

Metadata concerning the animal data stored under the directory “Animal CSI”, where CSI means “Climate Sensitive Infections”.

Thierfelder T.¹, Omazic A.², Berggren C.², Albihn A.², 2019-09-20

1: *Dept. of Energy & Technology, Swedish University of Agricultural Sciences, Uppsala, Sweden*

2: *Dept. of Chemistry, Environment, and Feed Safety, National Veterinary Institute, Uppsala, Sweden*

Contact person and citations

For any questions concerning the published data, please make contact with Tomas Thierfelder at Tomas.Thierfelder@slu.se.

Kindly refer to the title, date, and list of authors provided above in all publications that utilise CLINF animal diseases data.

Diseases

With most diseases addressed by CLINF being zoonotic, meaning that they may be transmitted from animals to humans, and that they hence occur in humans as well as in animals, the term “Animal CSI” indicates that the corresponding diseases have been observed in animals and reported via animal diseases report systems. In the CLINF repository, animal CSI data are stored separately, where the respective diseases data are stored in separate folders. These folders have been given abbreviated names of the associated diseases, where the names of folders and diseases are:

ANA = Anaplasmosis, BAB = Babesiosis, BRU = Brucellosis, ERY = Erysipelothrix, LEP = Leptospirosis, LIS = Listeriosis, RAB = Rabies, TRI = Trichinellosis.

Animal diseases data are either reported clinically or via laboratories, in written form or via digital report systems. This introduces a potential of overlay error since single cases of diseases may be reported twice, and since older written report systems may overlap with the implementation of digital dittos. CLINF has spent much effort into reducing such sources of error to their minimum.

The spatiotemporal coverage of animal diseases data is relatively uncertain as compared with human diseases data. Where relatively few cases of reportable human diseases are supposedly missed, the number of unreported cases of animal diseases is largely unknown. This implies that human diseases data may be considered as actually approximating the generic situation, at least to a larger extent than what may be expected by animal diseases data. A proper interpretation of CLINF animal diseases data hence implies that animal diseases “at least” have been observed as indicated in the associated database, whereas the situation outside these indications is more uncertain than in the case of human diseases data.

Spatial resolution

The data covering animal diseases are constituted by empirical observations concerning individual cases reported to authorities in the six nations of:

GRE = Greenland, ICE = Iceland, NOR = Norway, SWE = Sweden, FIN = Finland, RUS = Russia.

The national administration of reported diseases is managed per report district, where the size of a typical report district approximately equals the size of counties everywhere except in Russia, where diseases mainly are reported per oblast, republic, or autonomous region. Hence, the smallest possible spatial resolution of CLINF animal CSI data is report district – see *Appendix A* for a list of CLINF report districts. In Greenland and Iceland, the entire nations constitute one report district,

although confined to their respective coastal regions. In Greenland, animal diseases are predominantly reported in the southern and western coastal regions, whereas Iceland is very sparsely covered with CLINF animal CSI data.

Temporal resolution

In the temporal domain, an attempt has been made to cover the 30-year climate reference period 1987 - 2017 with diseases data. This ambition has been more or less successful depending on national differences regarding the inclusion of different diseases in their respective lists of reportable diseases (diseases that should be reported by law). As a consequence, the time-period covered by CLINF animal CSI data varies with diseases and nation, but is presented through the entire period of 1987 – 2017 in the CLINF database. In a few cases, zero annual diseases frequencies have been reported in some report districts. These zeros are meant to represent the absence of the particular disease in the official records, but have been removed by CLINF since the unrecorded presence of diseases cannot be precluded. As already discussed, the overall interpretation of this is that CLINF animal diseases “at least” have been observed as indicated with non-zero values, whereas the situation besides these indications is “rather uncertain”.

With diseases reported case-by-case, the exact reporting dates of every single case are available in the CLINF diseases database, although not publicly disseminated. This temporal case-by-case resolution may be scaled up into any other temporal resolution such as monthly or annual cumulations of cases. The CLINF temporal standard resolution of animal CSI data is *the number of annually cumulated cases per report-district*. With the standard report-format of human diseases being “incidences” that principally express the number of cases per capita, incidences cannot be reported in the case of animal diseases since population numbers are basically unknown.

Data disposition

In the animal disease’s directory of the CLINF GIS database, in the respective XXX diseases folders, two different Excel files are to be found: “XXX_list_freq_all”, and “XXX_matrix_freq_all”.

XXX_list_freq_all

In “XXX_list_freq_all”, the total 1987 - 2017 cumulation of cases (Count) are stored per nation, report district, and year, where Distr_code is the official abbreviated identity of report districts. In addition, WGS84 longitudes (Long) and latitudes (Lat) of the geometric centroid of report districts are provided. XXX is symbolising the respective diseases listed above.

XXX_matrix_freq_all

In “XXX_matrix_freq_all”, data are dispositioned with annual data per column, with one column per year during the period of 1987 - 2017. The head of annual columns hence indicate the exact year of accumulation, which means that the associated variable name (Year) is provided in the integer interval [1987, 2017].

In the files “XXX_matrix_freq_all, data may be sorted with respect to the longitude and latitude of district centroids, from western Greenland to Eastern Siberia, and from the south towards the north. If so, any eventual migration of diseases along climate gradients, through the observed period of time, may be visible as more or less diagonal traces of non-zero counts. When the migration of diseases is formally tested, the associated statistical model should infer the existence of such traces.

List of CLINF animal CSI variables

The following factors/variables are included in the CLINF standard disposition of animal diseases data:

CSI = [ANA, BAB, BRU, ERY, LEP, LIS, RAB, TRI]

Nation = [GRE, ICE, NOR, SWE, FIN, RUS]

District = Diseases report districts, see *Appendix A*.

Distr_code = Official abbreviated identities of diseases report district, see *Appendix A*.

Year = Year of annually cumulated disease cases.

Count = Number of disease cases reported per districts across the entire reference period 1987 - 2017.

Long = WGS84 longitude of the geometric centroid of report districts.

Lat = WGS84 latitude of the geometric centroid of report districts.

Appendix 1

[CLINF diseases report districts](#) (simply click the hyperlink).