



HAL
open science

Multi-source datasets acquired over Toulouse (France) in 2021 for urban microclimate studies during the CAMCATT/AI4GEO field campaign

Laure Roupioz, Xavier Briottet, K Adeline, A Al Bitar, Delphine Barbon-Dubosc, R Barda-Chatain, Philippe Barillot, S Bridier, E Carroll, Charlène Cassante, et al.

► To cite this version:

Laure Roupioz, Xavier Briottet, K Adeline, A Al Bitar, Delphine Barbon-Dubosc, et al.. Multi-source datasets acquired over Toulouse (France) in 2021 for urban microclimate studies during the CAMCATT/AI4GEO field campaign. *Data in Brief*, 2023, 48, 10.1016/j.dib.2023.109109 . hal-04064296

HAL Id: hal-04064296

<https://hal.science/hal-04064296>

Submitted on 11 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License



Data Article

Multi-source datasets acquired over Toulouse (France) in 2021 for urban microclimate studies during the CAMCATT/AI4GEO field campaign



L. Roupioz^{a,*}, X. Briottet^{a,*}, K. Adeline^a, A. Al Bitar^f,
 D. Barbon-Dubosc^a, R. Barda-Chatain^a, P. Barillot^a, S. Bridier^h,
 E. Carroll^h, C. Cassante^a, A. Cerbelaud^a, P. Déliot^a, P. Doublet^a,
 P.E. Dupouy^a, S. Gadal^h, S. Guernouti^b, A. De Guilhem De
 Lataillade^g, A. Lemonsu^e, R. Llorens^d, R. Luhache^c, A. Michel^a,
 A. Mousous^a, M. Musy^b, F. Nerry^c, L. Poutier^a, A. Rodler^b,
 N. Riviere^a, T. Riviere^a, J.L. Roujean^f, A. Roy^e, A. Schilling^a,
 D. Skokovic^d, J. Sobrino^d

^a ONERA DOTA, Université de Toulouse, 2 Av. Edouard Belin BP74025, 31055, Toulouse Cedex 4, France

^b Equipe de Recherche BPE, Cerema & LOCIE - UMR 5271, CNRS & Université Savoie, 9 Rue René Viviani BP 46223, 44262, Nantes, France

^c ICube Laboratory, UMR 7357, CNRS-University of Strasbourg, 300 Bd Sébastien Brant CS 10413, 67412, Illkirch Cedex, France

^d Global Change Unit, University of Valencia, C/Catedrático José Beltrán 2, ES-46980 Paterna, Spain

^e CNRM, Centre National de Recherches Météorologiques, Université de Toulouse, Météo-France, CNRS, 42 Av. Gaspard Coriolis, 31057, Toulouse Cedex 1, France

^f CESBIO, Centre d'Etudes Spatiales de la Biosphère, CNES, CNRS, INRA, IRD, UPS, 18 Av. Edouard Belin BPI 2801, 31401, Toulouse Cedex 9, France

^g DGA-TA, 47 rue Saint-Jean, Balma, 31044, France

^h Aix Marseille Univ, Université Côte d'Azur, Avignon Université, CNRS, ESPACE, UMR 7300, 74 rue Louis Pasteur, 84029, Avignon Cedex, France

* Corresponding authors.

E-mail addresses: laure.roupioz@onera.fr (L. Roupioz), xavier.briottet@onera.fr (X. Briottet).

ARTICLE INFO

Article history:

Received 1 March 2023

Revised 24 March 2023

Accepted 27 March 2023

Available online 3 April 2023

Dataset link: [CAMCATT-AI4GEO \(Original data\)](#)*Keywords:*

Land surface temperature

Spectral emissivity

Spectral reflectance

Air temperature

Airborne LiDAR

Atmospheric data

Urban area

ABSTRACT

The CAMCATT-AI4GEO extensive field experiment took place in Toulouse, a city in the southwest of France, from 14th to 25th June 2021 (with complementary measurements performed on the 6 September 2021). Its main objective was the acquisition of a new reference dataset on an urban site to support the development and validation of data products from the future thermal infrared (TIR) satellite missions such as TRISHNA (CNES/ISRO), LSTM (ESA) and SBG (NASA). With their high spatial (between 30–60m) and temporal (2–3 days) resolutions, the future TIR satellite data will allow a better investigation of the urban climate at the neighbourhood scale. However, in order to validate the future products of these missions such as LST, air temperature, comfort index and Urban Heat Island (UHI), there is a need to accurately characterise the organisation of the city in terms of 3D geometry, spectral optical properties and both land surface temperature and emissivity (LST and LSE) at several scales. In this context, the CAMCATT-AI4GEO field campaign provides a set of airborne VISNIR-SWIR (Visible Near InfraRed - ShortWave InfraRed) hyperspectral imagery, multispectral thermal infrared (TIR) imagery and 3D LiDAR acquisitions, together with a variety of ground data collected, for some of them, simultaneously to the flight. The ground dataset includes surface reflectance measured spectrally with ASD spectroradiometers and in six spectral bands spreading from shortwave to thermal infrared and for two viewing angles with a SOC410-DHR handheld reflectometer. It is completed with LST and LSE retrieved from thermal infrared radiance acquired in six spectral bands with CIMEL radiometers. It also includes meteorological data coming from four radio soundings (one of which was taken during the flight), data routinely collected at the Blagnac airport reference station as well as air temperature and humidity acquired using instrumented cars following two different itineraries. In addition, a link is provided to access the data routinely collected by the network of weather stations set up by Toulouse Metropole in the city and its surroundings. This data paper describes this new reference urban dataset which can be useful for many applications such as calibration/validation of at-surface radiance, LST and LSE data products as well as higher level products such as air temperature or comfort index. It also provides valuable opportunities for other applications in urban climate studies, such as supporting the validation of microclimate models.

© 2023 The Author(s). Published by Elsevier Inc.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Specifications Table

Subject	Earth-surface processes
Specific subject area	Radiometric and meteorological data collected over Toulouse (France) by ground and airborne sensors to create a reference dataset to study the urban microclimate
Type of data	Table Image (raster) Text file
How the data were acquired	<p>Most of the data were collected during the CAMCATT-AI4GEO experiment in Toulouse (France) from 14 to 25 June 2021. Several datasets were collected during this campaign:</p> <p>Airborne data acquired on the 15th June 2021 around 11am UTC, with three cameras on-board the Safire ATR42 aircraft (flight altitude 1500 m above ground level):</p> <ul style="list-style-type: none"> • Hyperspectral imagery acquired with a AISA FENIX 1K camera (spectral range 0.4 to 2.5 μm, swath 1024 m, Ground Sampling Distance (GSD) 1 m) • Multispectral Thermal Infrared (TIR) imagery acquired with a Telops camera equipped with a filter wheel in five spectral bands (8.1-8.9 μm, 8.7-9.9 μm, 10.5-12.9 μm, 6-14 (broadband) μm, 10-11.5 μm) (swath 593 \times 472 m², GSD 1.95 m) • 3D point cloud acquired with a high resolved 3D-LiDAR (RIEGL VQ-1560i) collocated with two cameras (RGB and Near-Infrared) (point density approximately 100 points/m²) <p>Ground data acquired during and outside the flight:</p> <ul style="list-style-type: none"> • Hyperspectral surface reflectance in the range 0.4-2.5 μm acquired with three ASD spectrometers using a spectralon as reference • Multispectral TIR radiance measured with three CIMEL 312 radiometers in six spectral bands (8-14 μm (broadband), 8.1-8.5 μm, 8.5-8.9 μm, 8.9-9.3 μm, 10.3-11 μm, 11-11.7 μm). The Temperature-Emissivity Separation (TES) method was applied to retrieve Land Surface Emissivity (LSE) and Land Surface Temperature (LST) using the downwelling irradiance estimated with an Infragold reference panel. The three CIMELs were intercalibrated at the ONERA laboratory over 14 samples (RMSE~1%) • Integrated reflectance measured with a spectroradiometer, SOC410, for two viewing angles (20 and 60°) and in six spectral bands (0.9-1.1 μm, 1.9-2.4 μm, 3.0-3.9 μm, 3.0-5.0 μm, 4.0-5.0 μm, 8.0-12.0 μm) • Air temperature and relative humidity acquired during and outside the flight by sensors (SHT85 sensirion + Thermocouple RS Pro type T) mounted on two vehicles circulating in the city <p>Atmospheric data:</p> <ul style="list-style-type: none"> • Radio sounding operated by Météo-France (14/06/21, 10h23 UTC, 15/06/21, 10h44 UTC, 24/06/21, 21h27 UTC, 25/06/21, 10h47 UTC) using a radiosonde VAISALA RS41, maximum measurement height 10 km with data acquired during both sensor ascent and descent • Blagnac operational weather station data provided hourly from 13 to 27 June (Lat. 43°37'15"N Lon. 1°22'43"E) <p>Secondary data:</p> <ul style="list-style-type: none"> • Meteorological data collected by the Toulouse Metropole network, which consists of around 60 DAVIS micro stations installed in Toulouse and its surrounding area to cover a wide variety of land covers. Each station measures the air temperature, pressure, humidity, precipitation, direction and wind speed

(continued on next page)

Data format	<p>Airborne data:</p> <ul style="list-style-type: none"> • Hyperspectral imagery: orthorectified and atmospherically corrected, ENVI file format • Multispectral TIR imagery: orthorectified and mosaicked, geoTIFF format • 3D point cloud georeferenced, LAZ format <p>Ground data:</p> <ul style="list-style-type: none"> • Spectral reflectance in the range 0.4-2.5 μm: raw, ASCII (.txt) • Spectral emissivity with CIMEL312: raw, ASCII (.txt) • Integrated emissivity with SOC 410: raw, ASCII (.txt) • Air temperature and relative humidity from vehicles: raw, ASCII (.txt) <p>Atmospheric data:</p> <ul style="list-style-type: none"> • Radio sounding: raw, netcdf file format • Blagnac RADOME operational station: raw, ASCII (.data) <p>Secondary data:</p> <ul style="list-style-type: none"> • Meteorological data from Toulouse Metropole weather stations network: raw, CSV file format
Description of data collection	<p>The dataset was acquired over Toulouse in June 2021 (completed in September 2021) with ground measurements collected within the urban area and airborne data covering the city and its rural surroundings. The data collection took place under clear sky conditions, at mid-day to ensure small sun zenith angles. The ground measurements were checked to eliminate noisy or corrupted data based on visual inspection of the data. No further filtering was applied to the data.</p>
Data source location	<p>City: Toulouse (city centre and surroundings) Country: France Latitude and longitude: 43° 36' 16.2" N; 1° 26' 38.4" E</p>
Data accessibility	<p>Repository name: SEDOOData identification number:</p> <ul style="list-style-type: none"> • Airborne 3D point cloud: doi.org/10.6096/1010 • Airborne hyperspectral data: doi.org/10.6096/1011 • Airborne thermal infrared data: doi.org/10.6096/1012 • ASD surface reflectance spectra: doi.org/10.6096/1017 • CIMEL LSE and LST measurements: doi.org/10.6096/1013 • Meteorological and atmospheric datasets: doi.org/10.6096/1014 • SOC410 reflectance measurements: doi.org/10.6096/1015 • Vehicle air temperature and humidity measurements: doi.org/10.6096/1016 <p>Direct URL to data: https://camcatt.sedoo.fr/catalogue/Instructions for accessing these data: To access the data, the user must create an account. Setting up an account is free and unrestricted.</p>

Value of the Data

- Reference datasets for the study of urban climate, providing radiometric and meteorological data from ground and airborne acquisitions collected simultaneously, are rare
- The dataset provides reference data to validate atmospheric correction and estimation of both LST and LSE in 3D urban environment in the thermal infrared domain
- The dataset provides reference acquisitions for urban land cover mapping using hyperspectral imagery, thanks to the large database of spectral optical properties
- The dataset can be used to calibrate and to support the validation of urban microclimate models, in particular for the estimation of air temperature
- The whole dataset, from in situ to airborne data, can benefit current studies in many fields of methodological research: machine learning, scaling up, multi-sensor data fusion, but also applications such as urban sprawl, Urban Heat Island (UHI), urban vegetation monitoring...

1. Objective

This new reference urban dataset has been acquired to support the development and validation of future data products from the upcoming thermal infrared (TIR) satellite missions such as TRISHNA (CNES/ISRO), LSTM (ESA) and SBG (NASA). These missions will allow for unprecedented investigations of the urban climate at the neighbourhood scale. From the state of the art, the challenge is to understand the interaction between the urban surface and the measured signal at the scale of these new missions in order to produce accurate data products such as LST, air temperature, comfort index and Urban Heat Island (UHI). This requires an accurate characterisation of the organisation of the city in terms of 3D geometry, spectral optical properties and both land surface temperature and emissivity (LST and LSE) at multiple scales. The datasets collected during the CAMCATT-AI4GEO field experiment aim to provide the necessary data to investigate this interaction. The campaign was designed to provide relevant data to calibrate and validate at-surface radiance, LST and LSE data products as well as higher level products such as air temperature or comfort index, and to support the validation of microclimate models.

2. Data Description

Fig. 1 provides an overview of the eight datasets provided in the framework of the CAMCATT-AI4GEO experiment and available in the online data catalogue, while Fig. 2 shows the extent of the area covered by these datasets. Each of the datasets shown in Fig. 1 is described individually below.

Airborne data	Ground data	Atmospheric data
Airborne 3D point cloud	ASD surface reflectance spectra	Meteorological and atmospheric datasets
Airborne hyperspectral data	CIMEL land surface emissivity and temperature measurements	
Airborne thermal infrared data	SOC410 reflectance measurements	
	Vehicle air temperature and humidity measurements	

Fig. 1. Presentation of the eight CAMCATT-AI4GEO datasets available in the data catalogue (Figure not in the repository).

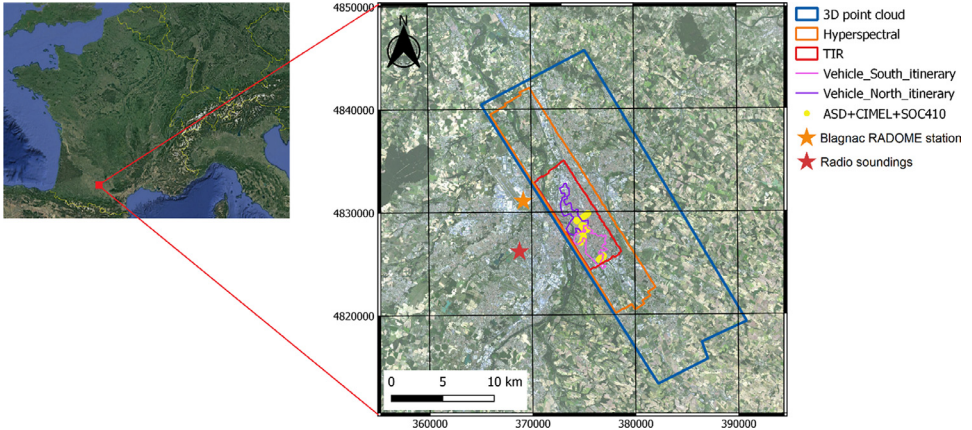


Fig. 2. Location of Toulouse in France together with a zoom on the city presenting an overview of the area covered by the different datasets acquired during the CAMCATT-AI4GEO field experiment (coordinates in UTM 31N) (Figure not in the repository).

2.1. Airborne data

2.1.1. Airborne 3D point cloud

The 3D point cloud density is approximately 100 points/m². In the provided data file, each 3D-point (x, y, z) is associated with additional data:

- Reflectance (intensity) at the laser wavelength (1064 nm, decibels (dB))
- Echo number
- Total number of echoes returned from the laser pulse
- Acquisition viewing angle (degrees)
- Point source information
- GPS acquisition time (UTC)
- Point elevation (above sea level, metres)
- RGB + NIR (800 nm) (camera Digit Number, for display only)

The “Airborne 3D point cloud” dataset shared in the repository contains several files:

- a 3D point cloud georeferenced in .LAZ format, with x, y and z coordinates and the above mentioned information associated to each 3D-point. Fig. 3 shows a portion of the point cloud displayed using true colour information
- a text file containing the data acquisition and processing metadata (including coordinate reference system information)
- a README file describing the dataset



Fig. 3. 3D point cloud displayed in RGB colours (Figure not in the repository).

2.1.2. Airborne hyperspectral data

The hyperspectral images cover the visible to the shortwave infrared domain (0.38 - 2.5 μm) with 420 spectral bands (camera characteristics in [Table 1](#)). The provided dataset consists of 25 orthorectified tiles (UTM 31N), provided in reflectance at surface level.

The “Airborne hyperspectral data” dataset shared in the repository contains several files:

- 25 orthorectified tiles (UTM 31N) provided in surface reflectance in ENVI (.bsq) format with their associated header files (.hdr). A mosaic of the 25 tiles is provided in [Fig. 4](#)
- a README file describing the dataset

2.1.3. Airborne thermal infrared data

The raw multispectral TIR images were acquired in five spectral bands (camera characteristics in [Table 2](#)). The data are provided, after pre-processing, as georeferenced mosaic (UTM 31N) in brightness temperature and radiance at sensor level together with a Land Surface Temperature (LST) map.

The “airborne thermal infrared data” dataset shared in the repository contains several files:

- a georeferenced mosaic (UTM 31N) of at sensor level brightness temperature in GeoTIFF format
- a georeferenced mosaic (UTM 31N) of at sensor level radiance in GeoTIFF format
- a Land Surface Temperature (LST) map in GeoTIFF format. This map is shown in [Fig. 5](#)
- an Excel file provided the spectral response function for each spectral band
- a README file describing the dataset



Fig. 4. Hyperspectral imagery displayed in RGB (Figure not in the repository).

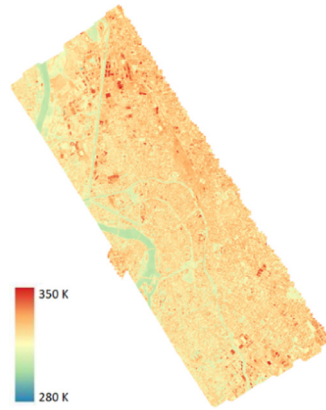


Fig. 5. LST map derived from the Telops data (Figure not in the repository).

2.2. Ground data

2.2.1. ASD surface reflectance spectra

The ASD dataset provides the spectral reflectance measurements acquired in 2151 spectral bands ranging from the visible to the shortwave infrared (0.4 - 2.5 μm). A total of 145 materials have been measured with typically three to four spectra acquired per material. Pictures and graphs are provided for each sample. An example is shown in Fig. 6 for the asphalt sample “210615_Busca_66-70”.

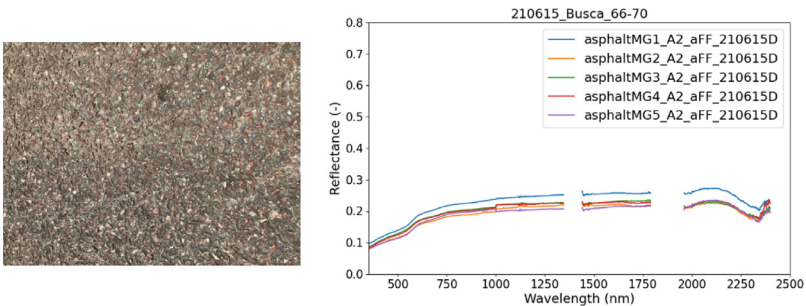


Fig. 6. Example of spectral reflectance measured with ASD and corresponding sample picture.

The “ASD surface reflectance spectra” dataset shared in the repository contains several files:

- a folder gathering 145 subfolders, one for each measured sample. The name of the subfolder is made up of the date of acquisition, district name and original ASD file number. Each subfolder contains:

- (1) xxx.txt: one or more text files derived from the original ASD measurement files acquired over the considered sample. Each text file consists of two columns: wavelengths in nanometres and measured reflectances. Several measurements are available to allow analyses of variability
 - (2) a picture (two in some cases) of the measured sample, as shown in Fig. 6
 - (3) a graph drawing the spectra provided in the text files, as shown in Fig. 6
- a text file providing an atmospheric absorption bands mask. It consists of two columns: wavelengths in nanometres and mask binary code (0 = atmospheric absorption band). This atmospheric band mask is indicative, each user can use its own mask
 - an Excel document containing three sheets:
 - (1) “all_samples”: list of all the measured samples organised according to an urban material classification. The taxonomy of urban material applied to this dataset is presented in Fig. 7. Column B: short description of the sample; Column C: name of the folder in which the data are located; Column D: list of the data files; Column E: ASD used for data acquisition; Column F: acquisition date
 - (2) “classification”: overview of the number of spectra available for each category of the urban classification
 - (3) “name_convention”: data text file naming convention
 - a README file describing the dataset

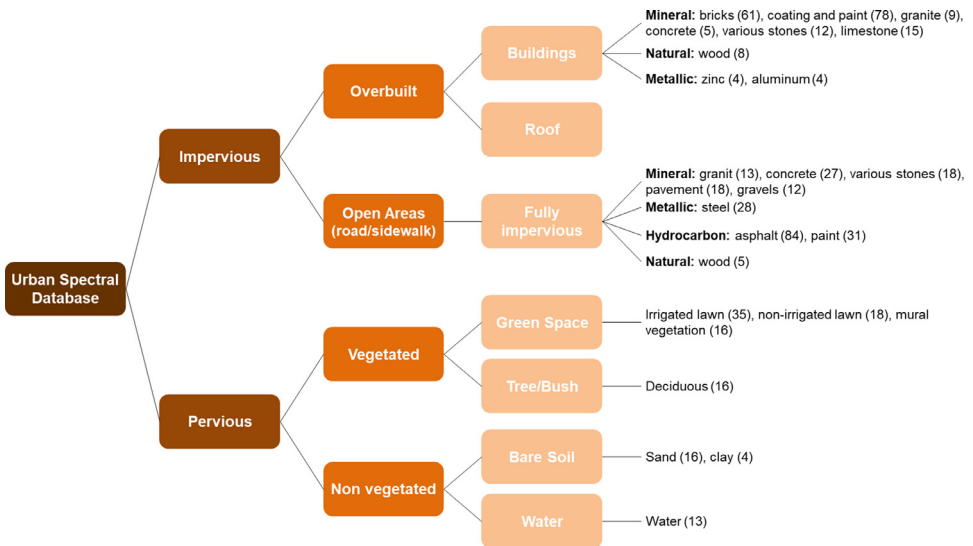


Fig. 7. Taxonomy of urban materials (the number of collected spectra by classes is indicated between brackets) (Figure not in the repository).

2.2.2. CIMEL land surface emissivity and temperature measurements

The CIMEL dataset provides LSE and LST derived from the raw TIR radiance data acquired in six spectral bands between 8 to 14 microns (8-14 μm (large band), 8.1- 8.5 μm , 8.5- 8.9 μm , 8.9- 9.3 μm , 10.3-11 μm , 11-11.7 μm) for nine types of materials: 9 asphalts, 3 concrete, 6 grasses, 1 gravel, 4 pavements, 1 sand, 1 synthetic floor, 2 tiles and 1 waters. Pictures are provided for each sample.

The “CIMEL land surface emissivity and temperature measurements” dataset shared in the repository contains several files:

- a text file for each material type containing a list of measurements acquired over that material type. In some cases, multiple measurements are available to allow variability analyses. For each sample, it provides:
 - (1) The sample name and the reference name of the corresponding picture
 - (2) A brief description including the district where the measurements were performed and which instrument was used
 - (3) LSE in the six spectral bands + LST in Kelvin. An example of LSE spectra for grass and asphalt is shown on the right hand side of Fig. 8
- a folder containing pictures of the measured samples, as shown in Fig. 8. The pictures are named using the reference name indicated in the measurement text files
- a folder containing three files with the band spectral response function for each CIMEL sensor used during the experiment. They are referenced using the sensor name indicated for each sample in the text files
- a README file describing the dataset

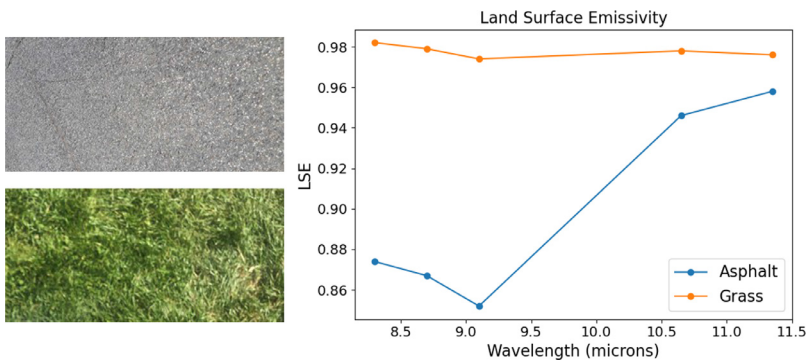


Fig. 8. Example of LSE derived from CIMEL measurement and corresponding sample picture (the graph on the right hand side is not in the repository).

2.2.3. SOC410 reflectance measurements

The SOC410 dataset provides integrated hemispherical-directional reflectance values directly measured for two viewing angles (20 and 60°) and in six spectral bands spreading from the shortwave to the thermal infrared (0.9-1.1 μm , 1.9-2.4 μm , 3.0-3.9 μm , 3.0-5.0 μm , 4.0-5.0 μm , 8.0-12.0 μm).

The “SOC410 reflectance measurements” dataset shared in the repository contains several files:

- a text file listing the measurements made over several samples. In some cases, multiple measurements are available to allow variability analyses. For each sample it provides:
 - (1) The sample name and measurement number
 - (2) A short description (optional)
 - (3) Reflectance measurements in the six spectral bands for a viewing angle of 20°. A graph showing the reflectance values for asphalt is presented on the right hand side of Fig. 9
 - (4) Reflectance measurements in the six spectral bands for a viewing angle of 60°
- a folder containing pictures of some of the measured samples. An example is given for an asphalt sample on the left hand side of Fig. 9. The pictures are referenced using point numbers, which are also mentioned in the data text file
- a README file describing the dataset

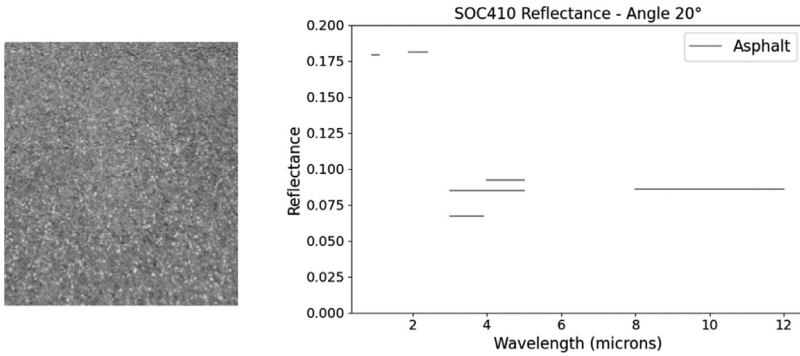


Fig. 9. Example of SOC410 reflectance measurement and corresponding sample picture (the graph on the right hand side is not provided in the repository).

2.2.4. Vehicle air temperature and humidity measurements

The vehicle dataset provides raw air temperature and humidity measurements acquired using two instrumented cars during and between airborne acquisitions following two itineraries: (1) South and (2) North. The planned itineraries are provided as KML files. The route of each vehicle was tracked by GPS. Due to some unexpected events on the road, some itineraries may vary slightly from time to time.

The “Vehicle air temperature and humidity measurements” dataset shared in the repository contains several files:

- a text file for each itinerary, consisting of height columns:
 - (1) Unixtime: UTC time of acquisition in UNIX format
 - (2) Temp_SHT85: air temperature measured by the SHT85 sensor air temperature and humidity sensor (°C). Air temperatures measured along the North and South itineraries on the 16 June 2021 during daytime are shown in [Fig. 11](#)
 - (3) Humidity: relative humidity acquired by the SHT85 sensor air temperature and humidity sensor (%), shown on the left hand side of [Fig. 10](#)
 - (4) ThermoC: air temperature acquired by the Thermocouple RS Pro type T (°C)
 - (5) Temp_In: air temperature in the sensor box (cold junction temperature of the thermocouple) (°C)
 - (6) LON: longitude of the vehicle during data acquisition (°E)
 - (7) LAT: latitude of the vehicle during data acquisition (°N)
 - (8) ALT: altitude of the vehicle during data acquisition (above sea level in meter)
- Text file naming convention:
 - YYYY-MM-DD: acquisition date
 - time: day or night for daytime and night-time acquisitions respectively
 - itinerary: SOUTH or NORTH
- North and South itineraries in vector format
- two figures showing North and South itineraries respectively, shown in the middle and right hand sides of [Fig. 10](#)
- a README file describing the dataset

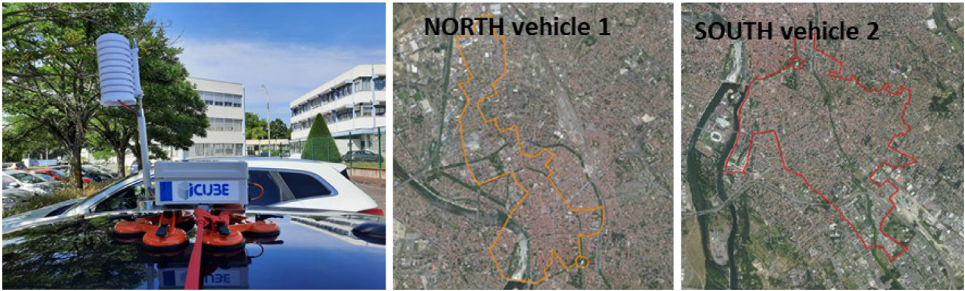


Fig. 10. Left: Picture of the air temperature and humidity equipment mounted on the roof of the vehicle (picture not in the repository). Right: North and South vehicles itineraries.

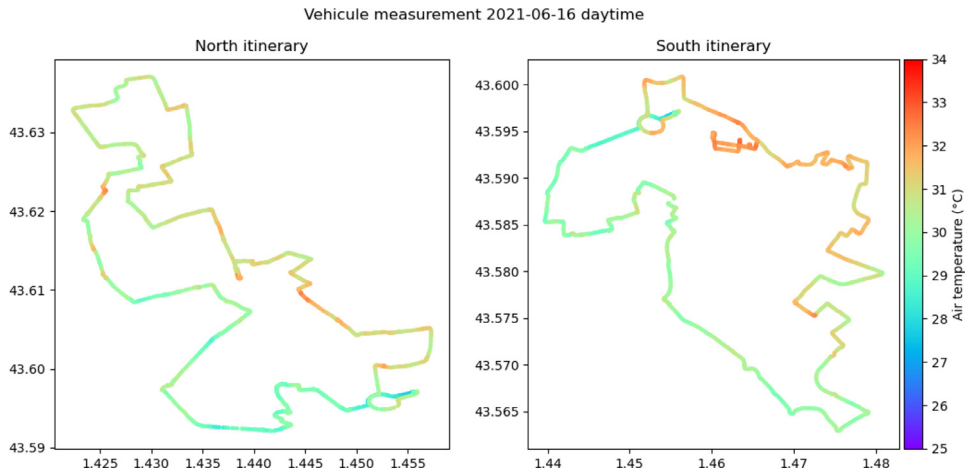


Fig. 11. Example of air temperatures ($^{\circ}\text{C}$) measured along the North and South itineraries on the 16 June 2021 during daytime (Figure not in the repository).

2.3. Atmospheric data

2.3.1. Meteorological and atmospheric datasets

The “Meteorological and atmospheric datasets” shared in the repository includes radio-soundings operated by Météo-France and the data acquired by the RADOME operational weather station at Toulouse Magonac airport according to Météo-France ground measurement standards. The dataset covers only the period of the field campaign. It consists of two folders:

(1) “Soundings”: this folder contains data acquired during four radio soundings. The data were acquired during both the ascent and descent of the sensor, with a maximum measurement height of 10 km. A Python script is provided to read the files. Depending on the date, the measurements were performed at the ONERA-Fauga Mauzac centre or at the Météo-France headquarters-Toulouse centre. Four radio soundings are available:

- ONERA-Fauga Mauzac centre on the 14 June 2021, starting acquisition hour 10h23 UTC
- Météo-France-Toulouse centre on the 15 June 2021, starting acquisition hour 10h44 UTC
- Météo-France-Toulouse centre on the 24 June 2021, starting acquisition hour 21h27 UTC
- Météo-France-Toulouse centre on the 25 June 2021, starting acquisition hour 10h47 UTC

The “Soundings” folder contains several files:

- a netcdf file for each sounding made of 14 columns:
 - (1) flight_time: UTC data acquisition time (YYYY-MM-DD HH:MM:SS)
 - (2) launch_time: UTC radiosonde launch time (YYYY-MM-DD HH:MM:SS)
 - (3) ascentRate: ascent/descent rate of measuring device (meter/second)
 - (4) pressure: atmospheric pressure (millibar)
 - (5) altitude: geopotential height (m)
 - (6) temperature: air temperature (°C), shown in red in Fig. 12
 - (7) humidity: relative humidity (%), shown in blue in Fig. 12
 - (8) dewPoint: dew point temperature (°C)
 - (9) mixingRatio: mass ratio of water defined as the mass of water divided by the mass of dry air in a given air parcel (g/kg)
 - (10) windSpeed: wind speed (meter/second)
 - (11) windDirection: wind direction as from which it originates reported as azimuthal angles (degree, 0° being the North)
 - (12) latitude: latitude of the radiosonde during data acquisition (°N)
 - (13) longitude: longitude of the radiosonde during data acquisition (°E)
 - (14) altitude_WGS84: height_above_reference_ellipsoid (m)
- File netcdf naming convention:
 - XX: measurement location ONERA-FAUGA or METEPOLE-RS
 - YYYYMMDD: acquisition date
 - HHMM: acquisition time in UTC hour
- a README file describing the dataset

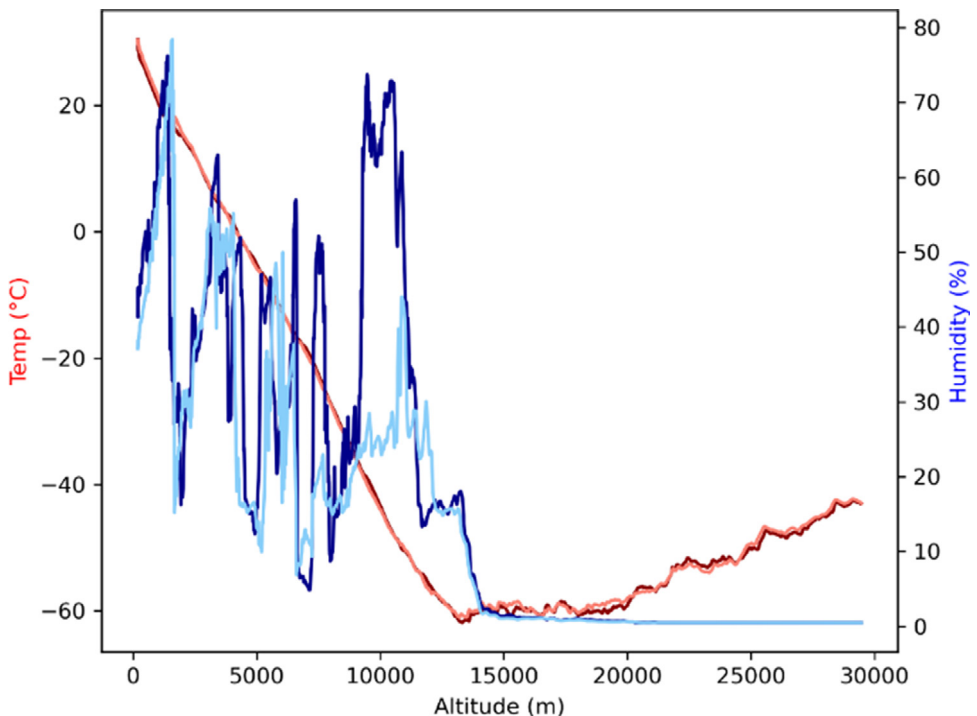


Fig. 12. Example of humidity and temperature profiles acquired by the radiosonde during the ascent (dark colour line) and the descent (light colour line) (Figure not in the repository).

(2) “station_ope_bagnac”: this folder contains meteorological measurements from Météo-France RADOME Blagnac station recorded hourly from 13 June 00am UTC to 27 June 11pm UTC (43°37′15″N, 1°22′43″E).

The “station_ope_bagnac” folder contains several files:

- 20210613-20210627_Toulouse-Blagnac_StationRADOME.data: text file composed of 16 columns separated by “;”, with the following content:
 - (1) POSTE: Weather station reference (31069001)
 - (2) DATE: UTC data acquisition time formatted as YYYYMMDDHHmn
 - (3) RR1: Hourly rainfall amount (millimetre)
 - (4) QRR1: related quality code
 - (5) T: Hourly shelter temperature (°C)
 - (6) QT: related quality code
 - (7) PSTAT: Hourly pressure (hPascal)
 - (8) QPSTAT: related quality code
 - (9) FF: Hourly wind speed (metre/second)
 - (10) QFF: related quality code
 - (11) DD: Hourly wind direction at 10m (Azimuth degree)
 - (12) QDD: related quality code
 - (13) U: Hourly relative humidity (%)
 - (14) QU: related quality code
 - (15) GLO: Hourly global radiation (Joules/cm²)
 - (16) QGL: related quality code
- a PDF file (in French) providing a description of the different parameters and related quality codes
- a figure showing some of the parameters plotted over the campaign time period
- a README file describing the dataset

2.4. Secondary data

2.4.1. Toulouse Metropole weather station network

In addition to the data collected during the field campaign, the data routinely collected by the network of weather stations set up by Toulouse Metropole are available on the following website:

<https://data.toulouse-metropole.fr/explore/?sort=modified&refine.publisher=Toulouse+M%C3%A9tropole&refine.keyword=m%C3%A9t%C3%A9o&refine.keyword=station+m%C3%A9t%C3%A9o>

Up to 60 DAVIS micro weather stations are installed in Toulouse and its surrounding area, as shown in Fig. 13. Each station can measure the air temperature, pressure, humidity, precipitation, direction and wind speed. Their locations have been chosen to account for the large diversity of land cover in the Toulouse Metropole area.

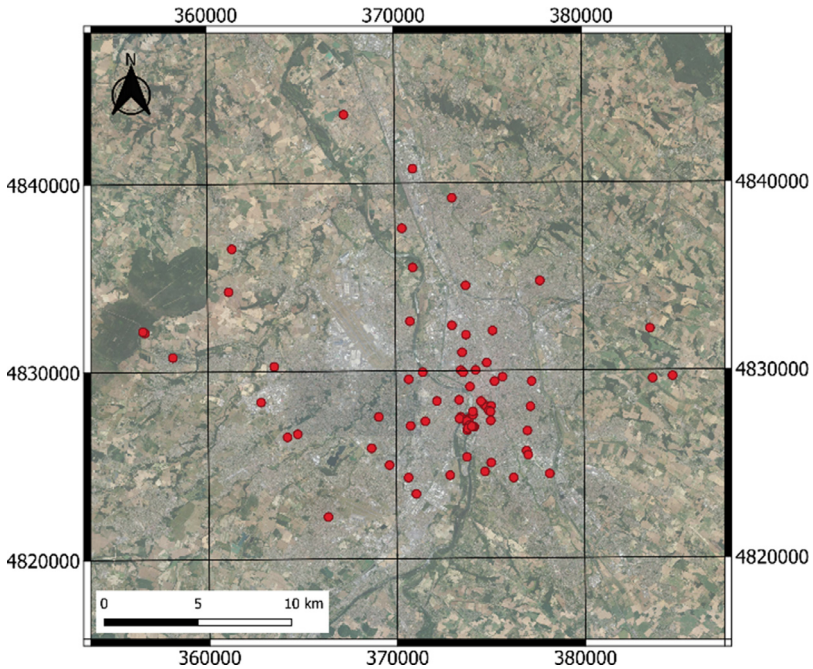


Fig. 13. Location of the weather stations (red dots) of the Toulouse Metropole network (coordinates in UTM 31N) (Figure not in the repository).

3. Experimental Design, Materials and Methods

The CAMCATT-AI4GEO field campaign took place over two weeks (14-25 June 2021 + complementary measurements on 6 September 2021), during which airborne and ground acquisitions were performed. Prior to the field campaign, the different instruments were calibrated, as described for each dataset in the following sections. The airborne acquisitions were carried out on 15 June 2021 at around 11am (UTC) over Toulouse, France, using two passive cameras and a 3D-LiDAR scanner system mounted on-board the Safire ATR42 aircraft (flight altitude ~1500 m AGL). The airborne data include the hyperspectral VISNIR-SWIR imagery, the multispectral TIR imagery and the 3D point cloud datasets, for which pre-processed data are provided. Several ground teams performed in-situ measurements during and outside the flight to collect four datasets in different spectral domains. In the reflective domain, three teams equipped with ASD spectroradiometers collected urban material spectra in different locations of the city during four days (14-16 June + 6 September). From the SWIR to TIR domain, the integrated reflectance was measured in six spectral bands and for two viewing angles with a SOC410-DHR handheld reflectometer, simultaneously as one of the ASD team. In the TIR domain, the LST and LSE were estimated from radiance measured by three CIMEL multiband radiometers set at different locations in the city. During the flight, the three CIMELs stayed at a fixed location to acquire LST over cold and hot targets for TIR airborne data calibration and validation purposes (14-15 June, first flight attempt on the 14 but unsuccessful). Outside the flight, the CIMEL were used to measure LSE over several urban materials (16 June + 6 September). Finally, air temperature and relative humidity were acquired by sensors mounted on top of two vehicles that simultaneously drove two opposite itineraries through the city, both starting and ending at the same location (43.5967438°N, 1.4546737°E). These itineraries were repeated six times, four times during day-time (14, 15, 16 and 25 June) and twice during the night (16 and 24 June).

Further details of the experimental designed and methods specific to each dataset are given below.

3.1. Airborne data

3.1.1. 3D point cloud

The 3D point cloud was acquired with a high resolution 3D-LiDAR (RIEGL VQ-1560i) collocated with two cameras (RGB and Near-Infrared). The full waveform signal was acquired by the laser scanner and multi-echo data is provided after post-processing. The final point density is approximately 100 points/m². Based on the data acquired by the two cameras, additional radiometric information (reflectance at the laser wavelength (1064 nm, decibel (dB)) + raw RGB) is associated with the coordinates and characteristics of each 3D-point.

The 3D-LiDAR calibration (geometry and radiometry) was achieved in the laboratory and controlled during a first test-flight of the CAMCATT-AI4GEO campaign using referenced ground targets deployed on the ONERA test site. The data provided are raw data with limited filtering for aerosol effects.

3.1.2. Hyperspectral data

The hyperspectral imagery was acquired with an AISA FENIX 1K camera measuring data from the visible to the shortwave infrared (0.38 - 2.5 μm) in 420 spectral bands (camera characteristics in [Table 1](#)). The camera was characterised in the laboratory prior to the flight and controlled by means of an in-flight calibration using well-characterised ground reference targets. After radiometric and geometric corrections, the data were converted in radiance units at aircraft level. The orthorectification was performed using the DEM provided by the RGE ALTI 1m database of IGN. Finally, the surface reflectance was retrieved after applying an atmospheric correction performed with COCHISE [1] using the atmospheric profile and aerosols in situ measurements from a CIMEL spectroradiometer acquired simultaneously to the flight. Due to computational issues, the full dataset was divided into 25 tiles, which were processed individually.

Table 1

AISA FENIX 1K camera characteristics.

Swath:	1024 m
Ground Sample Distance:	1 m
Spectral resolution	* ~3.6 nm between 380 - 970 nm * ~7.8 nm between 970 - 2500 nm

3.1.3. Thermal infrared data

The multispectral TIR images were acquired with a Telops camera equipped with a filter wheel in five spectral bands (camera characteristics in [Table 2](#)). The camera was first characterised in the laboratory to allow for radiometric and geometric correction of the data. The spectral response function for each band was then measured for the entire optical system, including sensors and lenses. The data were also controlled during the flight by comparison with ground measurements. Due to variations of the filter wheel temperature during data acquisition, a post processing correction was applied based on the wheel temperature recorded during the flight and wheel characterisation data acquired in a controlled environment (in the laboratory). The orthorectification was achieved using the DEM (Digital Elevation Model) provided by the RGE ALTI 1m database of IGN and the data were mosaicked using correlator 3D. Based on calibration data, the data were then converted into brightness temperature and radiance at sensor level. Finally, the LST retrieval was carried out by first estimating the Bottom Of Atmosphere (BOA) radiance using COMANCHE [2] and then applying the TES method (MMD, Minimum-Maximum emissivity Difference, relation trained on the ECOSTRESS spectral library [3]).

Table 2

Telops camera characteristics.

Swath:	593 × 472 m ²
Ground Sample Distance:	1.95 m
Spectral bands (band width and equivalent wavelength):	<ul style="list-style-type: none"> * Band 1: 8.1-8.9 (8.55) microns * Band 2: 8.7-9.9 (9.10) microns * Band 3: 10.5 - 12.9 (11.74) microns * Band 4: 6 - 14 (9.72) microns (broadband) * Band 5: 10 - 11.5 (10.73) microns

3.2. Ground data

3.2.1. ASD surface reflectance spectra

The spectral reflectance was measured in 2151 spectral bands ranging from the visible to the shortwave infrared (0.4- 2.5 μm) using three ASD spectroradiometers. The measurements were acquired in reflectance mode, based on regular measurements of a white spectralon plate. The ASD instrument used for each measurement is indicated in the data file name. In the case of the ASD named "CESBIO2", a first instrument was used on 14 June and a second one on 15 and 16 June but the same name was kept for a sake of simplicity.

In general, three to four spectra were acquired per material to allow variability analyses. Raw data without any filtering are provided to allow each user to apply its own correction methods. A preliminary check was performed to remove erroneous data but spectra with a significant leap at 1 micron or noise in the SWIR domain were kept as some part of the spectrum may be of interest for some applications. It is then strongly recommended to check the data carefully using the graph associated with each sample.

3.2.2. CIMEL land surface emissivity and temperature measurements

Thermal infrared (TIR) radiance was acquired in six spectral bands between 8 to 14 microns (8-14 μm (large band), 8.1- 8.5 μm , 8.5- 8.9 μm , 8.9-9.3 μm , 10.3-11 μm , 11-11.7 μm) using three CIMEL radiometers. A TES algorithm was used to retrieve the LSE in each band and the LST from the measured TIR radiance of the sample and the measurement of downward atmospheric radiance via an Infragold target. The three CIMELs were intercalibrated at the ONERA laboratory over 14 samples (RMSE~1%).

Several acquisitions were performed over each material to allow variability analyses.

3.2.3. SOC410 reflectance measurements

Integrated reflectance was measured for two viewing angles (20 and 60°) and in six spectral bands spreading from the shortwave to the thermal infrared (0.9-1.1 μm , 1.9-2.4 μm , 3.0-3.9 μm , 3.0-5.0 μm , 4.0-5.0 μm , 8.0-12.0 μm) using a SOC410-DHR handheld reflectometer. The acquired data are hemispherical-directional reflectance.

More details on the instrument can be found here: <https://surfaceoptics.com/wp-content/uploads/2020/11/410-DHR-Datasheet.pdf>. In total, 63 materials were measured. Pictures are available for some of the samples.

3.2.4. Vehicle air temperature and humidity measurements

Measurements of air temperature and humidity were acquired during and between airborne acquisitions using two instrumented vehicles following two different routes: the South and the North itineraries. The route of each vehicle was tracked by GPS. Due to some unexpected events on the road, some routes may vary slightly from time to time.

Each vehicle was equipped with a SHT85 sensirion air temperature and humidity sensor and a temperature sensor Thermocouple RS Pro type T, which acquired data every two seconds. The sensors were mounted approximately 2m above the ground.

The acquired data were compared with measurements taken by several weather stations from the Toulouse Metropole network at the time of vehicle's passage. The stations used for

these analyses are: Valade (00), Compans (42), St Cyprien (27) and Carmes (28). The comparison showed that the measurements were consistent, with an RMSE of 0.9 K and 6.4 % for air temperature and humidity respectively on the North itinerary.

3.3. Atmospheric data

Radio soundings were carried out simultaneously with the flight by Météo-France using a radiosonde VAISALA-RS41. The data were acquired during both the sensor ascent and descent with a maximum measurement height of 10 km. Depending on the date, the measurements were performed at the ONERA-Fauga Mauzac centre or at the Météo-France headquarters-Toulouse centre.

More information concerning the Météo-France RADOME Blagnac station are available here: https://donneespubliques.meteofrance.fr/?fond=contenu&id_contenu=37.

3.4. Secondary data

The data routinely collected by the network of weather stations set up by Toulouse Metropole are described in [4].

Table 3

Field experiment timeline (RS indicates the date of the radio sounding, and IC the date of the CIMEL and SOC410 inter-calibration exercise).

Field campaign planning	14/06	15/06	16/06	17/06	18/06	19/06	20/06	21/06	22/06	23/06	24/06	25/06	06/09
Airborne data		~11am											
Ground data	All	All	All					IC			Vehicles	Vehicles	ASD + CIMEL
Atmospheric data	RS	RS									RS	RS	

Ethics Statement

The authors declare that this work did not involve human subjects nor animal experiment.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

Data Availability

CAMCATT-AI4GEO (Original data) (SEDOO).

CRediT Author Statement

L. Roupioz: Conceptualization, Methodology, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Visualization, Supervision, Project administration; **X. Briottet:** Conceptualization, Methodology, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Supervision, Project administration, Funding acquisition; **K. Adeline:** Investigation, Resources, Writing – review & editing; **A. Al Bitar:** Conceptualization, Methodology, Investigation, Resources; **D. Barbon-Dubosc:** Investigation; **R. Barda-Chatain:** Formal analysis, Investigation, Visualization; **P. Barillot:** Methodology, Formal analysis, Investigation, Resources; **S. Bridier:** Investigation; **E. Carroll:** Investigation, Resources; **C. Cassante:** Investigation; **A. Cerbelaud:** Investigation; **P. Déliot:** Methodology, Formal analysis, Investigation, Resources; **P. Doublet:** Resources; **P.E. Dupouy:** Investigation, Resources; **S. Gadal:** Conceptualization, Methodology, Investigation; **S. Guernouti:** Conceptualization, Methodology, Investigation, Resources; **A. De Guilhem De Lataillade:** Methodology, Investigation, Resources; **A. Lemonsu:** Methodology, Investigation, Resources; **R. Llorens:** Formal analysis, Investigation; **R. Luhahe:** Methodology, Formal analysis, Investigation, Resources; **A. Michel:** Investigation, Writing – review & editing; **A. Moussous:** Investigation, Resources; **M. Musy:** Conceptualization, Methodology, Resources; **F. Nerry:** Conceptualization, Formal analysis, Investigation, Resources, Writing – review & editing; **L. Poutier:** Methodology, Formal analysis, Investigation, Resources, Writing – review & editing; **A. Rodler:** Conceptualization, Formal analysis, Investigation, Resources, Writing – review & editing; **N. Riviere:** Methodology, Formal analysis, Investigation, Resources, Writing – review & editing; **T. Riviere:** Investigation; **J.L. Roujean:** Conceptualization, Methodology, Investigation, Resources; **A. Roy:** Methodology, Formal analysis, Investigation, Resources, Writing – review & editing; **A. Schilling:** Formal analysis, Writing – review & editing; **D. Skokovic:** Formal analysis, Investigation, Resources; **J. Sobrino:** Conceptualization, Methodology, Investigation, Resources.

Acknowledgments

This work was supported by BPI (AI4GEO) and CNES (APR 2021 CAMCATT).

References

- [1] C. Miesch, L. Poutier, V. Achard, X. Briottet, X. Lenot, Y. Boucher, Direct and inverse radiative transfer solutions for visible and near-infrared hyperspectral imagery, *IEEE Trans. Geosci. Remote Sens.* 43 (7) (2005) 1552–1561, doi:10.1109/TGRS.2005.847793.
- [2] L. Poutier, C. Miesch, X. Lenot, V. Achard, Y. Boucher, COMANCHE and COCHISE: two reciprocal atmospheric codes for hyperspectral remote sensing, in: 2002 AVIRIS Earth Science and Applications Workshop Proceedings, 2002.
- [3] S.K. Meerdink, S.J. Hook, D.A. Roberts, E.A. Abbott, The ECOSTRESS spectral library version 1.0, *Remote Sens. Environ.* (2019) 230, doi:10.1016/j.rse.2019.05.015; A.M. Baldridge, S.J. Hook, C.I. Grove, G. Rivera, The ASTER spectral library version 2.0, *Remote Sens. Environ.* 113 (2009) 711–715.
- [4] G. Dumas, V. Masson, J. Hidalgo, V. Edouart, A. Hanna, G. Poujol, Co-construction of climate services based on a weather stations network: application in Toulouse agglomeration local authority, *Clim. Serv.* (2021) 24, doi:10.1016/j.cliser.2021.100274.