Note on the ATC-Caption-2025 dataset

trise5631@gmx.com, jean-luc.marchand@mh370-caption.net

Version 1.4 – 09/10/2025

The reader is encouraged to contact the authors by email if he or she has a specific idea about a particular and relevant use of the raw data.

Presentation of the ATC-Caption-2025 Dataset

The decoding explanations and the detailed quality analysis of the ATC-Caption-2025 dataset are well documented in the recently published report "Analysis of the 7 march 2014 raw ATC radar data" by Trise5631, Jean-Luc Marchand and Captain Blelly from CAPTION.

Both the dataset and the detailed report are available at www.mh370-caption.net

The ATC-Caption-2025 dataset includes the following data:

1. Data directly concerning flight MH370 are gathered in the folder « MH370 CSV » which includes 11 CSV files. Each file contains target reports data of one type only and from one source only as presented in the following table:

ADSB	Genting	Genting_ADSB_MH370.csv
	KT	KT_ADSB_MH370.csv
SSR	BW	BW_SSR_MH370.csv
	BW*	BW_SSR_MH370_corrected.csv
	Genting	Genting_SSR_MH370.csv
	KB	KB_SSR_MH370.csv
	Kuantan	Kuantan_SSR_MH370.csv
	Senai	Senai_SSR_MH370.csv
	Subang	Subang_SSR_MH370.csv
	Langkawi	Langkawi_SSR_MH370.csv
PSR	BW	BW_PSR_MH370.csv
	BW*	BW_PSR_MH370_corrected.csv
	KB-in	KB_PSR_IN_MH370.csv
	KB-out	KB_PSR_OUT_MH370.csv

* Same data, but "theta" values have been corrected for the azimuth angle bias estimated by comparing SSR measurements with ADS-B data. "Lat" and "lon" values are changed accordingly, but "x", "y", "lat2", and "lon2" values are retained.

KB-in and KB-out designate the track when the aircraft flew towards Kota Bharu radar and away from it respectively.

2. Data from Comm-B messages from MH370 received by Senai and KB “Mode S” SSR radars are stored in the “CommB” folder. One binary data file is provided per BDS type. These data are also encompassed in the CSV and JSON files.
3. Data of all identified flights within the time window from 16h to 19h UTC. One csv file has been created per aircraft and it includes the aircraft position every 10 seconds. The full set is stored in the “Merging” folder.
4. For the sake of universality, usefulness and completeness, JSON formatted files are provided including MH370 data coming from Genting and KT ADSB as well as from Genting, Kuantan, BW, Langkawi, KB and Senai SSR/PSR. They can be found in the JSON folder. Service messages data is also provided as it is an essential information for timing and position determination and in particular in BW (Butterworth) data.
5. ADS-B data from KT is also provided in the RAW format.
6. Additionally, to help reconstructing the trajectory, a CSV file is provided including the wind and temperature information around MH370 flight during the Mode-S coverage of KB and Senai as well as the same type of information but provided by the Copernicus ERA5 data[1].
7. Furthermore, for completeness, data concerning the flights mentioned in the detailed analysis report is provided in CSV format. They are namely VH-PFS, CES5093, BBC086, SIA68, UAE343 and IGO053. Even though these flights did not see MH370, they were most likely seen by MH370 and somehow influenced the decisions of the person in command.
8. An Excel file containing context data from all the flights, with some details on the aircraft

1. MH370 CSV files

Data from target positions of MH370 from Genting, Subang, Kuantan, BW, Langkawi, KB and Senai SSR, and Genting and KT ADS-B receivers, as well as BW and KB PSR after IGARI.

[1] The results contain modified Copernicus Climate Change Service information 2020. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains.
Copernicus Climate Change Service, Climate Data Store, ERA5 hourly data on pressure levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), 2023.
Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J-N. (2023): ERA5 hourly data on pressure levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), DOI: 10.24381/cds.bd0915c6

For radar files:

Name	*Description*	*Unit*
ToD	Time of day	Seconds
type	Report type	➔
rho	Slant range	NM
theta	Azimuth	°
x	X coordinate	NM
y	Y coordinate	NM
FL	Flight level	
CGS	Tracker calculated ground speed	Knots
CHdg	Tracker calculated heading	°
lat	Latitude (from rho/theta)	Degrees
lon	Longitude (from rho/theta)	Degrees
lat2	Latitude (from x/y)	Degrees
lon2	Longitude (from x/y)	Degrees
MCPAlt	MCP selected altitude*	Feet
roll	Roll angle*	°
trackAngle	Track angle*	°
trackAngleRate	Track angle rate*	°/s
GS	Ground speed*	Knots
heading	Heading*	°
IAS	Indicated Air Speed*	Knots
TAS	True Air Speed*	Knots
mach	Mach*	
modeA	Mode A code	
CS	Call Sign*	
ICAO24	ICAO24 address	
40	BDS 4,0 content	
50	BDS 5,0 content	
60	BDS 6,0 content	
10	BDS 1,0 content	
17	BDS 1,7 content	
1d	Unclear	
spf1	Special Field 1	
spf2	Special Field 2	
spf3	Special Field 3	
spf4	Special Field 4	
spf5	Special Field 5	
spf6	Special Field 6	
spf7	Special Field 7	
spf8	Special Field 8	

type	*Report type*
0	No detection
1	Single PSR detection
2	Single SSR detection
3	SSR + PSR detection
4	Single ModeS All-Call
5	Single ModeS Roll-Call
6	ModeS All-Call + PSR
7	ModeS Roll-Call +PSR

*: data extracted from BDS content

For ADS-B files:

Name	*Description*	*Unit*
ToD	Time	Seconds
ModeA	Mode A code	
lat	Latitude	Degrees
lon	Longitude	Degrees
galt	Geometric altitude	Feet
FL	Flight level	
BVR	Barometric vertical rate	Feet/min
GS	Ground Speed	Knots
TA	Track Angle	Degrees
CS	Call sign	
tas	True Airspeed	Knots
mh	Magnetic heading	Degrees
trd	Target Report Descriptor	
DCR	Differential correction bit	
GBS	Ground bit	
SIM	Simulation bit	
TST	Test bit	
RAB	Fixed transponder bit	
SAA	Selected altitude bit	
SPI	Special Position Identification bit	
ATP	Address Type	
ARC	Altitude Reporting Capability	
fom	Figure of merit	
AC	ACAS State	
MN	Multiple navigational aids	
DC	Differential correction	
PA	Position Accuracy	
DTI	Cockpit Display of Traffic Information	
MDS	Mode-S Extended Squitter	
UAT	UAT	
VDL	VDL mode 4	
OTR	Other technology	
VA	Velocity Accuracy	
Sig	Signal Amplitude	0-255[2]
ec	Emitter Category	
SP	Special Field	

[2] The amplitude seems be reversed (maximum signal for Sig=0), the variations in KT values may be due to the fact that the KT ADS-B system uses multiple receivers.

2. MH370 Comm-B Messages

Data from Comm-B messages from MH370 received by Senai and KB Mode S SSR radars.

The CSV files columns are:

1) Time of Day
2) 56 bit Comm-B message in hexadecimal

The messages are separated by BDS codes, one file for each:

- BDS code 1,0 — Data link capability report
- BDS code 1,D — unknown, possibly used by the radar to transmit other data
- BDS code 1,7 — Common usage GICB capability report
- BDS code 4,0 — Selected vertical intention
- BDS code 5,0 — Track and turn report
- BDS code 6,0 — Heading and speed report

3. Merged files

"Merged files" are CSV files produced using the original ASTERIX data. Each file contains the interpolated position every 10s of the same aircraft during the recording, based on multiple sources. As some aircraft serve multiple flights during the 6-hour recording, the corresponding files have multiple flight numbers. Some aircraft could only be identified by their mode A (flights that are just tracked by “mode A” SSR during the recordings -> names starting with a “!”), others only by their flight numbers (as the ADS-C data does not contain ICAO24 addresses -> names starting with a “+”). The files include some ground vehicles, for which no flight number is associated, therefore, these “flights” only contain the ICAO24 address. The columns are:

Column	*Description*
1	Time of day of the interpolation
2	Source of the closest data point
3	Latitude
4	Longitude
5	Ground Speed
6	Track Angle
7	Time from the closest data point

For several flights for which the flight route is known, the data has been extrapolated using the hypothesis of a constant Mach, based on GDAS weather data.

The source code numbers correspond to:

Code	*Description*
1	ADS-B (KT or Genting) data
2	ADS-C data (mainly in the Northwest of the FIR)
3	Langkawi Mode A/C SSR data
4	KT and Senai Mode A/C/S SSR data, (KB/BW PSR radar for MH370)
5	MH370 trajectory: Geoscience points before 18h UTC, Blelly/Marchand trajectory after.
6	AMDAR data (Aircraft Meteorological Data Relay), ACARS positions every 6-8 minutes (only for SAA287 and DLH779)

This is also the order of priority of interpolation.

Other sources have been excluded, as they would only increase the coverage of some flights at low altitudes, during take-off or landing.

The file !!flightIDs.csv contains all the flight numbers associations. The columns are:

Column	*Description*
1	ICAO24 address of the aircraft
2	Flight number
3	First Time the aircraft sends this flight number
4	Last Time the aircraft sends this flight number

4. JSON files

The "JSON files" are transcripts of ASTERIX messages from the Genting and KT ADS-B stations, as well as the Genting, Kuantan, Subang, Langkawi, Senai, KB, and BW PSRs/SSRs.

The "JSON files" contain all service messages between 16:00 UTC and 19:00 UTC, all target reports bearing either Mode A code 2157, ICAO24 address 75008f, or runway number 1744 or 1827 (for KB entries and exits). In addition, they contain points without Mode A codes that appear to originate from MH370 (BW PSRs and some SSR positions without Mode A codes).

Each line of the file is a JSON message corresponding to an ASTERIX message. In the case of BW, some messages appear to have been repeated. Additionally, keep in mind that BW target reports do not contain time-of-day information, and that the time must be calculated from the theta angle value, the northbound service message, and the rotation rate, keeping in mind that the north crossing service messages are ahead of the target reports in the recordings, so much that a target report with a theta of 213° or more is recorded after the north crossing service message of the next rotation. Other service messages also seem to be out of sync.

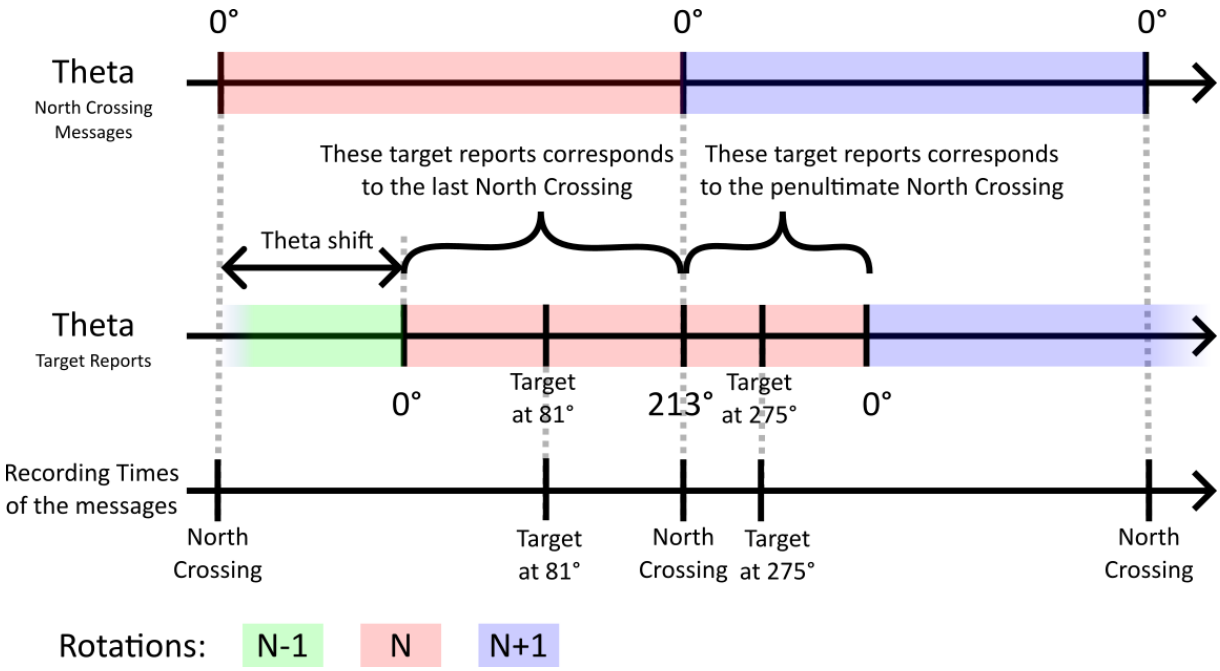

Figure 1 Illustration of the shift between the north crossing messages and the target reports

The structure and labels of the items are similar to those of the CroatiaControl ASTERIX parser, but the data was produced by a specifically in-house designed Python parser.

With the exception of ADS-B KT, for which information is already publicly available, the SIC and SAC codes have been removed. It should be noted that, in the dataset, these codes do not comply with ICAO code assignments.

5. ADS-B KT files

The ADSB_KT.raw file contains the raw data of the entire recording of the ADS-B KT receiver. The messages are in ASTERIX CAT021 version 0.26 format.

The ADSB_KT_MH370_hex.txt file contains all ASTERIX CAT021 records for flight MH370. Each CAT021 message (called a data block) contains multiple records for different aircraft. This file is an extract of all MH370 records from the ADSB_KT.raw file, and was added for ease of use. In practice, it could be parsed manually by treating each line as a CAT21 message without the data category octet.

Additionally, two PNG screen snapshots of the KT system from 2013 from a Prezi presentation are available.

6. MH370 Mode S wind CSV files

Wind and temperature values calculated from Mode S data from KB SSR.

Name	*Description*	*Unit*
ToD	Time of day	Seconds
FL	Flight level	
lat2	Latitude (from x/y)	Degrees
lon2	Longitude (from x/y)	Degrees
GS	Ground speed*	Knots
trackAngle	Track angle*	°
TAS	True Air Speed*	Knots
heading	Heading*	°
WindSpd	Wind Speed*	Knots
WindDir	Wind Direction*	°
WindNS	Wind Speed NS*	Knots
WindEW	Wind Speed EW*	Knots
temp	Temperature*	Kelvin
windSpdERA5	Wind Speed (from ERA5)	Knots
windDirERA5	Wind Direction (from ERA5)	°
tempERA5	Temperature (from ERA5)	Kelvin

*: data extracted/calculated from BDS content

Wind speed and direction are calculated by subtracting the groundspeed vector (GS and track angle) from the airspeed vector (TAS and heading). Temperature is calculated from the Mach and TAS values. ERA5 values are interpolated (4D linear interpolation) with the pressure level calculated with the flight level.

7. Influential flights CSV files

Data concerning the flights mentioned in the detailed analysis report is provided in CSV format. They are namely VH-PFS, CES5093, BBC086, SIA68, UAE343 and IGO053. Even though these flights did not see MH370, they were most likely seen by MH370 and somehow influenced the decisions of the person in command.

8. Flights and aircraft details

This Excel file (Flights.xlsx) contains the list of flights tracked during the recordings. The first sheet contains data based on Mode S (Senai and KB) and ADS-B (Genting and KT) data, the second sheet contains data originating from the ATC system (which could contain flights scheduled before or after the recordings).