind speed for return periods that do not exceed eight years.

In this work, we have nevertheless also calculated the annual maximum daily mean wind speed for a 10 year return period even if this result should be taken with precaution.

In **EatlasClimMod 1.0**, the Gumbel frequency analysis is applied in two steps:

- Application of the Gumbel frequency analysis on all the stations
- Correction and/or adjustment of the original dataset for unusual observations (typing mistakes, outliers,…).

The steps to be followed for these two steps are described in the coming sections.

3.3.2.1 Application of the Gumbel frequency analysis on all the weather stations

Here are the steps to be followed in order to apply the Gumbel frequency analysis on all the weather stations at the same time:

1. In EatlasClimMod 1.0, click on **Operation>Gumbel analysis**;

2. Choose the **All stations** option. This will open the window presented in Figure 11.

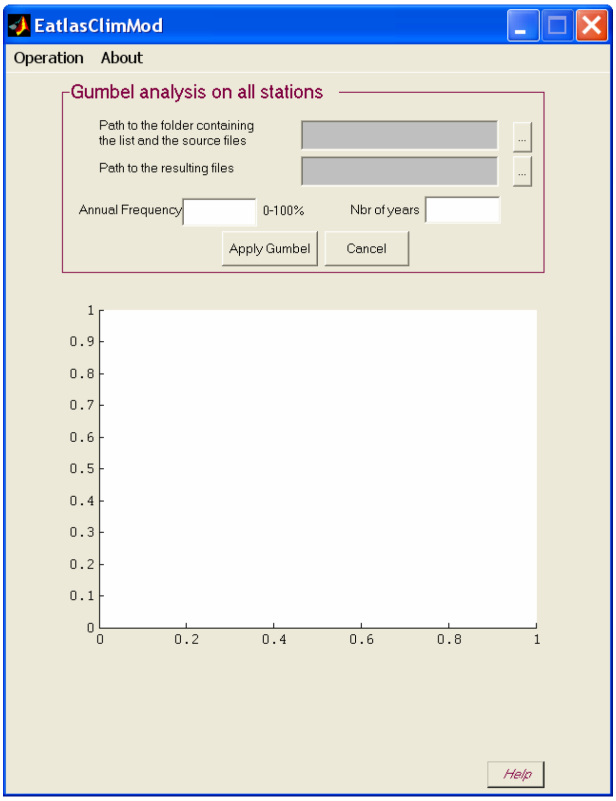


Figure 11. Interface window of the application of the Gumbel method to all stations

In this dialogue box, the user must specify:

a. the path to the folder containing the weather station specific files (see section 3.3.1),
b. the path to the folder where the resulting files will be saved
c. the two thresholds described in section 3.3.2 (annual frequency and minimum number of years of observations).

For a. and b. the path to the selected folder can be changed by clicking on the browse button 

3. Once all the parameters entered, click on the 

button to start the process, This will automatically run the Apply_Gumbel_allSTN() function which itself calls the following four functions:
    - fix_threshold(): used to remove the lines in the dataset for which the annual frequency is lower than the fixed threshold;
    - Gumbel(): produces the tables containing the values of the parameters involved in the Gumbel frequency method (Gumbel reduced variable, empirical frequency, mean and standard deviation)
    - fct_rank(): used to rank to the annual maximum daily mean wind speed (the rank will be the same in the case of similar values for different years)
    - index_return(): computes the annual maximum daily mean wind speed for two, five, eight and ten year return periods.

    In the window, the graph for each station is appearing one after the other once the analysis completed. The passage from a station graph to another one is done automatically. Figure 12 show one example of such window.

At the end of the treatment, the **EatlasClimMod 1.0** application will have produced a graph, plotting the annual maximum daily mean wind speed versus the Gumbel reduced variate, for each of the station and stored this graph as an image file, named ***numSTN*.jpg** (with *numSTN* = the station number, for example 85940.jpg, 600600.jpg, 400010.jpg), in the folder selected previously. Figure 13 present one example of such graph.

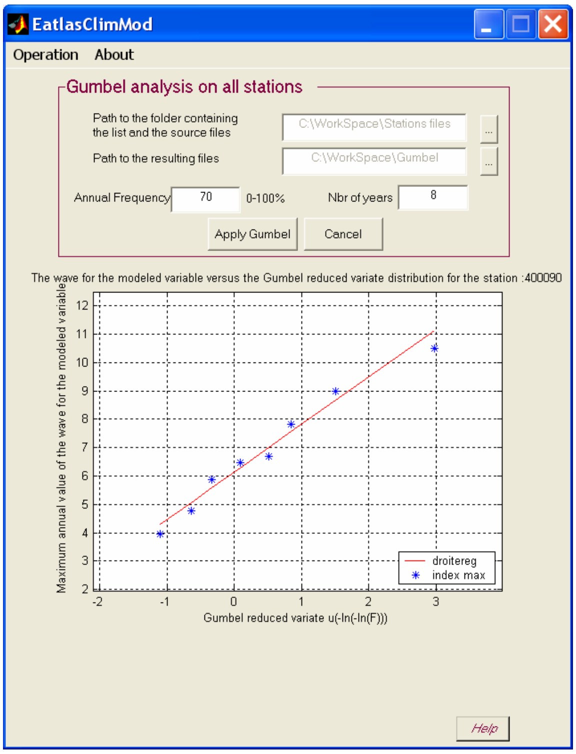


Figure 12. Example of window appearing as the gumbel analysis is completed on each weather station

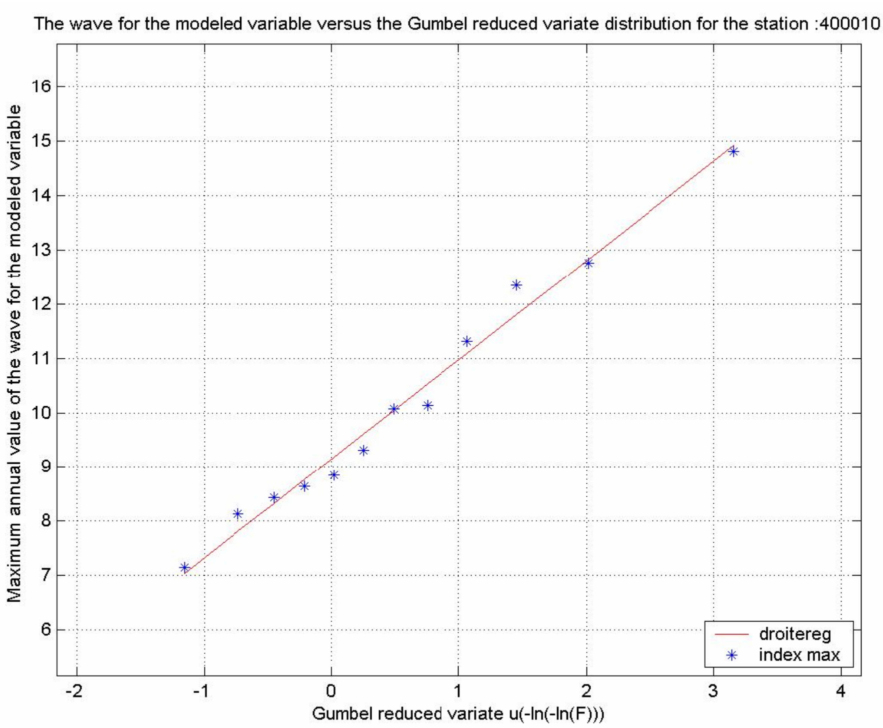
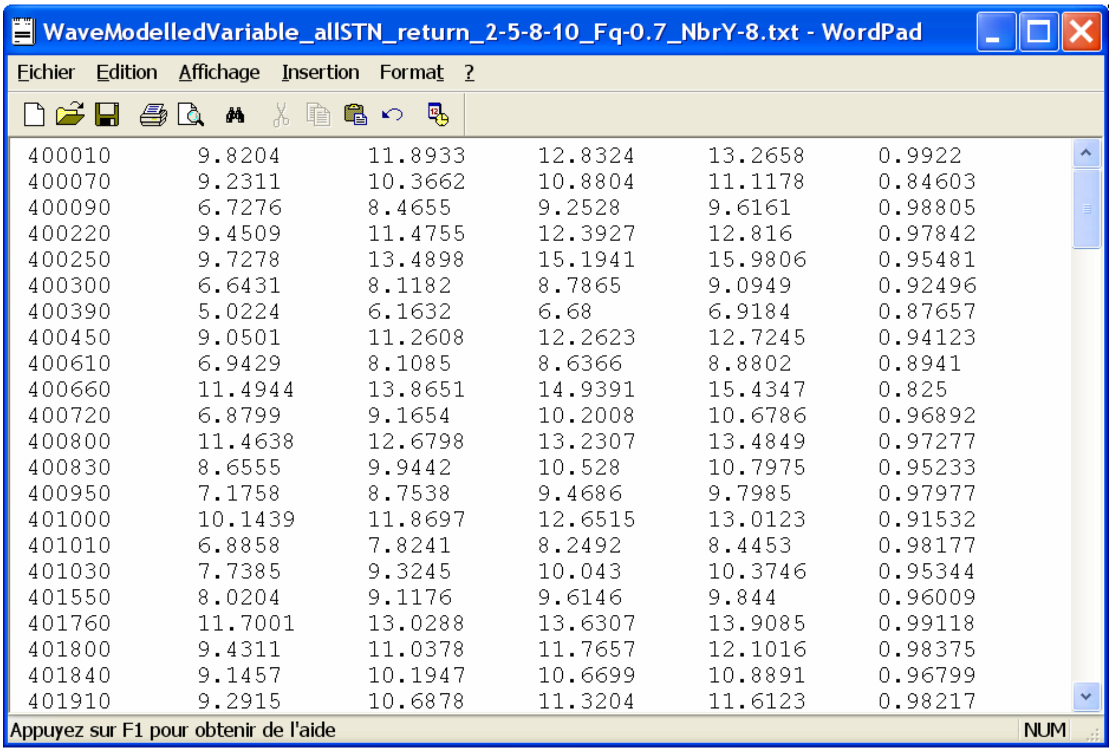


Figure 13. The annual maximum daily mean wind speed versus the Gumbel reduced variate distribution for weather station 400010 with the annual frequency threshold 70% and eight year return period

In addition to these graphs, the application also generates a summary file **WaveModelledVariable _allSTN_return_2-5-8-10_Fq-0.7_NbrY-8.txt** for each studied zone and places it in the same folder. This file contains, for each station, the annual maximum daily mean wind speed for two, five, eight and ten year return periods as well as the correlation value between the annual maximum daily mean wind speed and a Gumbel reduced variable (see figure 13). Figure 14 presents an example of such a file.



400010	9.8204	11.8933	12.8324	13.2658	0.9922
400070	9.2311	10.3662	10.8804	11.1178	0.84603
400090	6.7276	8.4655	9.2528	9.6161	0.98805
400220	9.4509	11.4755	12.3927	12.816	0.97842
400250	9.7278	13.4898	15.1941	15.9806	0.95481
400300	6.6431	8.1182	8.7865	9.0949	0.92496
400390	5.0224	6.1632	6.68	6.9184	0.87657
400450	9.0501	11.2608	12.2623	12.7245	0.94123
400610	6.9429	8.1085	8.6366	8.8802	0.8941
400660	11.4944	13.8651	14.9391	15.4347	0.825
400720	6.8799	9.1654	10.2008	10.6786	0.96892
400800	11.4638	12.6798	13.2307	13.4849	0.97277
400830	8.6555	9.9442	10.528	10.7975	0.95233
400950	7.1758	8.7538	9.4686	9.7985	0.97977
401000	10.1439	11.8697	12.6515	13.0123	0.91532
401010	6.8858	7.8241	8.2492	8.4453	0.98177
401030	7.7385	9.3245	10.043	10.3746	0.95344
401550	8.0204	9.1176	9.6146	9.844	0.96009
401760	11.7001	13.0288	13.6307	13.9085	0.99118
401800	9.4311	11.0378	11.7657	12.1016	0.98375
401840	9.1457	10.1947	10.6699	10.8891	0.96799
401910	9.2915	10.6878	11.3204	11.6123	0.98217

Figure 14. Example of summary file resulting from the application of Gumbel analysis on all the stations in a given region

On the basis of the graph and summary file, the user is then identifying any potential errors in the original datasets, potential outliers (points far from the regression line on the graph) and/or correlation value lower than 0.80 for example.

In these cases, the user writes down the number of the concerned stations and follows the steps reported in section 3.3.2.2.

3.3.2.2 Correction and/or adjustment of the original dataset for unusual observations

The application of the Gumbel frequency analysis on all the weather stations might reveal some data entry mistakes and/or outliers (see section 3.3.2.2). When identified, such cases needs to be corrected in the original dataset in order to improve the correlation in the analysis and therefore reduce the error on the final values for the annual maximum daily mean wind speed for these stations.

The following case illustrates how to modify an error in the original dataset and run again the Gumbel analysis on that particular weather station.

The graph generated by the analysis for the weather station n°407040 (Figure 15) shows an isolated point (red circle on the graph). Let's for example consider that this particular measure correspond to a year (2004) where the measurement instruments at the weather station has been changed and that this resulted in wrong measurement of some of the climatic variable.

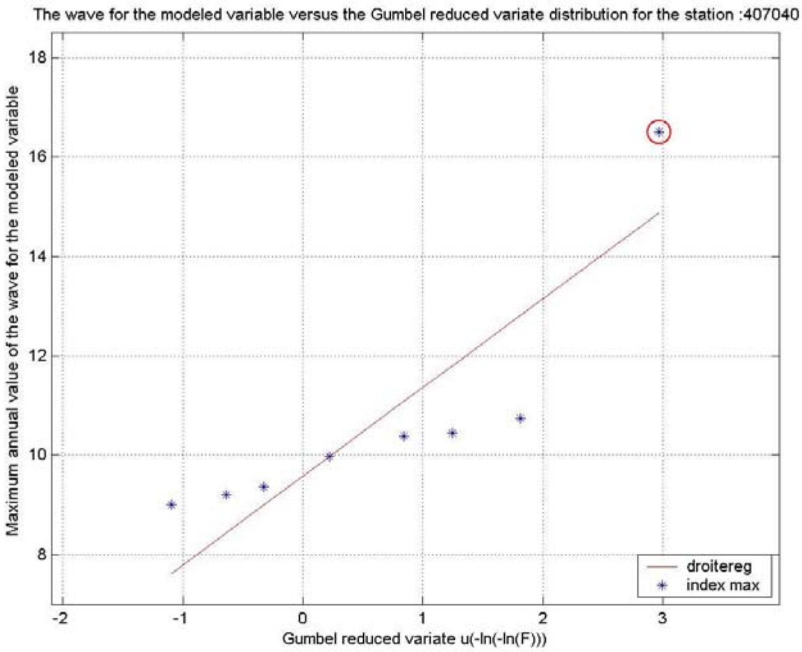


Figure 15. Graph resulting from the Gumbel frequency analysis for weather station 407040 with the isolated point indicated by the red circle

It would therefore be appropriate to remove this point from the analysis. To do so:

1. Make a copy of the file for that station (164290.txt) located in the "Stations files" folder
2. Using WordPad, open the version of the file (164290.txt) located in the "Stations files" folder (Figure 16),

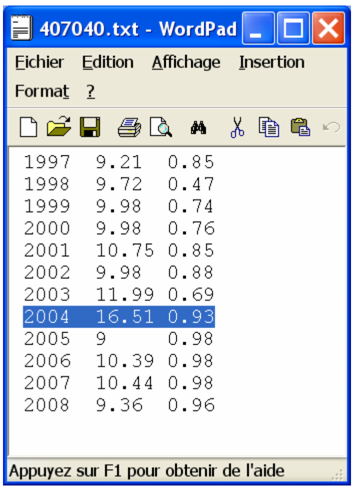


Figure 16. Weather station n° 407040 specific file in which the year 1999 record is highlighted in blue

3. delete the record (the all line) corresponding to year 2004 (16.51 m/s)
4. Save the file in the same location (do not move it); it is always preferable to make a copy before modifying a file

Once the record deleted, it is possible to run again the Gumbel frequency analysis but on that weather station only. For that:

1. in **EatlasClimMod 1.0**, go to **Operation>Gumbel analysis>One station** (Figure 17).

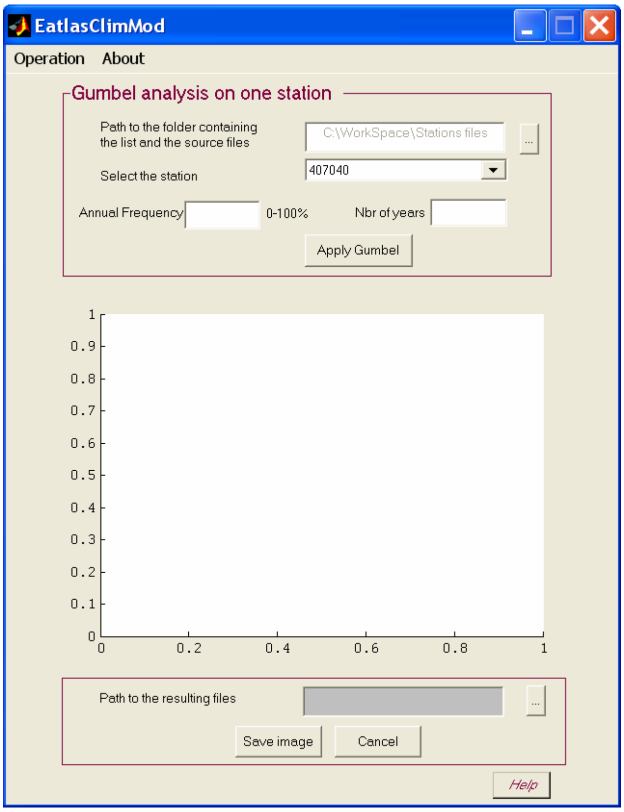


Figure 17. Interface window of the application of the Gumbel method to one station

In this window, the user must specify:

a. the path to the folder containing the weather station specific files (see section 3.3.1),
b. the number of the station to be corrected (407040) in this case
c. the two thresholds described in section 3.3.2 (annual frequency and minimum number of years of observations).

2. After completing all the boxes, click on the button Apply Gumbel.
3. The new graph is then displayed (Figure 18) . If the result is satisfactory, save the graph and the associated text file by selecting the path to the resulting files with the button ..., then by validating the change by clicking on the Save image button.

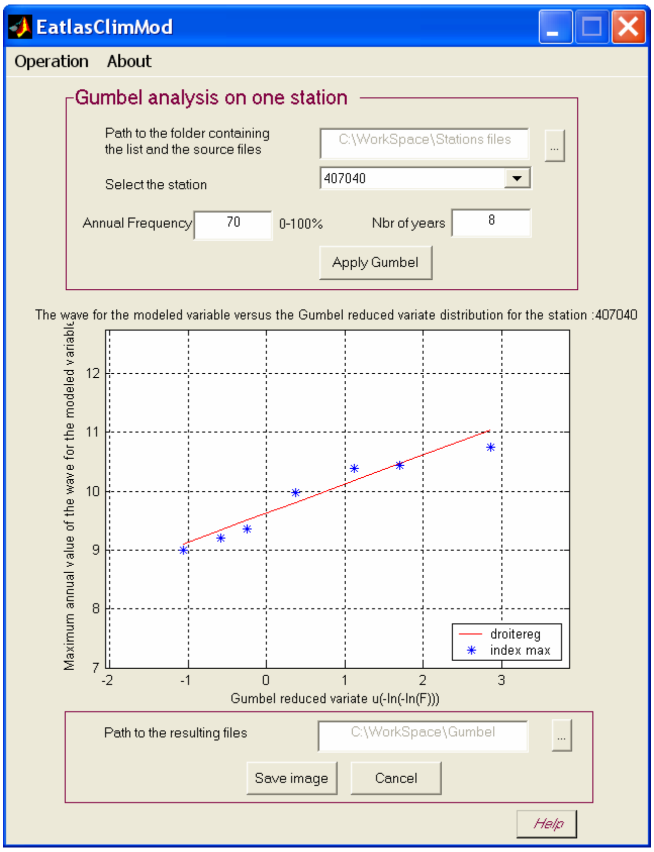


Figure 18. Interface result of Gumbel application after correction

This operation will have for result to modify the previous version of the summary file (Figure 14), file which is necessary for the rest of the process.

3.4 Selection of the parameters and identification of the regression model

As described in section 2.3, stepwise regression analysis was used to identify the variables explaining the variation of the annual maximum daily mean wind speed for each return period over the region covered in this version of the e-atlas and therefore derive the regression model to be used for the spatialization.

This process has been implemented in three steps, which are explained in the following sections:

- preparation of the GIS layers containing the spatial distribution of the causal factors and the dependant variable (annual maximum daily mean wind speed for two, five, and ten year return periods
- preparation of the stepwise regression analysis
- application of the stepwise regression analysis.

Due to the climatological characteristics of the study area, the stepwise regression analysis has been applied separately to the five zones described in section 2.3.

3.4.1 Preparation of the GIS layers containing the spatial distribution of the causal factors and dependant variable

As reported in section 2.3, the causal factors retained to explain the spatial variability of the annual maximum daily mean wind speed are:

1. *Elevation (Z)*
2. *Slope in percents (*Slp_pr)
3. *Distance to the nearest coastline (*d_Coast*)*
4. *Distance from the relative longitude (d_X) and latitude (d_*Y*)*
5. *Mean elevation within* a 3×3 pixel window around each *cell (*Z9*)*
6. *Aspect (*Asp*)*.

The spatial distribution of the first 2 causal factors can be directly and easily derived from the dataset generated for the implementation of the different models (see the Methodology document for the WHO e-atlas of disaster risk.Volume 1. Exposure to natural hazardsVersion 2.0, **Preparation of the datasets**). When it comes to the last 4 factors, specific process had to be applied in order to get the appropriate GIS layer for the analysis.

Before applying the steps reported in the next sections, each of the layers used in this process first had to be projected into a metric projection system. This has been done using the process presented in Annex 2.

In addition to that, and as mentioned in section 2.3, the all region covered in this version of the e-atlas had to be cut into 5 zones during the analysis to take their respective climatologic characteristics into account.

Each of the layers used in the regression analysis therefore had to be cut according to the extent of the respective zone to which a 300 km buffer has been added in order to ensure a good interpolation at the edge of each of the zone. This process is described in Annex 3.

3.4.1.1 Preparation of the distance to the nearest coastline layers

First, the coastlines have been extracted from the projected version of the international boundaries layer (**st_ar_int_bord_km.shp**) using the following steps:

1. Make sure that the XTools Extension is uploaded in ArcView.

2. Display in the view both the international boundaries (**st_ar_int_bord_km.shp**) and the layer containing the global coastline border coming from the SALB project **Un_coast_01.shp** reprojected into metric projection system (Annex 2) (**Un_coast_01_km.shp**).

3. Use the **XTools> Clip with polygon(s)** function.

4. Select the **Un_coast_01_km.shp** as the theme that contains features that you wish to clip.

5. Select **st_ar_int_bord_km.shp** as the polygon theme that contains the polygons that will be used as the reference for the clipping.

6. Specify the name for the new shapefile to be created as **st_ar_coast_km**.

The resulting projected coastline layer has then cut according to the 5 climatic zones (see Annex 3) to give the following layers: **zone1_coast_buffer_300.shp, zone2_coast_buffer_300.shp, zone3_coast_buffer_300.shp, zone4_coast_buffer_300.shp, zone5_coast_buffer_300.shp.**

From there, the distance from the coastline was computed as follows starting with the first climatic zone:

1. In ArcView, display the **zone1_coast_buffer_300.shp**

2. Navigate to **Analysis>Find Distance** to create a grid of distances from the coasts in kilometres:

3. in the next menu, select Output Grid Specification and specify as follow:
   a. set Output Grid Extent = Same As **zone1_int_bord_buffer_300.shp**
   b. set Output Grid Cell Size = As Specified Below
   c. set Cell Size = 1 km
   d. use the default number of rows and columns
   e. save the result as **zone1_dist_coast**.

4. Repeat steps 1 to 3 on the buffered coastline shapefiles for the other zones changing the name of the resulting files accordingly.

3.4.1.2 Preparation of the distance from the relative latitude/longitude layer

The following process was followed for generating the climatic zone specific distance from the relative latitude layers:

1. Create a line shapefile that will pass by the point located at the extreme South of the of the frist climatic zone as follows:

   a. In ArcView, add the **zone1_int_bord_buffer_300.shp** file in the view (See annex 3)
   b. create a new line theme using **View>New Theme,**
   c. manually digitize a straight horizontal line passing by the point located at the extreme South of the of this particular zone. This is going to be the zero degrees relative latitude for this zone,
   d. save the editing work as **zone1_Y.shp**.

2. Navigate to **Analysis>Find Distance** to create a grid of distances from the line generated in the **zone1_Y.shp** file.

3. in the next menu, select Output Grid Specification and specify as follow:
   a. set Output Grid Extent = Same As **zone1_int_bord_buffer_300.shp**

b. set Output Grid Cell Size = As Specified Below
c. set Cell Size = 1 km
d. use the default number of rows and columns
e. save the result as **zone1_dist_Y**.

4. Repeat steps 1 to 3 on the other zones.

The process used to create a layer containing the relative longitude for study area is identical to the process described for the relative latitude drawing this time a vertical line passing by the point located at the extreme West of each climatic zone. This would represent the zero degrees relative longitude. The resulting output is then saved as **zone*n*_dist_X** (*n* corresponds to the number associated with each climatic zone).

3.4.1.3 Preparation of the mean elevation distribution layer

The mean elevation distribution grid was derived from the DEM using the following steps:

1. In ArcView, make sure that the Grid Analyst extension is.
2. Add the **st_ar_dem*** grid in the view and activate it.
3. Use the **Analysis>Neighborhood Statistics** function specifying the following in the windows that appears:
   a. statistic = Mean
   b. under neighbourhood, type of neighbourhood for analysis = Rectangle
   c. select the “cell” dial; and set “Width” and “Height” to three cells
   d. save the resulting grid as **st_ar_Z9**.

The resulting grid has then been projected using the steps reported in Annex 2 before being cut according to the extent of each climatic zones to which a 300 km has been added and this following the process reported in Annex 3. The grids resulting from these operations are named: **Zone1_ Z9, Zone2_ Z9, Zone3_ Z9, Zone4_Z9** and **Zone5_Z9.**

3.4.1.4 Preparation of the aspect layers

The aspect, or slope direction, layer has been included in the regression analysis under the form of a dummy variable. This is being done as we don’t want to consider broad directions (N, NE, E, SE,…) and not any small variations in the direction of the slope.

In order the use the aspect distribution layer as a dummy variable in the regression analysis, eight grids named **Aspect_*X*** (where *X* is N, NE, E, SE, S, SW, W, and NW) have been derived according to the following classification:

- 0°–22.5° and 337.5°–360°: North
- 22.5°–67.5°: North East
- 67.5°–112.5°: East
- 112.5°–157.5°: South East

- 157.5°–202.5°: South
- 202.5°–247.5°: South West
- 247.5°–292.5°: West
- 292.5°–337.5°: North West

For example, in **Aspect_N** cells which slope is directed towards the North are given a value of 1 while any other cells are given the value 0.

The procedure used to create these grids is outlined in the following steps.

1. In ArcView, upload the aspect distribution grid **st_ar_aspect*** into the view.
2. Select the **Analysis>Map Calculator** function and enter the following formulas in the Calculator window: [st_ar_aspect]>=67.5 and [st_ar_aspect]<112.5.
3. Save the output grid as **st_ar_asp_E**. This corresponds to the spatial distribution of the aspect with eastern slopes.
4. Repeat these steps for the other directions using the above classification to generate **st_ar_asp_N**, **st_ar_asp_NE**, **st_ar_asp_SE**, **st_ar_asp_S**, **st_ar_asp_SW**, **st_ar_asp_W** and **st_ar_asp_NW**.

The resulting grids have then been projected using the steps reported in Annex 2 before being cut according to the extent of each climatic zones to which a 300 km has been added and this following the process reported in Annex 3.
The resulting files used in the regression analysis are therefore the following for the first climatic zone: **Zone1_asp_N**, **Zone1_ asp_NE**, **Zone1_ asp_E**, **Zone1_ asp_SE**, **Zone1_asp_S**, **Zone1_asp_SW**, **Zone1_asp_W** and **Zone1_asp_NW**.

3.4.2 Integration of the annual maximum daily mean wind speed figures into the weather stations location GIS layer

In order to complete the regression analysis, the annual maximum daily mean wind speed calculated for the different return periods, for each weather station and for each zone (see section 3.3.2) needed to be attached to the GIS layer containing the location of these stations. This was done using the following steps:

1. In Microsoft Excel, open the **WaveModelledVariable_allSTN_return_2-5-8-10_Fq-0.7_NbrY-8.txt** file created in section 3.3.2.1. This file contains the annual maximum daily mean wind speed for two, five, eight and ten year return periods for the weather stations for each zone

2. Add a line on top of the table and name each column as follow (from left to right):: "**STN**" (Station number), "**WDSP_2**" (annual maximum daily mean wind speed for a two year return period), "**WDSP_5**" (annual maximum daily mean wind speed for a five year return period), "**WDSP_8**" (annual maximum daily mean wind speed for an eight year return period), "**WDSP_10**" (annual maximum daily mean wind speed for a ten year return period). Save the result as **Zone1_wdsp_2_5_8_10.dbf**.

3. Repeat step 2 for the other zones.

4. Select the C-Tables Tools>Appends Tables Together function to create one unique table from the five created previously and save the result as **st_ar_wdsp_2_5_8_10.dbf**.

5. Merge the **st_ar_wdsp_2_5_8_10.dbf** with the **st_ar_stations.shp*** shapefile as follows:

   a. in ArcView add the **st_ar_wdsp_2_5_8_10.dbf** table in the table window
   b. add the **st_ar_stations.shp*** in the view and open its attribute table
   c. Select the header of the STN column in the **st_ar_wdsp_2_5_8_10.dbf** table and the header of the Number column in the attribute table of the **st_ar_stations.shp*** shapefile
   d. keeping the attribute table of the **st_ar_stations.shp*** file active, join the two tables by clicking the join button on the tool bar
   e. use the **C-Tables Tools>Make Joins Permanent** function to fix all the columns added in the attribute table of the **st_ar_stations.shp*** shapefile
   f. save the resulting shapefile as **st_ar_wdsp.shp**.

The resulting shape file has then been projected using the steps reported in Annex 2 before being cut according to the extent of each climatic zones to which a 300 km has been added and this following the process reported in Annex 3. The files resulting from these operations are named as follow: **zone_1_wdsp.shp, zone_2_wdsp.shp, zone_3_wdsp.shp, zone_4_wdsp.shp, zone_5_wdsp.shp**.

3.4.3 Preparation of the stepwise regression analysis

Before performing the stepwise regression analysis on each of the five zones it was necessary to prepare a table which contained, for each weather station, the annual maximum daily mean wind speed for each return period as well as the variables extracted from each grid prepared in section 3.4.1. The procedure used to create this table is outlined in the following steps.

1. In ArcView, make sure that the **Grid Analyst** extension is uploaded,
2. In a view, add the seven causal factor distribution grids for the first climatic zone **zone1_dem**, **zone1_Z9**, **zone1_slope**, **zone1_asp_Y**, **zone1_dist_coast**, **zone1_dist_X**, **zone1_dist_Y** and the shapefile containing the distribution of the weather stations to which the annual maximum daily mean wind speed for the four return periods have been associated for the first zone (**zone1_wdsp.shp**).
3. Make the **zone1_wdsp.shp** shapefile the active theme and use the **Grid Analyst>Extract X, Y and Z Values for Point Theme from Grid** functions
4. Select the first grids listed in step 2 from the drop list. The function will add and then populates three new fields in the attribute table of **zone1_wdsp.shp** (Xval, Yval and Zval), the last one storing the value extracted from the grid layer.
5. Open the attribute table of the **zone1_wdsp.shp** shapefile and click on the header of the "Zval" column.

6. Rename this field to correspond to the name of the raster layer (i.e. Z for elevation, Z9 for mean elevation, SLP for slope, ASP_*X* for aspect (where *X* is N, NE, E, SE, S, SW, W and NW respectively), d_Coast for the distance from the coastline, d_X for the distance from the relative longitude, d_Y for the distance from the relative latitude) using the C-Tables Tools>Rename/Resize/Copy Field(s) function.
7. Repeat steps 3 to 5 on the other causal factor distribution grids until the Zval for each of them is integrated into the attribute table of **the zone_1_wdsp.shp** shapefile.
8. Save the resulting table as **zone1_wdsp_regression.dbf**.
9. Repeat steps 2 to 8 on the other zones, changing the names of the resulting files accordingly.

3.4.4 Application of the stepwise regression analysis

Once the stepwise regression table is ready (see section 3.4.3), it is possible to perform the stepwise regression analysis on each zone and for each return period. This procedure is done using the S-Plus software as follows:

1. Launch the S-Plus software.

2. Choose **File>Import Data>From File** function to import the stepwise regression table created above for the first climatic zone (**zone_1_wdsp_regression.dbf**).

3. Choose **Statistics>Regression>Stepwise** function.

4. In the "Stepwise Linear Regression" dialogue box that appears:

    a. under Data Set scroll down the list and click on **zone_1_wdsp_regression**
    b. click on the Create Formula box for the Upper Model and use wdsp_2 as the response and add all explanatory variables (Z, Z9, SLP, ASP_*X,* d_Coast, d_X, d_Y) as Main Effects and Quadratic

       with: wdsp_2 = annual maximum daily mean wind speed for a two year return period
       Z = elevation
       Z9 = mean elevation
       SLP = slope
       ASP_*X* = aspect (where *X* is N, NE, E, SE, S, SW, W and NW)
       d_Coast = distance from coastline
       d_X = Distance from the relative longitude
       d_Y = Distance from the relative latitude
    c. click OK to run the procedure.

    A report is created that shows the result of this selection procedure with the coefficient of the variables selected and their significance, residual standard error, multiple R^2 and probability (F statistic). Table 5 shows for the report obtained for a two year return period in zone 2.

The regression equation explaining the maximum wind speed for this particular return period (two years) and zone 2 can be read as follows:

$$wdsp\_2 = 0.002482*Z+0.006111*d\_Y–0.012497*d\_Coast–0.000001*Z^2–0.000001*d\_Y^2+0.000019*d\_Coast^2$$

Table 5. Results of the regression analysis for the annual maximum wind speed over zone 2 for a two year return period

Variable	Regression coefficient	Standard error	*t* value	Probability Pr(>|*t*|)
Z	0.002482	0.001229	2.018447	0.005871
d_Y	0.006111	0.000365	16.7108	0.000000
d_Coast	–0.012497	0.004348	–2.873828	0.004831
Z^2	–0.000001	0.000000	–1.895487	0.006538
d_Y^2	–0.000001	0.000000	–9.343988	0.000000
d_Coast^2	0.000019	0.000007	2.751341	0.006897

Residual standard error	0.43629
Degrees of freedom	115
Multiple R^2	0.95
F statistic	132.94746
Probability (*F* statistic)	0.0000

5. Repeat steps 3 to 4 for the other return periods in the same zone.

6. Repeat steps 2 to 5 for the other zones, changing the names of the files accordingly.

3.5 Spatialization of the annual maximum daily mean wind speed for each return period

The annual maximum daily mean wind speed distribution map for each return period and zone is created by applying the regressions found in section 3.4.4 on the corresponding grids as follows (example for the second climatic zone and two year return period):

1. Make sure that all the seven causal factor distribution layers for the first zone are uploaded in the view.

2. Select the **Analysis>Map Calculator** function and enter the following formula in the calculator window:

wdsp_2=(0.002482*[zone2_dem])+(0.006111*[zone2_dist_Y])–(0.012497*[zone2_dist_coast])–(0.000001*[zone2_dem]*[zone2_dem])–(0.000001*[zone2_dist_Y]*[zone2_dist_Y])+(0.000019*[zone2_dist_coast]*[zone2_dist_coast])

3. Save the output grid as **zone2_wdsp_2.** This corresponds to the spatial distribution of the annual maximum daily mean wind speed over zone 2 and for a two year return period.

4. Unproject the **zone2_ wdsp_2** layer from the Equal-Area Cylindrical projection to the geographic one using the following steps:

   a. click either the  button or use the Grid Projector>Grid and Theme Projector function

   b. select **zone2_ wdsp_2** from the list as the grid to project

   c. In the Grid Projector window:
      - Specify the parameters for the current projection as follows:
        - Category = projection of the world
        - Type = Equal-Area Cylindrical projection
        - Current Projection Units = kilometres

      - specify the parameters for the new projection as follows:
        - Category = projection of the world
        - Type = Geographic
        - New Projection Units = decimal degrees

   d. In the next window, specify the new cell size = 0.008333.

   e. Save the output grids as **zone2_ wdsp_2_d**.

5. Repeat steps 2 to 4 for the other return periods using the corresponding regressions.

6. Repeat steps 1 to 5 on the other zones

The annual maximum daily mean wind speed distribution maps for each return period and each climatic zone were then merged to generate three grids, one for each return period, covering the study area using the following steps.

1. In ArcView, make sure that the Grid Transformation Tool extension is uploaded.

2. Select the Transform **Grid>Mosaic** function to create one unique grid from **zone1_wdsp_2_d**, **zone2_wdsp_2_d**, **zone3_wdsp_2_d**, **zone4_wdsp_2_d** and **zone5_wdsp_2_d**.

3. Save the result as **st_ar_wdsp_2_d**.

4. Repeat steps 2 and 3 for the other two return periods and save the output mosaic as **st_ar_wdsp_5_d**, **st_ar_wdsp_10_d**.

Finally the following process was applied to clip the annual maximum daily mean wind speed distribution mosaic map for each return period to the borders of the region covered in this version of the e-atlas.

1. In ArcView, make sure that the Grid Analyst extension is uploaded.

2. upload the AFRO international boundary level used for the e-atlas (**afro_int_bnd.shp**)

3. Activate **st_ar_wdsp_2_d** and select the **Grid Analyst>Extract Grid Theme Using Polygon** function

    - Click "yes" to continue

    - Select the **afro_int_bnd.shp** from the dropdown list to be used as the layer on which the grid needs to be clipped and click OK

    - Make the output grid the active theme and choose the **Theme>Convert to Grid** function to create the AFRO annual maximum daily mean wind speed distribution map for 2 year return period

    - Save it as **afro_wdsp_2**.

4. Repeat steps 2 and 3 for the others WHO Regions (EURO and EMRO) and others return periods (five and ten) and save the outputs as **afro_wdsp_5**, **afro_wdsp_10**, **emro_wdsp_2**, **emro_wdsp_5**, **emro_wdsp_10**, **euro_wdsp_2**, **euro_wdsp_5**, **euro_wdsp_10**.

The metadata associated with these grids are reported in Annex 4.

3.6 Creation of the wind speed hazard distribution maps

The last step of the method consisted in reclassifying the annual maximum daily mean wind speed distribution maps to correspond to the five intensity levels selected for this project using the following process.

1. Upload the three annual maximum daily mean wind speed distribution grids into the view.

2. Make the **afro_wdsp_2** grid the active theme.

3. Use the Analysis>Reclassify function to reclassify the grid according to the following classification:

Wind speed range (m/s)	Intensity level
<3.3	1
3.3–10.7	2
10.7–17.1	3
17.1–24.4	4
≥24.4	5

4. Save the output grid as **afro_wdsp_2_cl**.

5. Select Theme>Edit Legend in the Legend Editor window; change the legend to:
   1: very low
   2: low
   3: medium
   4: high
   5: very high.

6. Repeat steps 1 to 5 for the for the others WHO Regions (EURO and EMRO) and others return periods (five and ten years) and save the results as **afro_wdsp_5_cl**, **afro_wdsp_10_cl**, **emro_wdsp_2_cl**, **emro_wdsp_5_cl**, **emro_wdsp_10_cl**, **euro_wdsp_2_cl**, **euro_wdsp_5_cl**, **euro_wdsp_10_cl**.

The map resulting from the application of this approach for the European Region is reported in Figure 19. Please refer to the e-atlas DVD itself for the maps covering the other two WHO Regions and the other return period. The associated metadata for these layers can be found in Annex 5.

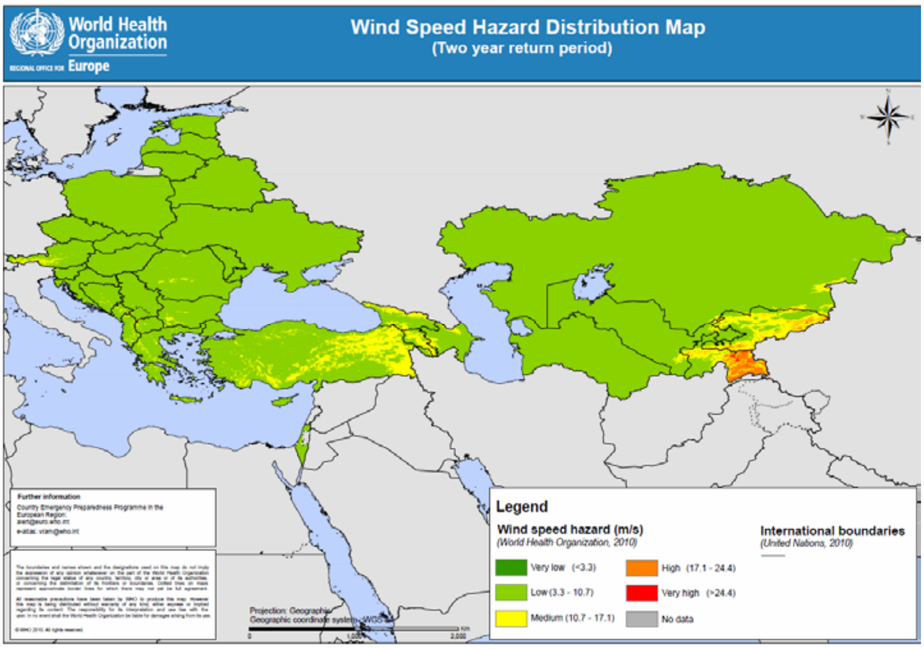


Figure 19. Wind speed hazard distribution map (Two year return period) for the countries of the European Region covered in this version of the *WHO e-atlas of disaster risk*

References and further reading

Agnew MD, Palutikof JP, 2000. GIS-based construction of baseline climatologies for the Mediterranean using terrain variables. *Climate research*, 14:115–27.

Ashraf M, Loftis JC, Hubbard KG, 1997. Application of geostatistics to evaluate partial weather station networks. *Agricultural and forest meteorology*, 84(3–4):255–71.

Benzi R, Deidda R, Marrocu M, 1997. Characterization of temperature and precipitation fields over Sardinia with principal component analysis and singular spectrum analysis. *International journal of climatology*, 17(11):1231–62.

Chessa PA, Delitala AM, 1997. Objective analysis of daily extreme temperatures of Sardinia (Italy) using distance from sea as independent variable. *International journal of climatology*, 17(13):1467–1485.

Collins FC Jr, Bolstad PV, 1996. A comparison of spatial interpolation techniques in temperature estimation. *Proceedings of the Third International Conference/Workshop on Integrating GIS and Environmental Modeling. January 21–25, 1996, Santa Fe, New Mexico, USA*.

Condie R, Lee KA, 1982. Flood frequency analysis with historic information. *Journal of hydrology*, 58:47–61.

Dodson R, Marks D, 1997. Daily air temperature interpolated at high spatial resolution over a large mountainous region. *Climate research*, 8(1):1–20.

Eckstein BA, 1989. Evaluation of spline and weighted average interpolation algorithms. *Computers & geosciences*, 15(1):79–94.

Eischeid JK, Baker FB, Karl TR, Diaz HF, 1995. The quality control of long-term climatological data using objective data analysis. *Journal of applied meteorology*, 34(12):2787–95.

El Morjani Z, 2003. Conception d'un système d'information à référence spatiale pour la gestion environnementale; application à la séléction de sites potentiels de stockage de déchets ménagers et industriels en région semi-aride (Souss, Maroc). Doctoral thesis, University of Geneva. *Terre et environnement*, vol. 42.

Fuller WE, 1914. Flood flows. *Transactions of the American Society of Civil Engineers*, 77:564–617.

Foster HA, 1935. Duration curves. *Transactions of the American Society of Civil Engineers*, 99:1213–35.

Gerard R, Karpuk EW, 1979. Probability analysis of historical flood data. *Journal of the Hydraulics Division, ASCE*, 105(HY9):1153–65.

Gumbel EJ, 1941. The return period of flood flows. *American mathematical statistics*, 12(2):163–90.

Gumbel EJ, 1942. Statistical control curves for flood discharges. *Transactions, American Geophysical Union*, 23:489–500.

Gumbel EJ, 1954. The statistical theory of droughts. *Proceedings of the American Society of Civil Engineers*, 80:1–19.

Gumbel EJ, 1960. Bivariate exponential distribution. *Journal of the American Statistical Association*, 55:698–707.

Hammond T, Yarie J, 1996. Spatial prediction of climatic state factor regions in Alaska. *Ecoscience*, 3(4):490–501.

Hargy VT, 1997. Objectively mapping accumulated temperature for Ireland. *International journal of climatology*, 17(9):909–27.

Holdaway MR, 1996. Spatial modeling and interpolation of monthly temperature using kriging. *Climate research*, 6:215–25.

Hudson G, Wackernagel H, 1994. Mapping temperature using kriging with external drift: theory and an example from Scotland. *International journal of climatology*, 14:77–91.

Hulme M, Conway D, Jones PD, Jiang T, Barrow EM, Turney C, 1995. Construction of a 1961–1990 European climatology for climate change modeling and impact applications. *International journal of climatology*, 15(12):1333–63.

Hutchinson MF, Gessler PE 1994. Splines—more than just a smooth interpolator, *Geoderma*, 62:45–67.

Jones T, Middelmann M, Corby N, 2005. *Natural hazard risk in Perth, Western Australia*. Canberra, Geoscience Australia.

Kite GW, 1977. *Frequency and risk analyses in hydrology*. Fort Collins, Colorado, Water Resources Publications.

Koutsoyiannis D, 2004. On the appropriateness of the Gumbel distribution for modeling extreme rainfall. In Brath A, Montanari A, Toth E, eds. *Hydrological risk: recent advances in peak river flow modeling, prediction and real-time forecasting. Assessment of the impacts of land-use and climate changes*:303–19. Bologna, Italy, Editoriale Bios, Castrolibero.

Landwehr JM, Matalas NC, Wallis JR, 1979. Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Water resources research*, 15:1055–64.

Legates DR, Willmott CJ, 1990. Mean seasonal and spatial variability in global surface air temperature. *Theoretical and applied climatology*, 41:11–21.

Lennon JJ, Turner JRG, 1995. Predicting the spatial distribution of climate: temperature in Great Britain. *Journal of animal ecology*, 64:370–92.

Li J, Huang JF, Wang XZ, 2006. A GIS-based approach for estimating spatial distribution of seasonal temperature in Zhejiang Province, China. *Journal of Zhejiang University science A*, 7(4):647–56.

MacEachren AM, Davidson JV, 1987. Sampling and isometric mapping of continuous geographic surfaces. *American cartographer*, 14(4):299–320.

Matheron G, 1963. Principles of geostatistics. *Economic geology*, 58:1246–66.

Meylan P, Musy A, 1998. Hydrologie fréquentielle. *Edition HGA Bucarest*, 413p

Moin SMA, Shaw MA, 1985. *Regional flood frequency analysis for Ontario streams: volume 1, single station analysis and index method*. Burlington, Ontario, Inland Waters Directorate, Environment Canada.

Myers RH, 1990. *Classical and modern regression with applications*, Boston, Massachusetts, PWS-Kent Publishing.

Sarma P, 1999. Flood risk zone mapping of Dikrong sub basin in Assam. At http://www.gisdevelopment.net/application/natural_hazards/floods/nhcy0006pf.htm; link checked 16 December 2010.

Stedinger JR, Vogel RM, Foufoula-Georgiou E, 1992. Frequency analysis of extreme events. In Maidment DA, ed. *Handbook of hydrology.* New York, McGraw-Hill.

Stolte W, Dumontier S, 1977. *Flood frequency analysis for mountain and prairie streams*. University of Saskatchewan and Alberta Environment.

USACE, 1993. *Hydrologic frequency analysis*. Engineer Manual 1110-2-1415. Washington DC, US Army Corps of Engineers.

Vogel RM, 1986. The probability plot correlation coefficient test for normal, lognormal, and Gumbel distributional hypothesis. *Water resources research*, 22(4):587–90; corrections, 23(10):2013.

Vogt JV, Viau AA, Paquet F, 1997. Mapping regional air temperature fields using satellite-derived surface skin temperatures. *International journal of climatology*, 17(14):1559–79.

Willmott CJ, Matsuura K, 1995. Smart interpolation of annually averaged air temperature in the United States. *Journal of applied meteorology*, 34(12):2577–86.

Annex 1. Description of the NCDC daily meteorological elements dataset

This annex provides the indication of the type of the data (int[eger], real or char[acter]) as well as a description of each of the fields of the daily meteorological data coming from the global surface summary of the day data produced by the National Climatic Data Center (NCDC).

Field	Type	Description
STN	Int	Station number
WBAN	Int	This is the historical "Weather Bureau Air Force Navy" number where applicable
YEARMODA	Int	Year, month and day
TEMP	Real	Mean temperature for the day in degrees fahrenheit to tenths. Missing = 9999.9
Count	Int	Number of observations used in calculating mean temperature
DEWP	Real	Mean dew point for the day in degrees fahrenheit to tenths. Missing = 9999.9
Count	Int	Number of observations used in calculating mean dew point
SLP	Real	Real mean sea level pressure for the day in millibars to tenths. Missing = 9999.9
Count	Int	Number of observations used in calculating mean sea level pressure
STP	Real	Mean station pressure for the day in millibars to tenths. Missing = 9999.9
Count	Int	Number of observations used in calculating mean station pressure
VISIB	Real	Mean visibility for the day in miles to tenths. Missing = 999.9
Count	Int	Number of observations used in calculating mean visibility
WDSP	Real	Mean wind speed for the day in knots to tenths. Missing = 999.9
Count	Int	Number of observations used in calculating mean wind speed

Field	Type	Description
MXSPD	Real	Maximum sustained wind speed reported for the day in knots to tenths. Missing = 999.9
GUST	Real	Maximum wind gust reported for the day in knots to tenths. Missing = 999.9
MAX	Real	Maximum temperature reported during the day in degrees fahrenheit to tenths. Missing = 9999.9
Flag	Char	Blank indicates maximum temperature was taken from the explicit maximum temperature report and not from the hourly data. * indicates maximum temperature was derived from the hourly data (i.e. highest hourly or synoptic-reported temperature)
MIN	Real	Minimum temperature reported during the day in degrees fahrenheit to tenths—time of minimum temperature report varies by country and region, so this will sometimes not be the minimum for the calendar day. Missing = 9999.9
Flag	Char	Blank indicates minimum temperature was taken from the explicit minimum temperature report and not from the hourly data. * indicates minimum temperature was derived from the hourly data (i.e. lowest hourly or synoptic-reported temperature)
PRCP	Real	Total precipitation (rain and/or melted snow) reported during the day in inches and hundredths; will usually not end with the midnight observation—i.e. may include latter part of previous day. 0.00 indicates no measurable precipitation (includes a trace). Missing = 99.99

Field	Type	Description
Flag	Char	A = one report of 6-hour precipitation amount
		B = summation of two reports of 6-hour precipitation amount
		C = summation of three reports of 6-hour precipitation amount
		D = summation of four reports of 6-hour precipitation amount
		E = one report of 12-hour precipitation amount
		F = summation of 2 reports of 12-hour precipitation amount
		G = one report of 24-hour precipitation amount
		H = station reported 0 as the amount for the day (e.g. from 6-hour reports), but also reported at least one occurrence of precipitation in hourly observations—this could indicate a trace occurred, but should be considered as incomplete data for the day
		I = station did not report any precipitation data for the day and did not report any occurrences of precipitation in its hourly observations—it is still possible that precipitation occurred but was not reported
SNDP	Real	Snow depth in inches to tenths—last report for the day if reported more than once. Missing = 999.9
		Note: most stations do not report 0 on days with no snow on the ground; therefore, 999.9 will often appear on these days
FRSHTT	Int	Indicators (1 = yes, 0 = no/not reported) for the occurrence during the day of: fog (‘F’—1st digit); rain or drizzle (‘R’—2nd digit).
		Snow or ice pellets (‘S’—3rd digit).
		Hail (‘H’—4th digit).
		Thunder (‘T’—5th digit).
		Tornado or funnel cloud (‘T’—6th digit).

Annex 2. Projection of a GIS layers into the metric projection system

This operation was used to switch the map units of the different layers used in the analysis from decimal degrees to kilometres in order to be able to measure distances.

The operation was performed separately on each the vector and raster layers.

Taking the international boundaries as an example (**st_ar_int_bord.shp***), all the vector layers have been projected using the following steps:

1. In ArcView, make sure that the Grid and Theme Projector v.2 extension are uploaded.
2. Click either the  button or use the Grid **Projector>Grid and Theme Projector** function from the menu.
3. In the window that open, select the **st_ar_int_bord.shp*** layer from the list as the theme to project.
4. In the next window that opens:

    a. Specify the following parameters for the current projection
    Category = projection of the world
    Type = Geographic
    Current Projection Units = decimal degrees

    b. Specify the following parameters for the new projection
    Category = projection of the world
    Type = Equal-Area Cylindrical
    New Projection Units = kilometers.

5. Save the output theme as **st_ar_int_bord_km.shp**.

Use the same process to reproject the **st_ar_WDSP.shp** shapefile and the global coastline border coming from the SALB project **Un_coast_01.shp** and then save the result respectively as **st_ar_WDSP_km.shp** and **st_ar_coast_km.shp.**

The raster layers have also been projected. Here is for example the process followed for the Digital Elevation Model (DEM):

1. Click either the  button or select **Grid Projector>Grid and Theme Projector**.

2. Select the **st_ar_dem** grid from the list in the next window.

3. In the next window that opens:

    a. specify the parameters for the current projection
    Category = projection of the world
    Type = Geographic

Current Projection Units = decimal degrees

b. specify the parameters for the new projection
Category = projection of the world
Type = Equal-Area Cylindrical
New Projection Units = kilometres

c. in the next window, specify the new cell size (in km) = 1

d. choose Interpolation Method = Bilinear Interpolation and Transformation Order = 4

e. save the output grid as **st_ar_dem_km**

Annex 3. Creation of a 300 km buffer around each climatic zone and clipping of the different layers for the regression analysis

In order to insure good interpolation at the edge of each of the 5 climatic zones considered in this version of the e-atlas to conduct the regression analysis, a 300 km buffer was added to each of these zones using the following process:

1. Extract the five zones (see section 2.3) from the projected version of the international boundaries layer (See Annex 2) using the following steps in ArcView:

    a. In ArcView, make sure that the XTools Extension is uploaded
    b. add the projected version of the international boundaries layer **st_ar_int_bord_km.shp**, and open its attribute table,
    c. select the countries part of the first zone (see section 2.3)
    d. In the view, make the **st_ar_int_bord_km.shp** shapefile the active theme
    e. use the **Theme>Convert To** Shapefile function to create a shapefile containing only the countries part of the first zone and save it as **zone1_int_bord_km.shp**
    f. repeat steps c) to e) for the other four zones.

2. Create the 300 km buffer around each climatic zone as follows:

    a. from the **View>Properties** change the Map Units and Distance Units to kilometres
    b. select **Theme>Create Buffer**; the Create Buffers wizard appears
    c. in the first wizard dialog box, choose **zone1_int_bord_km.shp** as the items to buffer
    d. in the second dialogue box, select *At a Specified Distance* as the method used to create the buffer and Width of the Buffer = 300. Make sure that the distance units are set to kilometres
    e. in the third dialogue box, create the buffer using *Only Inside of the Polygon* Parameter. Specify that the buffer be saved as a new theme and save as **zone_1_int_bord_buffer_300.shp**. The new buffer theme will be added to the current view
    f. Repeat steps c) to e) on the other four zones.

Taking the projected version of the weather station location layer (**st_ar_WDSP_km.shp**) as an example, the following steps have then been applied to clip each vector layer to the 5 buffered climatic zones.

1. In ArcView, make sure that the XTools Extension is uploaded,

2. In the view, display both the buffered international boundaries of the first zone **zone1_int_bord_buffer_300.shp** and the projected layer containing the distribution of the weather stations with the associated annual maximum daily mean wind speed **st_ar_WDSP_km.shp**.

3. Use the **XTools> Clip with Polygon(s)** function.

4. Select **st_ar_WDSP_km.shp** as the theme that contains features that you wish to clip.

5. Select **zone1_int_bord_buffer_300.shp** as the polygon theme that contains the polygons that will be used as the reference for the clipping.

6. Specify the name for the new shapefile to be created as **zone_1_WDSP.shp**.

7. Repeat steps 2 to 6 for the others zones

For the raster layers, taking the DEM as an example, the following steps are applied for the clipping.

1. In ArcView, make sure that the Grid Analyst extension is uploaded,

2. In the view, add the **st_ar_dem_km** grid and **zone1_int_bord_buffer_300.shp** shape file.

3. Make the first grid **st_ar_dem_km** active and use the **Grid Analyst>Extract Grid Theme Using Polygon** function.

4. Select first the **zone1_int_bord_buffer_300.shp** from the drop list to use in the clip.

5. Make the resulting grid the active theme and select the **Theme>Convert to Grid** function to save the output grid under **zone1_dem**.

6. Repeat steps 2 to 5 for the four other zones.

Annex 4. Metadata for the annual maximum daily mean wind speed distribution layers (two, five and ten year return periods)

Dataset title	Spatial distribution of annual maximum wind speed for the WHO Regions (Africa, Eastern Mediterranean and part of Europe)
Theme keywords	WHO, Africa, Eastern Mediterranean, Europe, natural disaster, Geographic Information System (GIS), natural hazard, wind speed, windstorm
Dataset topic category	Wind speed
Geographic location	The layer cover a total of 100 countries (22 for the Eastern Mediterranean, 46 for Africa and 32 for Europe)
Publication date	20110401
Data exchange format	ArcView grid
Filename	st_ws2_ucl, st_ws5_ucl, st_ws10_ucl
Dataset edition	Second edition
Abstract	This dataset contains the spatial distribution of the annual maximum wind speed for two, five and ten year return periods over the WHO Regions (Africa, Eastern Mediterranean and part of Europe)
Lineage	The process used to create the annual maximum wind speed for two, five and ten year return periods distribution layer is described in the *Methodology and implementation process for modelling the spatial distribution of wind speed hazard document that can be found in the first volume (2nd version) of the WHO e-atlas of disaster risk for the WHO Regions (Africa, Eastern Mediterranean and part of Europe)*
Data quality comments	Please refer to the data specific metadata for more information regarding the quality of the grids used as input for the creation of the annual maximum wind speed grid Because of the methods and resolution used (1 km) special care should be taken when using this dataset for application below the national level

	Even though the dataset covers all of the WHO Regions, the lack of historical data for many countries reduces the quality of the dataset for those areas
Distributor	WHO Mediterranean Centre for Health Risk Reduction (WMC)
Spatial representation type	grid
Map projection	Unprojected (Geographic)
Reference system	WGS 84 datum
Geographic box	X min: –25.358747°, X max: 91.8287° Y min: –46.978931°, Y max: 63.459827°
Resolution	30 arc-seconds (0.008333°)
Redistributions constraints	The annual maximum wind speed distribution layers are copyrighted. The owner of the data agrees to the use, reproduction, distribution, display, publication and dissemination at no cost to third parties of the annual maximum wind speed distribution layers, in any manner and in any form whatsoever, subject to the copyright and acknowledgement mentioned in these metadata
Access and use constraints	These layers may not be reproduced, changed, adapted, translated, stored in a retrieval system or transmitted in any form or by any means without prior permission of the copyright holder, except to make a security backup. Requests for permissions, with a statement of purpose and extent, should be address to the VRAM programme at the WHO Mediterranean Centre for Health Risk Reduction (VRAM@who.int)
Acknowledgement	*WHO e-atlas of disaster risk for the WHO Regions (Africa, Eastern Mediterranean and part of Europe)* 2nd edition. Copyright © WHO 2011. All rights reserved
Disclaimer	All reasonable precautions have been taken by WHO to produce these layers. However these layers are being distributed without warranty of any kind, either express or implied regarding their content. The responsibility for their interpretation and use lies with the user. In no event shall the World Health Organization be liable for damages arising

	from their use
Dataset language	English
Dataset character set	ASCII
Metadata provider	WHO Mediterranean Centre for Health Risk Reduction (WMC)
Metadata contact	El Morjani Zine El Abidine BP 3566 Poste Talborjt 80000 Agadir Morocco Telephone: +212 528 28 55 30 email: elmorjaniz@gmail.com
Metadata date	20110401
Metadata language	English
Metadata character set	ASCII
Metadata standard	ISO 19115

Annex 5. Metadata for the wind speed hazard distribution layers (two, five and ten year return periods)

Dataset title	Spatial distribution of the intensity level of wind speed hazard for the WHO Regions (Africa, Eastern Mediterranean and part of Europe)
Theme keywords	WHO, Africa, Eastern Mediterranean, Europe, natural disaster, Geographic Information System (GIS), natural hazard, wind speed, windstorm
Dataset topic category	Wind speed hazard
Geographic location	The layer covers a total of 100 countries (22 for the Eastern Mediterranean, 46 for Africa and 32 for Europe)
Publication date	20110401
Data exchange format	ArcView grid
Filename	st_ws2_cl, st_ws5_cl, st_ws10_cl
Dataset edition	Second edition
Abstract	This dataset contains the spatial distribution of the intensity level of wind speed hazard for two, five and ten year return periods over the WHO Regions (Africa, Eastern Mediterranean and part of Europe) according to five intensity levels (very low, low, medium, high and very high)
Lineage	The process used to create the intensity level of wind speed hazard distribution maps is described in the *Methodology and implementation process for modelling the spatial distribution of wind speed hazard document that can be found in the first volume (2nd version) of the WHO e-atlas of disaster risk for the WHO Regions (Africa, Eastern Mediterranean and part of Europe)*
Data quality comments	Please refer to the data specific metadata for more information regarding the quality of the grids used as input for the creation of the wind speed hazard grid Because of the methods and resolution used (1 km) special care should be taken when using this dataset for application

	below the national level Even though the dataset covers all of the WHO Regions, the lack of historical data for many countries reduces the quality of the dataset for those areas
Distributor	WHO Mediterranean Centre for Health Risk Reduction (WMC)
Spatial representation type	grid
Map projection	Unprojected (Geographic)
Reference system	WGS 84 datum
Geographic box	X min: –25.358747°, X max: 91.8287° Y min: –46.978931°, Y max: 63.459827°
Resolution	30 arc-seconds (0.008333°)
Redistributions constraints	The intensity level of wind speed hazard distribution layers are copyrighted. The owner of the data agrees on the use, reproduction, distribution, display, publication and dissemination at no cost to third parties of the intensity level of wind speed hazard distribution layers, in any manner and in any form whatsoever, subject to the copyright and acknowledgement mentioned in these metadata
Access and use constraints	These layers may not be reproduced, changed, adapted, translated, stored in a retrieval system or transmitted in any form or by any means without prior permission of the copyright holder, except to make a security backup. Requests for permission, with a statement of purpose and extent, should be address to the VRAM programme at the WHO Mediterranean Centre for Health Risk Reduction (VRAM@who.int)
Acknowledgement	*WHO e-atlas of disaster risk for the WHO Regions (Africa, Eastern Mediterranean and part of Europe)* 2nd edition. Copyright © WHO 2011. All rights reserved
Disclaimer	All reasonable precautions have been taken by WHO to produce these layers. However these layers are being distributed without warranty of any kind, either express or implied regarding their content. The responsibility for their

	interpretation and use lies with the user. In no event shall the World Health Organization be liable for damages arising from their use
Dataset language	English
Dataset character set	ASCII
Metadata provider	WHO Mediterranean Centre for Health Risk Reduction (WMC)
Metadata contact	El Morjani Zine El Abidine BP 3566 Poste Talborjt 80000 Agadir Morocco Telephone: +212 528 28 55 30 email: elmorjaniz@gmail.com
Metadata date	20110401
Metadata language	English
Metadata character set	ASCII
Metadata standard	ISO 19115