Met Office

Oasis Platform for Climate and Catastrophe Risk Assessment – Asia

Bangladesh Tropical Cyclone Historical Event Set: Data Description Documentation

Hamish Steptoe, Met Office

Supported by:

Federal Ministry
for the Environment, Nature Conservation
and Nuclear Safety

based on a decision of the German Bundestag

oasis

Supported by:
Federal Ministry for the Environment, Nature Conservation and Nuclear Safety
based on a decision of the German Bundestag

Contents

Introduction

This document describes the data that forms the historical catalogue of Bangladesh tropical cyclones, part of the Oasis Platform for Climate and Catastrophe Risk Assessment – Asia, a project funded by the International Climate Initiative (IKI), supported by the Federal Ministry for the Environment, Nature Conservation and Nuclear Safety, based on a decision of the German Bundestag

The catalogue contains the following tropical cyclones:

Name	Landfall Date (DD/MM/YYYY HH:MMZ)	IBTrACS ID
BOB01	30/04/1991 00:00Z	1991113N10091
BOB07	25/11/1995 09:00Z	1995323N05097
TC01B	19/05/1997 15:00Z	1997133N03092

Supported by:

Federal Ministry for the Environment, Nature Conservation and Nuclear Safety

based on a decision of the German Bundestag

Akash	14/05/2007 18:00Z	2007133N15091
Sidr	15/11/2007 18:00Z	2007314N10093
Rashmi	26/10/2008 21:00Z	2008298N16085
Aila	25/05/2009 06:00Z	2009143N17089
Viyaru	16/05/2013 09:00Z	2013130N04093
Roanu	21/05/2016 12:00Z	2016138N10081
Mora	30/05/2017 03:00Z	2017147N14087
Fani	04/05/2019 06:00Z	2019117N05088
Bulbul	09/11/2019 18:00Z	2019312N16088

Ensemble Configuration

Each tropical cyclone comprises of a nine-member, 3-hourly time-lagged ensemble. Each ensemble member covers a period of 48 hours, once the initial 24-hour model spin is removed (see Figure 1 for a visual representation), at resolutions of 4.4km and 1.5km based on the Met Office Unified Model dynamically downscaling ECMWF ERA5 data. Note that there is no data assimilation process in these model runs.

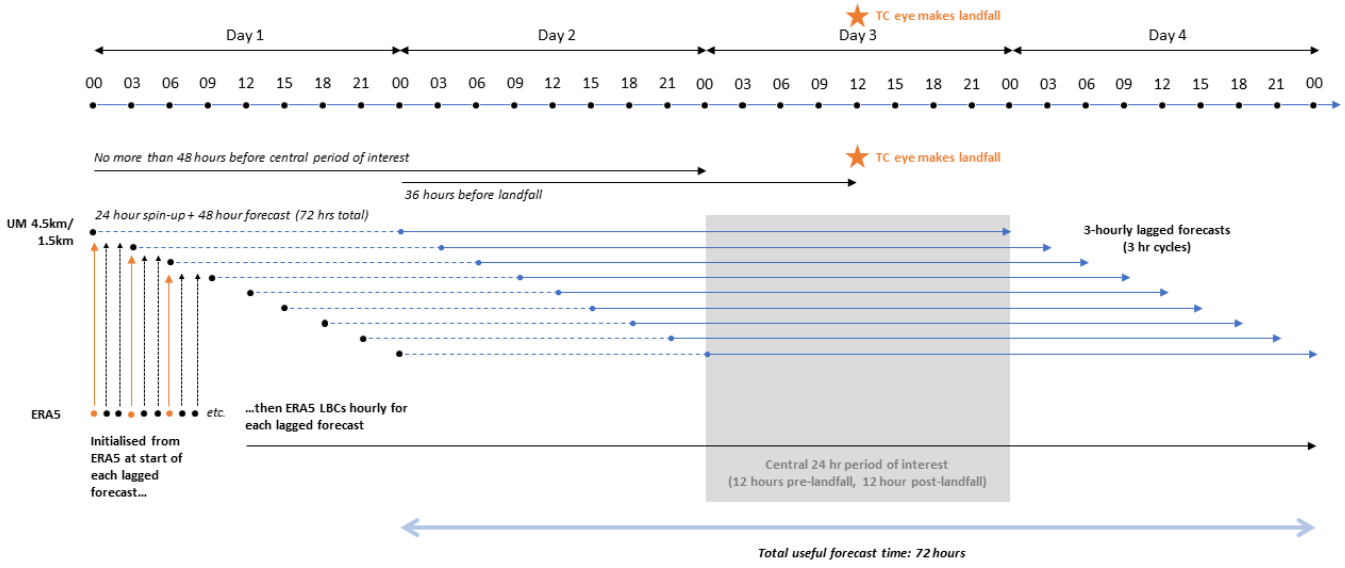

Figure 1 Ensemble configuration for the historical event set.

Domain

Model output is available at 4.4km (0.0405°) and 1.5km (0.0135°) over two nested domains (see Fig. 2): the 4.4km domain avoids placing model boundaries over the Himalayas and covers Nepal, Bhutan, Myanmar, most of India, and parts of the Tibetan plateau so as to facilitate hydrological modelling of the Ganges-Brahmaputra-Meghna river basins; the 1.5km domain is limited to Bangladesh only.

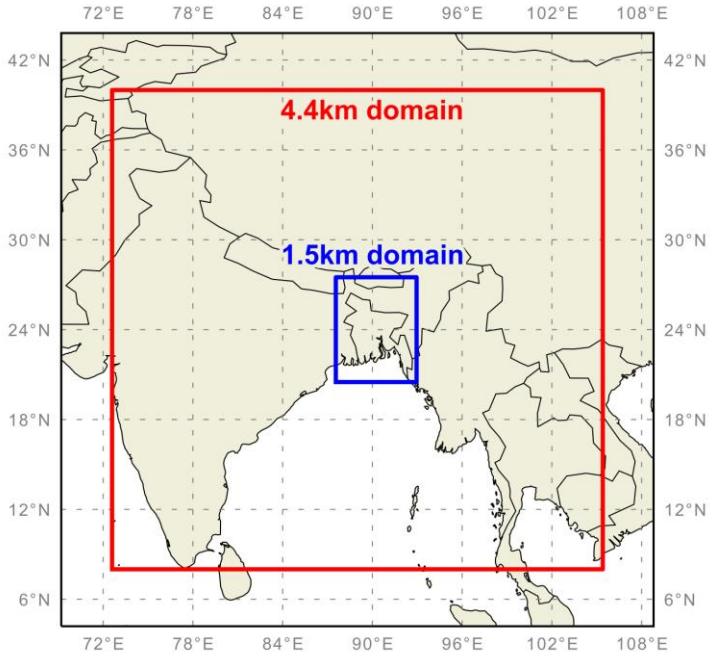

Figure 2 Model domains used for the 4.4km and 1.5km regional models. ERA5 data, with global coverage, provides initial conditions for the 4.4km domain. The 1.5km model takes its initial conditions for the 4.4km model.

Data Categories

The data in this catalogue are grouped into the following categories:

Name	Identifier	Description
Time series ensemble	tsens	Time series data for each ensemble member. Dimensions are typically: forecast_period, forecast_reference_time, latitude and longitude.
Ensemble footprints	fpens	Time-aggregated data for each ensemble member. Dimensions are typically: forecast_reference_time, latitude and longitude.
Best footprint	fp	A single best estimate footprint with lower, median and upper bounds accounting for ensemble variation.

Note that some variables exist on multiple pressure levels. For these cases, an additional 'pressure' dimension may also be present.

See the Data Metadata section for additional description of the dimensions.

Additional details covering the production of the footprint (fp) files will be available in a separate document.

Data Inventory

The follow data files are made available on Oasis Hub: https://oasishub.co/dataset/bangladesh-tropical-cyclone-historical-catalogue

File Name	Data Volume		Description
	Compressed	Un-compressed	
fpens.fg.T1Hmax.tar.gz	98MB	381MB	Ensemble Footprints - Wind Speed of Gust – 1-Hourly Maximum
fpens.psl.T1Hmin.tar.gz	114MB	352MB	Ensemble Footprints - Air Pressure at Sea Level – 1-Hourly Minimum
fpens.ua.T1Hpoint.tar.gz	77MB	364MB	Ensemble Footprints - X Wind – Hourly Instantaneous Measurement
fpens.va.T1Hpoint.tar.gz	75MB	364MB	Ensemble Footprints - Y Wind – Hourly Instantaneous Measurement
tsens.AILA.tar.gz	8.6GB	36.3GB	Cyclone Aila 2009 Time Series
tsens.AKASH.tar.gz	8.4GB	36.3GB	Cyclone Akash 2007 Time Series
tsens.BOB01.tar.gz	8.2GB	36.3GB	Cyclone BOB01 1991 Time Series
tsens.BOB07.tar.gz	7.7GB	36.3GB	Cyclone BOB07 1995 Time Series
tsens.BULBUL.tar.gz	8.0GB	37.3GB	Cyclone Bulbul 2019 Time Series
tsens.FANI.tar.gz	8.2GB	35.8GB	Cyclone Fani 2019 Time Series
tsens.MORA.tar.gz	8.9GB	37.5GB	Cyclone Mora 2017 Time Series
tsens.RASHMI.tar.gz	8.1GB	36.3GB	Cyclone Rashmi 2008 Time Series
tsens.ROANU.tar.gz	8.3GB	36.3GB	Cyclone Roanu 2016 Time Series
tsens.SIDR.tar.gz	8.2GB	37.5GB	Cyclone Sidr 2007 Time Series
tsens.TC01B.tar.gz	8.3GB	36.3GB	Cyclone TC01B 1997 Time Series
tsens.VIYARU.tar.gz	8.2GB	36.3GB	Cyclone Viyaru 2013 Time Series
set1.zip	953MB	4.9GB	Oasis format files (ensemble set)
set2.zip	98MB	514MB	Oasis format files (aggregated set)
TOTAL	**100GB**	**445GB**	

File Naming

Model output time-series (tsens) files are named according to the following convention:

VAR.TIMEMETHOD.UMRA2T.TIMEPERIOD.NAME.RES.nc

where:

VAR	A short variable identifier of the variable contained within the netCDF file. See full list of variables below.
TIMEMETHOD	The time method, specifying if the var is a mean, min, max or point and the period of time over which the mean, min, max or point measure is found. See below for a more detailed description.
UMRA2T	The Met Office regional model type. This is constant.
TIMEPERIOD	The time period that the data spans, in the form START_END formatted as YYYYMMDD.
NAME	The common name of the storm for the given time period.
RES	The resolution of the dataset. Either 4p4km = 4.4km or 1p5km = 1.5km grid size.
.nc	Indicates that this is a netCDF file.

Files relating to ensemble or best footprints have a simpler file naming structure:

fpens.VAR.TIMEMETHOD.NAME.RES.nc

or

fp.VAR.TIMEMETHOD.NAME.RES.nc

Variables

The full list of variables available, their shortened identifiers and their units are:

Variable	Identifier	Unit
net down surface sw flux corrected	rsnds	W/m^2
wet bulb potential temperature	wbpt	K
air pressure at sea level	psl	Pa
air temperature	tas	K
geopotential height	zg	m
relative humidity	hur	%
stratiform rainfall amount	prlst	kg/m^2
stratiform snowfall amount	prlssn	kg/m^2

surface downwelling shortwave flux in air	rsds	W/m^2
wind speed of gust	fg	m/s
x wind	ua	m/s
y wind	va	m/s

Note that where possible the variable names follow the CF Standard Names, http://cfconventions.org/standard-names.html.

Time Methods

Time methods are defined by the sampling period of the data and the sampling type applied to this period. The sampling period (or sampling interval) is one of: hourly (T1H), 3-hourly (T3H) or 24-hourly (T24H).

The sampling type is one of max (maximum), min (minimum), mean or point. Point sampling is an instantaneous sample taken from the model time-step (which is typically much less than the sample period) and is the sampling type that is most closely comparable to observational data.

Together then, T1Hmax is interpreted as hourly maximum data; T3Hmean is interpreted a 3-hourly mean data, and T1Hpoint are instantaneous measurements taken every hour.

Data Metadata

Each netCDF file has metadata that describes the data file. The metadata should be considered the definitive source of file information. **For a full understanding of the netCDF metadata conventions, please refer to:**
http://cfconventions.org/cf-conventions/cf-conventions.html.

What follows is a short description of the output from the UCAR Unidata ncdump utility, which prints this metadata to screen.

As an example:

```
$ ncdump -h psl.T1Hmin.UMRA2T.19910428_19910501.BOB01.1p5km.nc
```

returns:

```
netcdf psl.T1Hmin.UMRA2T.19910428_19910501.BOB01.1p5km {
dimensions:
```

```
       forecast_period = 48 ;
       forecast_reference_time = 9 ;
       latitude = 518 ;
       longitude = 400 ;
       bnds = 2 ;
variables:
       float air_pressure_at_sea_level(forecast_period, forecast_reference_time,
latitude, longitude) ;
              air_pressure_at_sea_level:standard_name = "air_pressure_at_sea_level" ;
              air_pressure_at_sea_level:units = "Pa" ;
              air_pressure_at_sea_level:um_stash_source = "m01s16i222" ;
              air_pressure_at_sea_level:cell_methods = "time: minimum (interval: 1
hour)" ;
              air_pressure_at_sea_level:grid_mapping = "latitude_longitude" ;
              air_pressure_at_sea_level:coordinates = "time" ;
       int latitude_longitude ;
              latitude_longitude:grid_mapping_name = "latitude_longitude" ;
              latitude_longitude:longitude_of_prime_meridian = 0. ;
              latitude_longitude:earth_radius = 6371229. ;
              latitude_longitude:proj4 = "+proj=longlat +ellps=WGS84 +datum=WGS84
+no_defs" ;
       double forecast_period(forecast_period) ;
              forecast_period:bounds = "forecast_period_bnds" ;
              forecast_period:units = "hours" ;
              forecast_period:standard_name = "forecast_period" ;
       double forecast_period_bnds(forecast_period, bnds) ;
       double forecast_reference_time(forecast_reference_time) ;
              forecast_reference_time:axis = "T" ;
              forecast_reference_time:units = "hours since 1970-01-01 00:00:00" ;
              forecast_reference_time:standard_name = "forecast_reference_time" ;
              forecast_reference_time:calendar = "gregorian" ;
       float latitude(latitude) ;
              latitude:axis = "Y" ;
              latitude:units = "degrees_north" ;
              latitude:standard_name = "latitude" ;
       float longitude(longitude) ;
              longitude:axis = "X" ;
              longitude:units = "degrees_east" ;
              longitude:standard_name = "longitude" ;
       double time(forecast_reference_time, forecast_period) ;
              time:bounds = "time_bnds" ;
              time:units = "hours since 1970-01-01 00:00:00" ;
              time:standard_name = "time" ;
              time:calendar = "gregorian" ;
       double time_bnds(forecast_reference_time, forecast_period, bnds) ;
```

```
// global attributes:
              :rim_removed = "13 point rim removed" ;
              :um_version = "11.1" ;
              :source = "Copernicus Climate Change Service Information (C3S) ECMWF ERA5
/ Met Office UM RA2T CON" ;
              :Conventions = "CF-1.7" ;
              :contact = "enquiries@metoffice.gov.uk" ;
              :comment = "Supported by the International Climate Initiative (IKI) and
the Federal Ministry for the Environment, Nature Conservation and Nuclear Safety,
based on a decision of the Germany Bundestag" ;
              :data_type = "grid" ;
              :date_created = "20192730T15:27:10" ;
              :geospatial_lat_max = "27.483002" ;
              :geospatial_lat_min = "20.503502" ;
              :geospatial_lat_resolution = "0.01" ;
              :geospatial_lat_units = "degrees_north" ;
              :geospatial_lon_max = "92.942" ;
              :geospatial_lon_min = "87.5555" ;
              :geospatial_lon_resolution = "0.01" ;
              :geospatial_lon_units = "degrees_east" ;
              :history = "(1.0) Initial release" ;
              :id = "psl.T1Hmin.UMRA2T.19910428_19910501.BOB01.1p5km.nc" ;
              :institution = "Met Office, UK" ;
              :keywords = "Bangladesh, dynamical downscaling, RA2T, Met Office" ;
              :product_version = "v1.0" ;
              string :project = "Oasis Platform for Climate and Catastrophe Risk
Assessment – Asia" ;
              :references = "" ;
              :standard_name_vocabulary = "NetCDF Climate and Forecast (CF) Standard
Names version 51" ;
              :summary = "Tropical cyclone data over Bangladesh downscaled using Met
Office RA2T_CON initiated from ERA5" ;
              :title = "Downscaled Tropical Cyclone data over Bangladesh" ;
              :type = "float" ;
              :spatial_resolution = "4.4km" ;
              :licence = "Creative Commons Attribution 4.0 International (CC BY 4.0)" ;
}
```

Notice that this is split into several sections, and describes the data as follows:

Dimensions

```
dimensions:
    forecast_period = 48 ;
```

```
    forecast_reference_time = 9 ;
    latitude = 518 ;
    longitude = 400 ;
    bnds = 2 ;
```

There are 5 dimensions used within this file:

- `forecast_period` - the time step of the data, in this case this dimension has length 48. This describes the time dimension of the data in hours relative to the forecast reference time.
- `forecast_reference_time` – the initialisation time of each model run. This dataset is made of 9 ensemble members, each of which are initialised at a different start time in units of 'hours since' a defined reference period (defined in the netCDF metadata), using the Gregorian calendar.
- `latitude` – The latitudinal position of each grid point relative to EPSG:4326 (WGS84)
- `longitude` - The latitudinal position of each grid point relative to EPSG:4326 (WGS84)
- `bnd` - the number of bounds the data has. In this case, two representing an upper and lower bound.

Additional dimensions that may be present in other files include:

- `pressure` – a vertical dimension, representing the pressure level at which the data exists, in units of hPa (hectopascals).

Variables

There is a long list of variables. Notice that each of these has a number of attributes associated with it, and information describing its dimensions. For example:

```
float air_pressure_at_sea_level(forecast_period, forecast_reference_time, latitude,
longitude) ;
        air_pressure_at_sea_level:standard_name = "air_pressure_at_sea_level" ;
        air_pressure_at_sea_level:units = "Pa" ;
        air_pressure_at_sea_level:um_stash_source = "m01s16i222" ;
        air_pressure_at_sea_level:cell_methods = "time: minimum (interval: 1 hour)" ;
        air_pressure_at_sea_level:grid_mapping = "latitude_longitude" ;
        air_pressure_at_sea_level:coordinates = "time" ;
```

Describes the variable `air_pressure_at_sea_level`, shows that the data is of type float, and is described using 4 of the 5 available dimensions (`forecast_period, forecast_reference_time, latitude, longitude`). There follows additional metadata specific to this variable, such as the `units`, and `standard_name`.

Global Attributes

The global attributes contain additional meta data information applicable to the whole file. Note that this includes the data licence, product version number and project information.

Oasis LMF Files

Binary files compatible with the Oasis loss model framework (LMF) have been derived from the *fpens.fg.T1Hmax* netCDF files. This is a two-step process, firstly converting the binary netCDF arrays into a flat CSV file, then using ktools to convert these into the Oasis binary format.

We provide two sets of the historical catalogue in this format: (1) assuming the ensemble of footprints are independent of each other (set 1, also referred to as the 'ensemble' set), and (2) footprints aggregated by ensemble with probabilities for the intensity bins derived from a naïve Bayesian interpretation the posterior predictive distribution of ensemble spread (set 2, also referred to as the 'aggregated' set). Steptoe et al. (2020) has further details on how this set was constructed.

Set 1 .csv files are preceded by `hc_oasis` and set 2 .csv files by `hcprob_oasis`.

In each case, each set of Oasis compatible files comprises of:

- `input/events.bin`
- `input/occurrence.bin`
- `static/footprint.bin`
- `static/footprint.idx`

Additional CSV files `*_areaperil_dict.csv` and `*_event_dict.csv` provide lookup tables relating unique AREAPERIL_IDs with their longitude and latitude values, and EVENT_ID with their storm names. File `*_intensity_bin_dict.csv` relates intensity bins with their corresponding gust speeds.

Note that `input/coverages.bin`, `input/items.bin`, `static/vulnerability.bin` and `static/damage_bin_dict.bin` files are provided solely for the purposes of ktools workflow validation only. These files contain dummy input and are not intended to be used for cat modelling purposes. These files do not constitute output of the IKI project. Users should replace these files with their own input data before using the Oasis loss model. In particular, a relationship between damage and intensity should not be derived from these files.

For the aggregated probabilistic event set representation (ie. set 2), probabilities are represented as 32-bit floating point numbers. Floating point representation of fractional probabilities reduces data fidelity, and in some cases the probabilities for a given AREA_PERIL_ID do not sum to 1. This only occurs in 2409 locations (across all 12 events), equivalent to 0.03% of possible AREA_PERILs. In these cases, they are typically ≥ 0.992.

ktools conversion

The ktools binary format conversion uses the following commands

Set 1

```
damagebintobin < hc_oasis_damage_bin_dict.csv > static/damage_bin_dict.bin
evetobin < hc_oasis_events.csv > input/events.bin
occurrencetobin -P 261 < hc_oasis_occurrence.csv > input/occurrence.bin
footprinttobin -i 1100 -n < hc_oasis_footprint.csv > static/footprint.bin
```

Set2

```
damagebintobin < hcprob_oasis_damage_bin_dict.csv > static/damage_bin_dict.bin
evetobin < hcprob_oasis_events.csv > input/events.bin
occurrencetobin -P 29 < hcprob_oasis_occurrence.csv > input/occurrence.bin
footprinttobin -i 110 < hcprob_oasis_footprint.csv > static/footprint.bin
```

Note that the number of intensity bins for set 2 footprints is reduce to 110 (i.e. 1 m/s bins).

Further details about the ktools conversion process can be found at https://github.com/OasisLMF/ktools/blob/develop/docs/md/DataConversionComponents.md#dataconversioncomponents

References

Steptoe, H. and Economou, T.: Improving our understanding of wind extremes from Bangladesh tropical cyclones: insights from a high-resolution convection-permitting numerical model, Nat. Hazards Earth Syst. Sci., (submitted), doi:10.5194/nhess-2020-299, 2020.

Supported by:
Federal Ministry for the Environment, Nature Conservation and Nuclear Safety
based on a decision of the German Bundestag

Document Control		
	Role Title	**Name**
Document Owner	Project Manager	Claire Souch (Oasis)
Author	Project Scientist	Hamish Steptoe
Reviewers	Project Manager Project Scientist Project Scientist	Claire Souch (Oasis) Kate Salmon Bernd Becker

Version	Date published	Review due	Change description
1.0	7/10/2019	-	Initial version
1.1	-	-	Added licence information. Added Data Inventory section.
1.2	-	-	Added information for Cyclone Bulbul. Added IBTrACS IDs for all storms.
1.3	13/05/2020	-	Added information covering Oasis binary files. Added information on the modelling domain.
1.4	4/12/2020	-	Added reference to Steptoe et al. paper. Minor changes to the names and descriptions of Set 1 and Set 2 events.

Met Office
FitzRoy Road, Exeter
Devon EX1 3PB
United Kingdom

Tel (UK): 0370 900 0100 (Int) : +44 1392 885680
Fax (UK): 0370 900 5050 (Int) :+44 1392 885681
enquiries@metoffice.gov.uk
www.metoffice.gov.uk