Journal: scientific data (nature)  
Article type: Data Descriptor (Descriptions of datasets shared in repositories)

Format

* Title
* Author list (does not need a heading)
* Abstract
* Background & Summary
* Methods
* Data Records
* Technical Validation
* Usage Notes (optional)
* Code Availability
* References
* Author Contributions
* Competing Interests
* Acknowledgements

* Title (recommended length 110 characters or fewer)

**A spatio-temporal Culicoides species dataset produced by the French surveillance program from 2009 to 2012**

Hammami P, Baltusyte I, Taconet P, Delécolle JC, Setier-Rio ML, Mathieu B, Venail R, Balenghien T\*, Garros C\*

*En plus sur le jeu de données GBIF : Delécolle D, Akaddar A, Allène X, Chavernac D, Baldet T*, *Rakotoarivony I, Lhoir J, Scheid B*

* Abstract (maximum 170 words)

*Culicoides* biting midges (Diptera: Ceratopogonidae) are vectors of pathogens of veterinary and public health importance, including bluetongue, Epizootic Hemorrhagic Disease, and Oropouche viruses. Following the incursion and spread of multiple bluetongue virus serotypes in France (2000–2008), a national entomological monitoring program (2009–2012) was launched to support livestock policies and meet European Commission requirements. This program generated the most comprehensive dataset on Culicoides midges in France, with over 6.34 million specimens from at least 83 species. Midges were identified to species, sexed, and categorized by gonotrophic and physiological status. Sampling was conducted in 210 farms spread across all the mainland French departments, and synchronized nationally with a weekly frequency during population fluctuations and monthly during periods of activity/inactivity. The dataset provides unprecedented spatiotemporal coverage, with over 66% of identified specimens belonging to the C. obsoletus/C. scoticus complex. This comprehensive resource enhances understanding of midge ecology, diversity, and phenology, supporting vector-borne disease modeling, surveillance, and control strategies at a national scale.

* Background & Summary (unlimited length)

*Culicoides* biting midges (Diptera: Ceratopogonidae) are small hematophagous insects recognized as vectors of numerous pathogens, including bluetongue virus (BTV), Schmallenberg virus, epizootic hemorrhagic disease virus (EHDV), and Oropouche virus1. Several of these pathogens pose significant threats to livestock health and agricultural economies2–6. In the 2000s, the epidemiological importance of *Culicoides* biting midges has driven extensive surveillance efforts to monitor their distribution, seasonal dynamics, and potential role in disease transmission.

In October 2000, the epidemiological landscape of France experienced its first incursion of bluetongue virus serotype 2 (BTV-2) in Corsica (39 outbreaks in 2000, followed by 335 outbreaks in 2001)7–9. This marked the beginning of a critical period for vector-borne disease surveillance in the country. In 2002, in response to this emergence, the French Ministry of Agriculture and Fisheries mandated the CIRAD (Centre de coopération internationale en recherche agronomique pour le développement) to coordinate an entomological monitoring program for *Culicoides* biting midges, the primary vectors of BTV10. The focus was on monitoring *Culicoides imicola*, a known BTV vector in Corsica, and on establishing sentinel traps in the Mediterranean coastal region to detect its potential establishment on the mainland.

The surveillance network intensified following the emergence of new BTV serotypes, leading to significant outbreaks that severely impacted French livestock. In Corsica, BTV-4 was detected in 2003 (17 outbreaks), followed by BTV-16 in 2004 (25 outbreaks). On the mainland, multiple incursions of BTV-8 and BTV-1 occurred between 2006 and 2009, resulting in over 50,000 outbreaks8,11,12. Entomological monitoring was extended to affected regions included the Pyrénées-Atlantiques (2005), northeastern departments (2006), central France (2007), and Brittany (2008). In parallel, the disease largely extended its range in northern areas of Europe8. Consequently, the European Commission regulation 2007/1266/EC12 mandated comprehensive surveillance of bluetongue disease to demonstrate the absence of specific serotypes, early detect any emergence, and monitor its spread, along with mandatory monitoring of its vectors to determine the seasonally vector-free period13.

In line with these comprehensive requirements, from 2009 to 2012, French populations of vectors species of *Culicoides* genus were monitored using 160 traps deployed across mainland France, with one or two traps per department. Trapping was carried out weekly in spring and autumn, and monthly for the rest of the year. The objectives of this network were multifaceted: to inventory *Culicoides* species, monitor population dynamics, and determine the start and end of vector activity periods (phenology). These temporal benchmarks had direct regulatory implications, as the cessation of vector activity permitted a relaxation of livestock movement restrictions imposed under European Union legislation. This national program also allowed for the collection of data across diverse ecoclimatic zones, highlighting significant variations in *Culicoides* diversity and seasonal abundance. Finally, this network contributed to raise awareness among farmers and veterinarians.

Comprehensive monitoring efforts, vaccination campaigns and trade restrictions gradually halted the spread of the disease, and transmission of BTV gradually declined, with only 83 outbreaks reported in 2009 and just one in 2010. By 2012, BTV transmission was successfully controlled, allowing France to regain its bluetongue free status. This national surveillance program distributed across mainland France and Corsica, was discontinued. However, the extensive data collected during this period has enabled *Culicoides* biting midges ecology to be studied at an unprecedented spatial and temporal resolution14,15.

The dataset described in this paper originates from this 2009-2012 entomological monitoring network and consists of three interconnected datasets documenting sampling events, Culicoides biting midges occurrences and abundance per species and per physiological state, and associated environmental conditions. The establishment of a synchronized and nationwide surveillance system provided an unprecedented opportunity to study Culicoides midges distribution and phenology across France. This dataset, unparalleled in its spatial and temporal resolution, enables the correlation of species distribution and dynamics with climatic and ecological factors, thereby improving our understanding of vector ecology and informing predictive models for future vector-borne disease outbreaks. It allows researchers to analyze species distribution, population dynamics, and ecological drivers of Culicoides midges activity 16–18. It has already contributed to vector ecology and distribution studies19 and can further support descriptive and predictive modeling of vector-borne disease risk under climate change scenarios. By integrating data on species composition, phenology, and environmental parameters, this dataset serves as a valuable resource for entomologists, epidemiologists, and policymakers involved in vector surveillance and disease prevention strategies.

* Methods (unlimited length)

Between 2009 and 2012, Culicoides midges population monitoring in mainland France and Corsica was conducted by national authorities and research groups in compliance with European requirements. The program deployed 160 traps with one or two traps allocated per department. The monitoring program targeted farms housing cattle, sheep, and horses. It should be noted that some sampling sites left the monitoring network during the program, and were replaced by a similar farm nearby in the same *department* (French administrative division).

**Sampling frequency and duration**

The sampling schedule varied by season. From mid-February to April and November to mid-December, traps were performed one night per week, whereas a monthly schedule was adopted for the remainder of the year (In January, and between May to October). In total, 14,895 collections were performed under the supervision of the Directions départementales de la cohésion sociale et de la protection des populations (DDecPP).

**Sample collection and preservation**

The traps used for the monitoring were black-light suction traps, also known as Onderstepoort traps (Onderstepoort Veterinary Institute, Pretoria, South Africa). These traps are known to capture higher numbers of Culicoides midges compared to other trap types 20,21.

The traps were positioned outside stables or near animal resting areas and operated from dusk to dawn. In these traps, insects are attracted by the UV light of the trap and collected using a fan, which directs them into a beaker containing soapy water. The soap allows the insects to sink and prevents them from drying out. A thin net was placed around the trap to prevent the collection of large insects. Once collected, the specimens were transferred to 70% ethanol at room temperature for storage and transport to identification centres.

**Specimen identification and processing**

The samples were processed at three specialized sorting centers: Cirad (Centre de coopération internationale en recherche agronomique pour le développement) and EID Med (Entente interdépartementale pour la démoustication du littoral méditerranéen) in Montpellier, and IPPTS (Institut de parasitologie et de pathologie tropicale) in Strasbourg. Expert entomologists performed species identification using morphological keys and the IIKC database 22,23. Morphological identification relied on key features such as sensory pit shape and size, wing spot patterns, and the number and morphology of spermathecae in females or the aedeagus and parameres in males. Damaged or incomplete specimens were excluded. Non-Culicoides specimens were not identified.

Identification was conducted to the species level or, where differentiation was not possible, to the complex level. Indeed, due to morphological indistinguishability, certain species were grouped together at a complex level: *i)* *C. cataneii* (Clastrier, 1957) and *C. gejgelensis* (Dzhafarov, 1964), *ii)* *C. obsoletus* (Meigen, 1818) and *C. scoticus* (Downes and Kettle, 1952), and *iii) C. sejfadinei* (Dzhafarov, 1958) and *C. tauricus* (Gutsevich, 1959).

In cases of excessively large samples, subsampling was performed to streamline the identification process, following the protocol outlined in Van Ark and Meiswinkel (1992) 18,24 . Specifically, if the total insect volume exceeded 3 mL, a subsample was taken, offering significant time savings. In some cases (in particular when the trap screen was damaged or misplaced allowing large insects to enter), before the subsampling, a preliminary sorting step was conducted under a stereomicroscope to isolate Culicoides midges from other insect species based on morphological traits, including body shape, antennal segment count, and wing patterns.

**Measurements and facts**

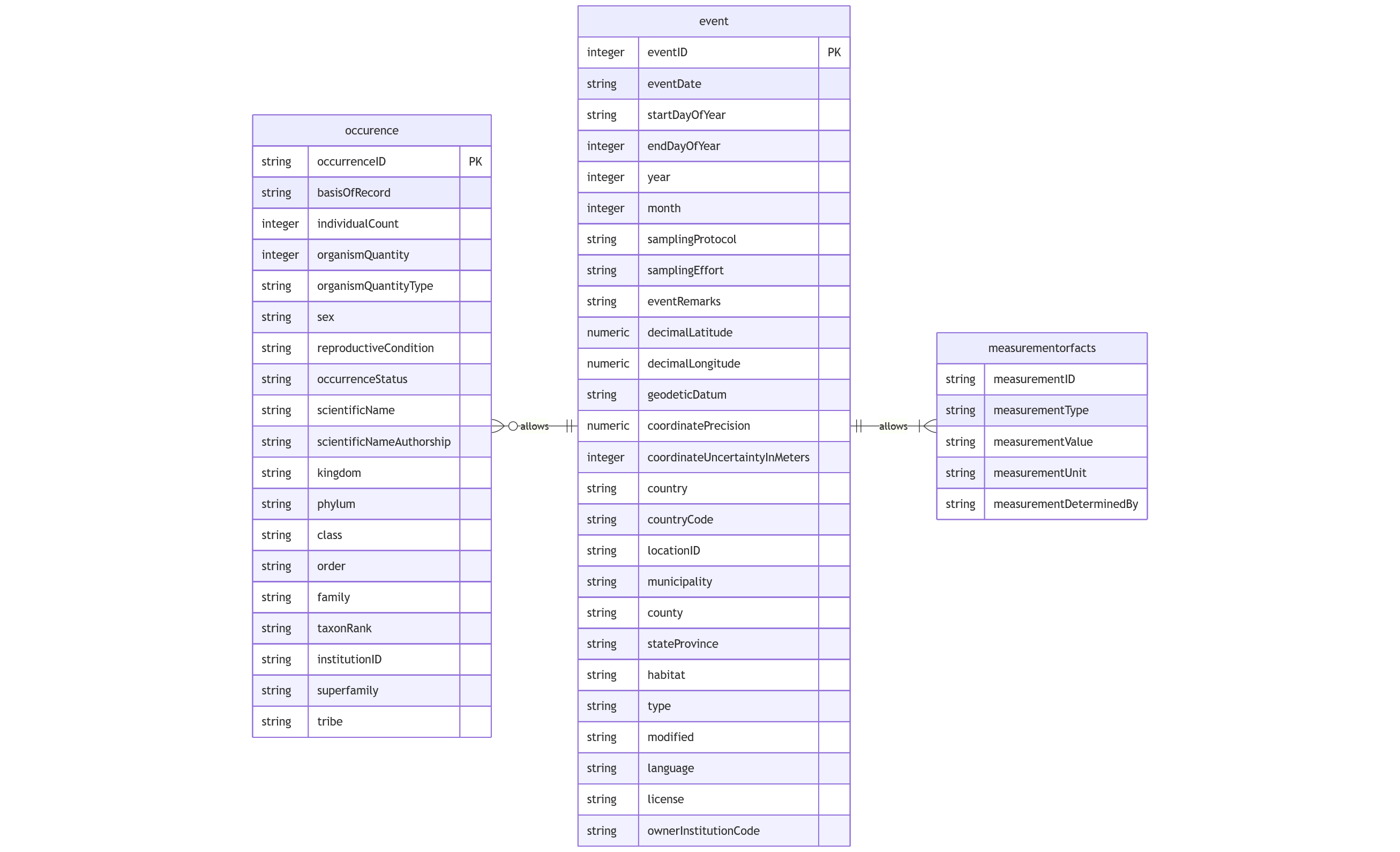
For each sampling campaign, additional information was recorded, either by on-site measurements or by further processing. On-site measurements included qualitative variables: building opening rate, trap position (inside or outside the stable), wind/cloud/rain at trap installation and collection, as well as quantitative ones: date and time of trap installation and collection, and GPS coordinates. The operator also had the opportunity to leave textual/non-formalized comments for each event. *A posteriori* measurement aimed at describing the environment of the trap and most of them were retrieved using GIS approaches intersecting exact GPS coordinates with spatial dataset collected from various sources such as ERA-5 land25 or MODIS26. Those data include elevation, livestock densities in the canton, land cover classes, biogeographical region, vegetation indexes, daily wind speed, daily temperatures (air and soil), daylight duration, surface net short-wave radiation flux, daily precipitation sum, and volume of water in soil layer 1 (0 - 7 cm). The data were restructured using R software to meet GBIF publication criteria.

* Data Records (unlimited length)

The database is openly accessible through the Global Biodiversity Information Facility (GBIF)27, where it can be downloaded as a Darwin Core Archive (DwC-A)29. The dataset was standardized to the Darwin Core structure as sampling-event data, and fit to the FAIR principles28, ensuring its findability, accessibility, interoperability, and reusability. It is presented as a set of files in tab-delimited txt format.

This sampling-event dataset comprises three interrelated tables based on the DarwinCore (DwC) standard30: the event core (**'event.txt')**, the occurrence extension (**'occurence.txt'),** and the measurement or facts extension (**'measurementorfacts.txt')**, respectively designed to capture data related to trapping events, *Culicoides* midges identifications, and complementary measurements or facts for each events, such as environmental data. The sampling event unique identifier (‘eventID') being the primary key defining the relational structure of the database and facilitating analyses of insect populations in relation to sampling conditions. The database structure is illustrated in Fig. X.

Figure X: PK being the primary key

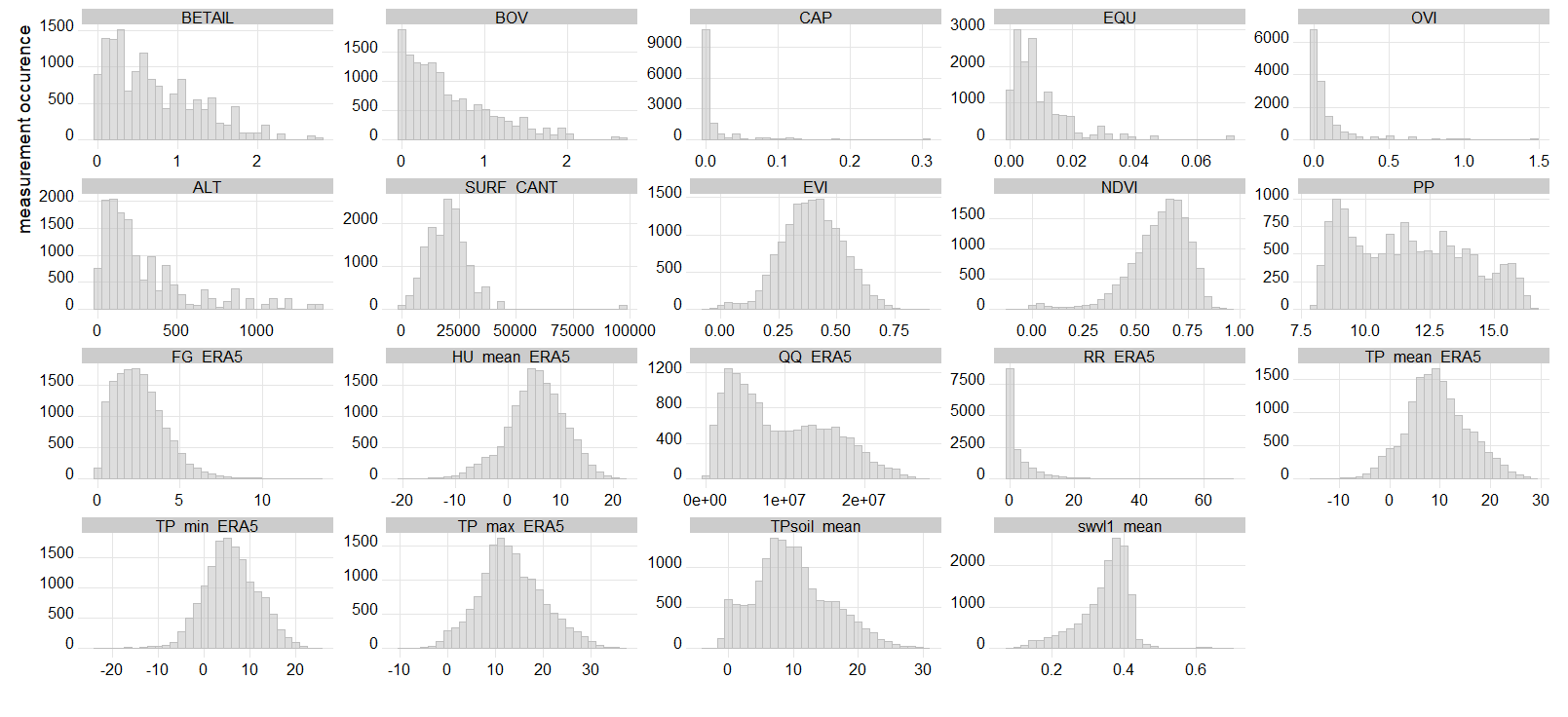


The field name in the database were also based on DwC: ‘eventID', 'eventType', 'eventDate', 'startDayOfYear', 'endDayOfYear', 'year', 'month', 'samplingProtocol', 'samplingEffort', 'eventRemarks', 'decimalLatitude', 'decimalLongitude', 'geodeticDatum', 'coordinatePrecision', 'coordinateUncertaintyInMeters', 'country', 'countryCode', 'locationID', 'municipality', 'county', 'stateProvince', 'habitat', 'type', 'modified', 'language', 'license', 'ownerInstitutionCode', 'week', 'occurrenceID', 'basisOfRecord', 'individualCount', 'organismQuantity', 'organismQuantityType', 'sex', 'reproductiveCondition', 'occurrenceStatus', 'scientificName', 'kingdom', 'phylum', 'class', 'order', 'family', 'taxonRank', 'institutionID', 'superfamily', 'tribe', 'scientificNameAuthorship', 'measurementID', 'measurementType', 'measurementValue', 'measurementUnit', 'measurementDeterminedBy’; for more details, see the Darwin Core Quick Reference Guide31.

The event core contains 14,895 sampling events carried out between 2009 and 2012 in 210 different farms, the occurrence extension has 8,683,785 occurrences belonging to more than 6,340,000 individuals, and the measurement or facts extension contains 491,535 event-associated measurements from 33 parameters described in Table Y. The occurrence records are arthropod insects belonging to Diptera order and the Ceratopogonidae family. It includes more than 80 species.

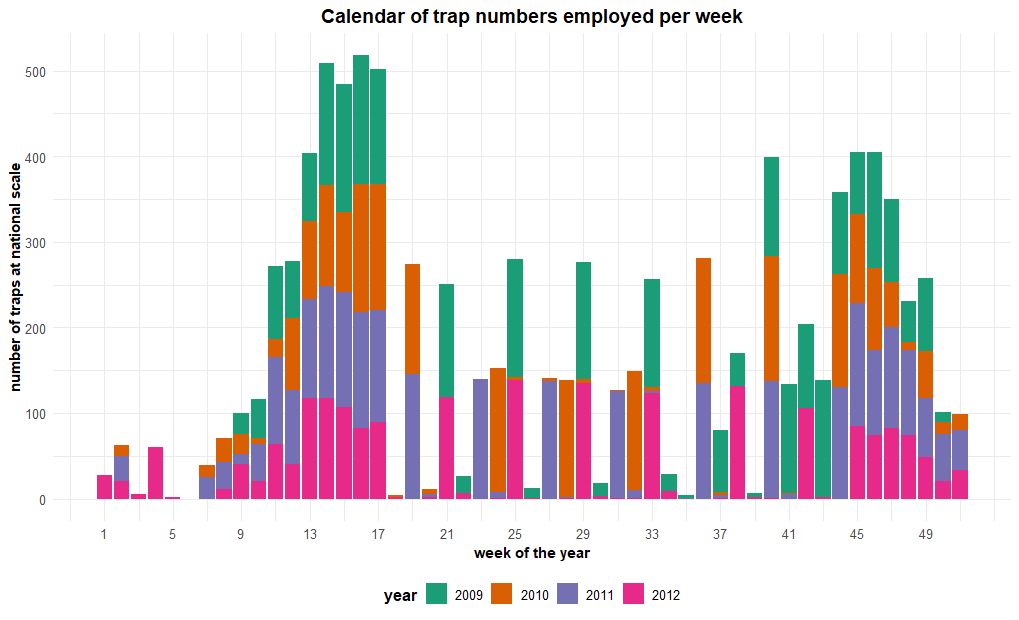
Table Y:

|  |  |  |  |
| --- | --- | --- | --- |
| measurementID | Variable description | Range of values | Source |
| ALT | Global 30 Arc-Second Elevation Data in meters (GTOPO30) | [1 ; 1395] | E-OBS 32,33 |
| BETAIL | Density of horse, cattle, sheep and goat in the canton (ind/ha) | [1 ; 2.76] | 2010 agricultural census 34 |
| BOV | Density of cattle in the canton (ind/ha) | [0 ; 2.67] |
| CAP | Density of goat in the canton (ind/ha) | [0 ; 0.31] |
| OVI | Density of sheep in the canton (ind/ha) | [0 ; 1.47] |
| EQU | Density of horse in the canton (ind/ha) | [0 ; 0.07] |
| SURF\_CANT | Surface of the canton (ha) | [725 ; 97440] |
| CLC | Thematic classes of land cover and land use | 15 categories | CORINE Land Cover 2018 (vector/raster 100 m), Europe, 6-yearly35 |
| ECO\_CLI | Biogeographical regions | 4 categories: Continental, Atlantic, Mediterranean, and Alpine | European Envrionment Agency36 |
| EVI | The enhanced vegetation index | [-0.10 ; 0.96]  -3000 values are to be considered as missing values | MOD13Q1 v061 from MODIS26,37 |
| NDVI | Normalized Difference Vegetation Index | [-0.07 ; 0.88]  -3000 values are to be considered as missing values |
| PP | Photoperiod / daylength (hours) | [7.94 ; 16.52] | Meteor package38,39 |
| FG\_ERA5 | Daily mean wind speed (m s-1) calculated from 10m u-component of wind (*windu*) and 10m v-component of wind (*windv*) | [0.01 ; 13.44] | ERA5-Land hourly data from 1950 to present25 |
| HU\_mean\_ERA5 | Daily averaged 2m dewpoint temperature. A measure of the humidity of the air (°C) | [-20.39 ; 21.56] |
| QQ\_ERA5 | Daily average surface net solar radiation (J m-2) | [197473 ; 28013636] |
| RR\_ERA5 | Daily precipitation sum (mm) | [0 ; 68] |
| TP\_mean\_ERA5 TP\_min\_ERA5 TP\_max\_ERA5 | Daily mean, min and max 2m temperatures (°C) | [-15 ; 29]  [-23 ; 25]  [-10 ; 37] |
| swvl1\_mean | Daily average volumetric water in the soil in layer 1 (0 - 7 cm) of the ECMWF Integrated Forecasting System (m3 m-3) | [0.08 ; 0.69] |
| TPsoil\_mean | Daily average temperature of the soil in layer 1 (0 - 7 cm) of the ECMWF Integrated Forecasting System (°C) | [-3.81 ; 30.24] |
| BUILD\_OP | Building opening rate  (%) | 4 categories: 0-5, 5-25, 25-75, and 75-100 | Communication for operators |
| SITUATIONPIEGE | Trap position | 2 categories : inside, and outside |
| VENTDEBUT / VENTFIN | Wind at trap installation / collection | 4 categories: no wind, slight breeze to light wind, medium wind, strong wind |
| NUAGEDEBUT / NUAGEFIN | Cloud at trap installation / collection | 4 categories: no cloud cover, low cloud cover, medium cloud cover,  high cloud cover |
| PLUIEDEBUT / PLUIEFIN | Rain at trap installation / collection | 4 categories: no rain, fog / drizzle, light rain, heavy rain |
| SOUSECH | Subsampling protocol | 2 categories : no, yes | Communication from the entomological expert in charge of identification |
| VTOT | In case of subsampling, volume total before subsampling (mL) | [3.4 ; 322] 0 must be considered as NA |
| VTRAIT | In case of subsampling, traited volume during subsampling (mL) | [1 ; 6]  0 must be considered as NA |
| PRETRI | Pre-sorting of *Culicoides* genus individuals before sub-sampling | 2 categories: no, yes |



The geographical coverage of the database is illustrated on the map displayed in Figure X. All the sampling sites extend from 41.5°N to 50.8°N in latitude, from 4.51°O to 9.5°E in longitude.

The annual collection calendar is presented in fig. X, with weekly trapping nights during the spring and autumn months, and a monthly schedule for the remainder of the year.

Figure X. Trap collection schedule during for the 2009 - 2012 study period.

The dataset contains data from 209 sampled livestock farms with a total of 14895 trap collections and a total of 19,209,261 specimens of *Culicoides* caught.

The majority of the trapping sites were cattle farms (65.5%, n = 137), followed by mixed livestock farms (20.6%, n = 43). Over the study period from 2009 to 2012, a total of 166, 164, 164, and 158 traps were deployed each year, respectively. The specific light trap locations, associated metropolitan France departments and their distribution across biogeographic regions are presented in Figure X1.

A map of france with different colored areas

AI-generated content may be incorrect.

Figure X1. Geographic distribution of *Culicoides* trapping sites across mainland France from 2009 to 2012. Trap locations are overlaid on biogeographic regions, with French department boundaries outlined for reference.

In total, species contributing the most to total catch volume belong to the *C. obsoletus* / *C. scoticus* complex in all regions except for Corsica, where *C. imicola* accounted for most individuals captured. *Culicoides* population composition varied noticeably between regions with slight shifts in less dominating species' relative abundance (Fig. X2).

Notably, *C. dewulfi* showed greater relative abundance in the north-western regions of France compared to other areas, while *C. newsteadi* was more prominent near the Mediterranean basin. Overall, despite the interannual variability, the population structure of the seven most abundant species remained stable in most regions, representing consistent habitat preference, with the *C. obsoletus* group species remaining dominant across the south-to-north gradient in the majority of mainland France areas.

A map of france with different colored circles

AI-generated content may be incorrect.

Figure X2. Spatial *Culicoides* population structure with seven most abundant *Culicoides* species across different administrative regions in France.

The Inter-annual trends of different *Culicoides* species are seen in the strong fluctuations observed across consecutive years (Figure X3). The notable peaks during the summer months highlight inconsistent presence and strong temporal dynamics. The *C. obsoletus* and *C. scoticus* group, as well as *C. punctatus* and *C. chiopterus* presented unimodal spikes in abundance during the 2009-2010 season, whereas multimodal patterns could be observed for other years. *C. newsteadi* population fluctuated moderately but consistently, suggesting more stable inter-annual trends. In contrast, *C. imicola* displayed high peaks and significant drops in population abundance, indicating more abrupt temporal dynamics.

A group of blue lines

AI-generated content may be incorrect.Figure X3. Intra-annual variability of seven most abundant *Culicoides* species found in France for the 2009 - 2012 period.

The finer time-scale seasonal dynamics of the most abundant *Culicoides* species are visualized in Figure X4. Each species showcased distinct phenology, with varying activity period length throughout the year. Species within the *C. obsoletus* and *C. scoticus* complex activity period was mostly confined to the warmer months, whereas *C. imicola* and *C. dewulfi* indicated delayed seasonal onset in relation to other species, with activity extending further into the autumn. In contrast, *C. newsteadi* populations appeared relatively consistent year-round. Overall, the large spread of collected *Culicoides* spp. values reflect the variability observed after each trapping session.

A group of blue bars

AI-generated content may be incorrect.

Figure X4. The general trends of the most abundant *Culicoides* species seasonal dynamics observed during the 2009 - 2012 period.

The 14895 trapping sessions produced data with large variability with the minimum of 0 and maximum of 491115 specimens caught by a single trap. Generally, the bulk of captures had low amounts of *Culicoides spp*. across all species with the median of around 6 individuals caught per trap, indicating that during the greater part of the year, a single trap would yield relatively few numbers. However, the mean capture value is significantly higher (mean = 168.4), which suggests that a small number of sessions captured an extremely high number of *Culicoides spp*., thus raising the overall average. This is consistent with the right-skewed distribution we can observe in fig. X5.

A graph of a number of blue and red bars

AI-generated content may be incorrect.Figure X5. Distribution of *Culicoides spp*. capture values from 2009 to 2012 in France. The values are represented on a logarithmic scale to reduce overdispersion.

* Technical Validation (unlimited length)

All identification included to the database were performed by entomological expert. The database was revised. Finally, all the authors carefully checked the complete database for possible technical failures and errors.

Des procédures de contrôle qualité sur l'identification des espèces ou la gestion des données ont-elles été mises en place ? Par exemple, avez-vous vérifié les identifications par plusieurs experts ou utilisé des méthodes alternatives comme la génétique ?

* Usage Notes (unlimited length)
* Code Availability
* References

1. Carpenter, S., Groschup, M. H., Garros, C., Felippe-Bauer, M. L. & Purse, B. V. Culicoides biting midges, arboviruses and public health in Europe. *Antiviral Res.* **100**, 102–113 (2013).

2. Tago, D., Hammitt, J. K., Thomas, A. & Raboisson, D. Cost assessment of the movement restriction policy in France during the 2006 bluetongue virus episode (BTV-8). *Prev. Vet. Med.* **117**, 577–589 (2014).

3. Häsler, B., Howe, K. S., Di Labio, E., Schwermer, H. & Stärk, K. D. C. Economic evaluation of the surveillance and intervention programme for bluetongue virus serotype 8 in Switzerland. *Prev. Vet. Med.* **103**, 93–111 (2012).

4. Gethmann, J., Probst, C. & Conraths, F. J. Economic Impact of a Bluetongue Serotype 8 Epidemic in Germany. *Front. Vet. Sci.* **7**, (2020).

5. Waret-Szkuta, A. *et al.* Economic assessment of an emerging disease: the case of Schmallenberg virus in France. *Rev. Sci. Tech. Int. Off. Epizoot.* **36**, 265–277 (2017).

6. Welby, S. *et al.* Effectiveness and Cost Efficiency of Different Surveillance Components for Proving Freedom and Early Detection of Disease: Bluetongue Serotype 8 in Cattle as Case Study for Belgium, France and the Netherlands. *Transbound. Emerg. Dis.* **64**, 1771–1781 (2017).

7. Kundlacz, C. *et al.* Bluetongue Virus in France: An Illustration of the European and Mediterranean Context since the 2000s. *Viruses* **11**, 672 (2019).

8. Mellor, P. S., Carpenter, S., Harrup, L., Baylis, M. & Mertens, P. P. C. Bluetongue in Europe and the Mediterranean Basin: History of occurrence prior to 2006. *Prev. Vet. Med.* **87**, 4–20 (2008).

9. Breard, E. *et al.* The epidemiology and diagnosis of bluetongue with particular reference to Corsica. *Res. Vet. Sci.* **77**, 1–8 (2004).

10. Thomas Balenghien *et al.* La surveillance des Culicoïdes en France. *Bull. Épidémiologique Santé Anim. Aliment.* 8–9 (2010).

11. Carpenter, S., Wilson, A. & Mellor, P. S. Culicoides and the emergence of bluetongue virus in northern Europe. *Trends Microbiol.* **17**, 172–178 (2009).

12. Gerbier, G. *et al.* Emergence Of Bluetongue In France 2000-2004. in (Cairns, Australia, 2006).

13. European Union (EU). Commission Regulation (EC) No. 1266/2007 on implementing rules for Council Directive 2000/75/EC as regards the control, monitoring, surveillance and restrictions on movements of certain animals of susceptible species in relation to bluetongue. *O.J. L* **L**, 37–52 (2007).

14. Mathieu, B. *et al.* Simultaneous quantification of the relative abundance of species complex members: application to Culicoides obsoletus and Culicoides scoticus (Diptera: Ceratopogonidae), potential vectors of bluetongue virus. *Vet. Parasitol.* **182**, 297–306 (2011).

15. Balenghien, T. *et al.* L’activité des populations de Culicoides en 2012 et bilan des quatre années du dispositif de surveillance. *Bull. Epidémiologique* (2013).

16. Cuéllar, A. C. *et al.* Modelling the monthly abundance of Culicoides biting midges in nine European countries using Random Forests machine learning. *Parasit. Vectors* **13**, 194 (2020).

17. Cuéllar, A. C. *et al.* Monthly variation in the probability of presence of adult Culicoides populations in nine European countries and the implications for targeted surveillance. *Parasit. Vectors* **11**, 608 (2018).

18. Ségard, A. *et al.* Schmallenberg virus in Culicoides Latreille (Diptera: Ceratopogonidae) populations in France during 2011-2012 outbreak. *Transbound. Emerg. Dis.* **65**, e94–e103 (2018).

19. Villard, P. *et al.* Modeling Culicoides abundance in mainland France: implications for surveillance. *Parasit. Vectors* **12**, 391 (2019).

20. Venter, G. J. *et al.* Comparison of the efficiency of five suction light traps under field conditions in South Africa for the collection of *Culicoides* species. *Vet. Parasitol.* **166**, 299–307 (2009).

21. Probst, C., Gethmann, J. M., Kampen, H., Werner, D. & Conraths, F. J. A comparison of four light traps for collecting Culicoides biting midges. *Parasitol. Res.* **114**, 4717–4724 (2015).

22. Mathieu, B. *et al.* Development and validation of IIKC: an interactive identification key for Culicoides (Diptera: Ceratopogonidae) females from the Western Palaearctic region. *Parasit. Vectors* **5**, 137 (2012).

23. Delecolle, J.-C. Nouvelle contribution à l’étude systématique et iconographique des espèces du genre Culicoides, (Diptéra): (Cératopogonidae) du Nord-Est de la France. (1985).

24. Van Ark, H. & Meiswinkel, R. Subsampling of large light trap catches of Culicoides (Diptera: Ceratopogonidae). *Onderstepoort J. Vet. Res.* **59**, 183–189 (1992).

25. Copernicus Climate Change Service. ERA5-Land hourly data from 1950 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS) https://doi.org/10.24381/CDS.E2161BAC (2019).

26. Didan, K. MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006. NASA EOSDIS Land Processes Distributed Active Archive Center https://doi.org/10.5067/MODIS/MOD13Q1.006 (2015).

27. GBIF. GBIF: The Global Biodiversity Information Facility (year) What is GBIF?. Available from https://www.gbif.org/what-is-gbif. https://www.gbif.org/what-is-gbif.

28. Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **3**, 160018 (2016).

29. Balenghien, T. *et al.* A spatio-temporal Culicoides species dataset produced by the French surveillance program from 2009 to 2012. doi:10.15468/8vepgt.

30. Wieczorek, J. *et al.* Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. *PLoS ONE* **7**, e29715 (2012).

31. Darwin Core Maintenance Group. Darwin Core Quick Reference Guide. Biodiversity Information Standards (TDWG). https://dwc.tdwg.org/terms/ (2025).

32. Danielson, J. J. & Gesch, D. B. *Global Multi-Resolution Terrain Elevation Data 2010 (GMTED2010)*. *Open-File Report* https://pubs.usgs.gov/publication/ofr20111073 (2011) doi:10.3133/ofr20111073.

33. Cornes, R. C., Van Der Schrier, G., Van Den Besselaar, E. J. M. & Jones, P. D. An Ensemble Version of the E‐OBS Temperature and Precipitation Data Sets. *J. Geophys. Res. Atmospheres* **123**, 9391–9409 (2018).

34. Ministère de l’Agriculture, de l’Agroalimentaire et de la Forêt (MAAF), Service de la statistique et de la prospective (SSP). Recensement Agricole 2010, France métropolitaine. (2010).

35. European Environment Agency. CORINE Land Cover 2018 (raster 100 m), Europe, 6-yearly - version 2020\_20u1, May 2020. European Environment Agency https://doi.org/10.2909/960998C1-1870-4E82-8051-6485205EBBAC (2019).

36. European Environment Agency. Biogeographical regions. https://www.eea.europa.eu/en/datahub/datahubitem-view/11db8d14-f167-4cd5-9205-95638dfd9618 (2016).

37. NASA Land Processes Distributed Active Archive Center (LP DAAC). Application for Extracting and Exploring Analysis Ready Samples (AppEEARS) API. https://appeears.earthdatacloud.nasa.gov/api/ (2025).

38. Forsythe, W. C., Rykiel, E. J., Stahl, R. S., Wu, H. & Schoolfield, R. M. A model comparison for daylength as a function of latitude and day of year. *Ecol. Model.* **80**, 87–95 (1995).

39. Hijmans, R. J. meteor: Meteorological Data Manipulation. 0.4-5 https://doi.org/10.32614/CRAN.package.meteor (2019).

* Author Contributions

Hammami P wrote the paper, collected the post-hoc data and formatted the dataset. Baltusyte I and Taconet P performed the descriptive presentation of the dataset, Delécolle JC, Setier-Rio ML, Mathieu B, Venail R, Balenghien T and Garros C carried out the entomological data collection. Balenghien T and Garros C managed the entomological surveillance network. All authors reviewed the paper.

* Competing Interests

The author(s) declare no competing interests.

* Acknowledgements (optional)

We acknowledge the E-OBS dataset from the EU-FP6 project UERRA (https://www.uerra.eu) and the Copernicus Climate Change Service, and the data providers in the ECA&D project (https://www.ecad.eu)

* Figures
* Tables