CLIDATA

In Ostrava 18/06/2013

Content

Introduction

The Clidata system is primary intended for archiving of climatology data, for the data quality control and for administration of climatology stations and station observations. The System was designed to replace the old CLICOM system, which has been used in Czech Republic from 1993-2000.

The system is designed for the Oracle database environment, which defines simple and secure access to stored data.

By virtue of the system flexibility, easy administration and multi-language support, the system is capable of set up in any foreign country and for any meteorological service. The system has been operationally used in Czech Hydrometeorological Institue for 13 years and it is successfully installed in more than 20 other countries.

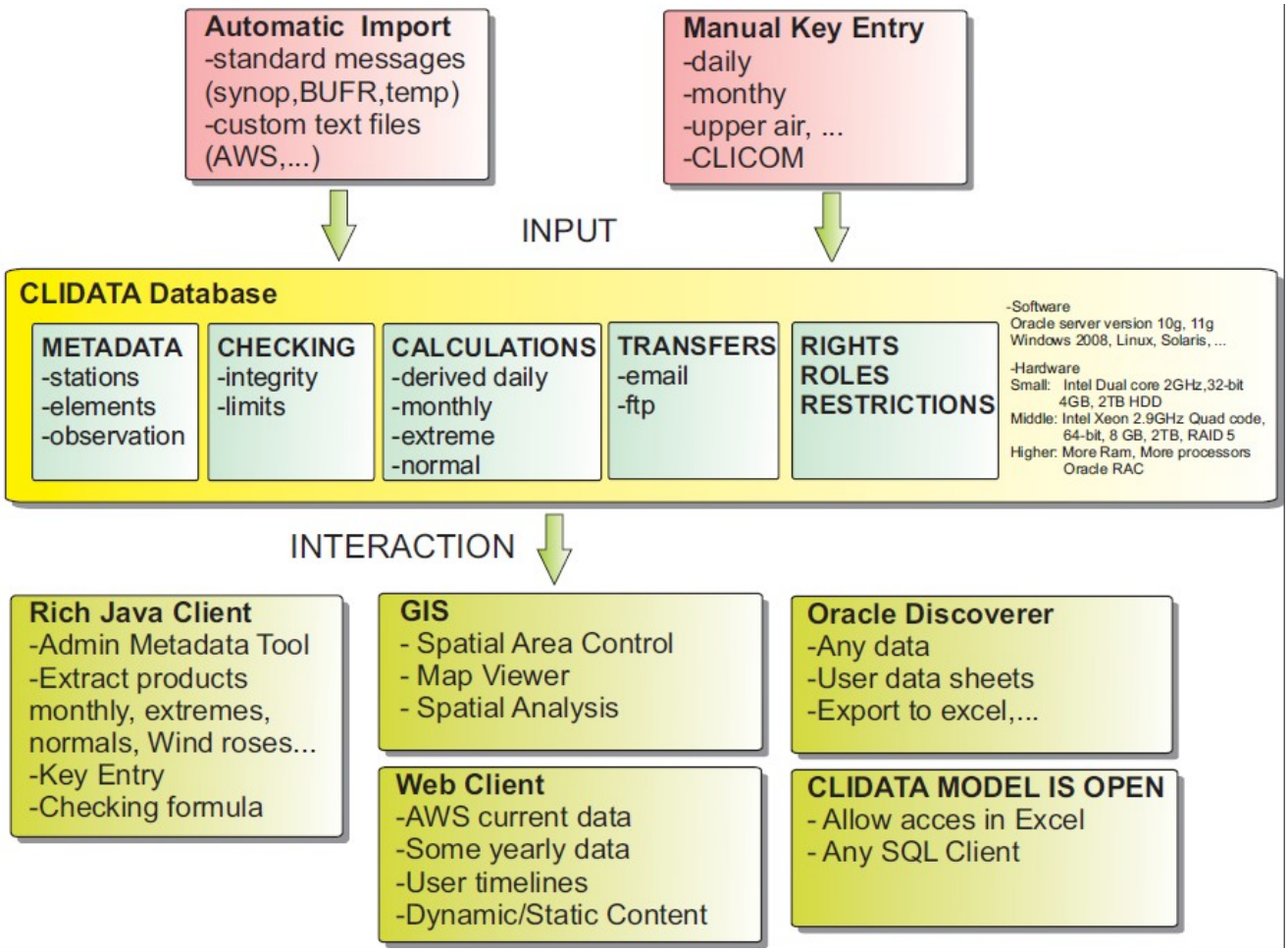
The system is particularly user-friendly during the definition of stations, station observations and manual key entry of the data. The system facilitates the population of data from automated (real-time) stations as well as the definition of personalised key entry forms.

Clidata New Features

Since New Clidata Client implemented in Java is available only for short time the following list summarizes the new feature of the Clidata Client Application:

- Rich Java Client – covers all functionality of old Oracle Developer Forms Client
- Web Client – optional tool for creation of charts and table reports
- Geography polygons
- Hydrological measurement profile curves
- Geography Obstacles
- The system of data quality checking is changed – the data checking is executed in days (no in month as before)
- New locking and quality control flags
- The smallest time unit for definition station geography and definition of station observation is one minute. As the open station observation the data 31.12.3999 23:59 is used
- New types of data checking and formulas were added
  - New variables in quality control and daily data formulas
  - Regular data checks
  - There is possible to execute batch data checking
- More tight security rules.
- New type of key entry forms
- Monitoring of database running
- Storing information about observers
- Possibility to store any kind of files for a station (Excel, Word …)
- Import of monthly data
- Spatial analysis was improved
- New GIS product was created
- Cross validation between rainfall and intensity rainfall
- More Wind roses charts
- Enhancements in import methods
- New Decode methods for standard WMO messages (BUFR, CLIMAT, etc..)
- Generating of standard WMO messages (BUFR CLIMAT, BUFR SYNOP, WWR)
- Regular data edit and copy from reference station
- WMO Information for station: isobaric level, elevation of barometer, WMO station name
- Automatic geography coordinate translation (Eg. Geography coordinates → Gauss)

Structure of Clidata Application



Clidata basic modules:

- Clidata Database – main storage, automatic processing
- Rich Java Client – main administrative tool
- Oracle Discoverer – data extraction tool

Clidata optional modules:

- Web Client – data reporting tool
- GIS – mapping tool
  - Map Viewer – simple mapping application
  - Spatial Analysis – advanced mapping application
  - Spatial Area Control – data checking in the map

Clidata Database

Clidata database is the basic module. In this module all database scheme (Tables, Views, Table Relations, etc.) is defined. In the tables all *data* and *metadata* is stored. Main metadata information for the system is:

- Stations geography
- Elements
- Stations observation

Main data information is:

- Daily data (Each 10,15,30,1hour,... data and daily values)
- Yearly, Monthly, 10 days, 5 days data
- Extremes
- Normals
- Upper Air Data
- Meteorological Phenomena
- Totalizers
- Intensity Rainfall (one minute precipitation)

Moreover the database stores *programs* which are responsible for data manipulations. These programs are running in the background as so called *jobs*. Jobs are responsible for:

- Automatic Imports
- Automatic Calculations, derived daily data, monthly data, extremes, normals, etc.
- Automatic Exports and Transfers

In order to streamline changes in individual application versions and source database it was decided to develop in ORACLE only. The focus on the Oracle provides the following advantages:

- High data security – data backup, data protection, data restrictions, disaster recovery features
- Optimized data processing - advanced oracle features for high performance, high scalability
- Multi-platform deployment (MS Windows, Linux, Solaris,...)
- Professional Support
- Proved by history
- High probability of future backward compatibility
- Utilization for many years without need of changes, core parts of the application did not face any problems when upgraded.

The source database versions changed gradually and now the system supports Oracle 9i to Oracle 11g versions.

Hardware requirements are optional and are depended on the system usage and amount of data stored. The system can be deployed from:

Small

Intel Dual Core, 1TB HDD, 4GB RAM, 32 bits OS

Middle

Intel XEON Quad Core, 2TB HDD, 8GB RAM, 64 bits OS

Large

Oracle RAC, more processors and more RAM

Rich Java Client

This is the main administration tool. The Rich Java client is developed in Java. The decision for Java was made for the following reasons:

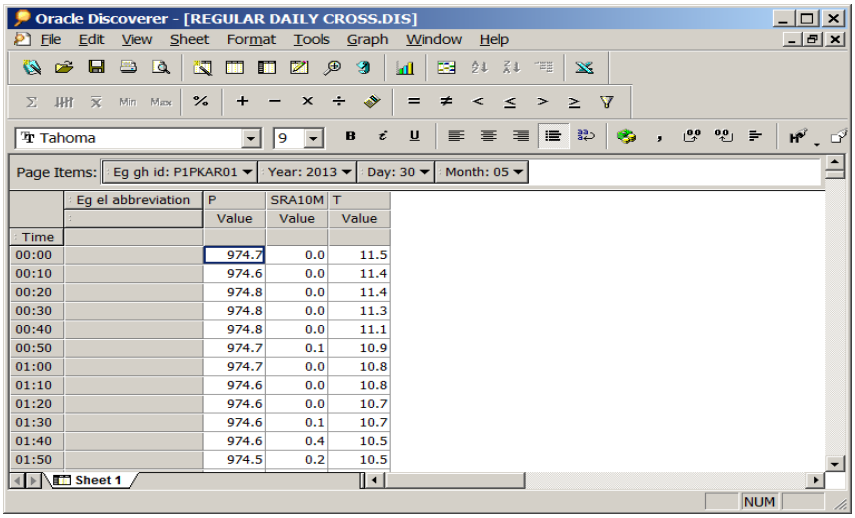
- Leading development technology
- Multi-platform deployment
- Huge Java development community
- Proved by history
- Owned by Oracle

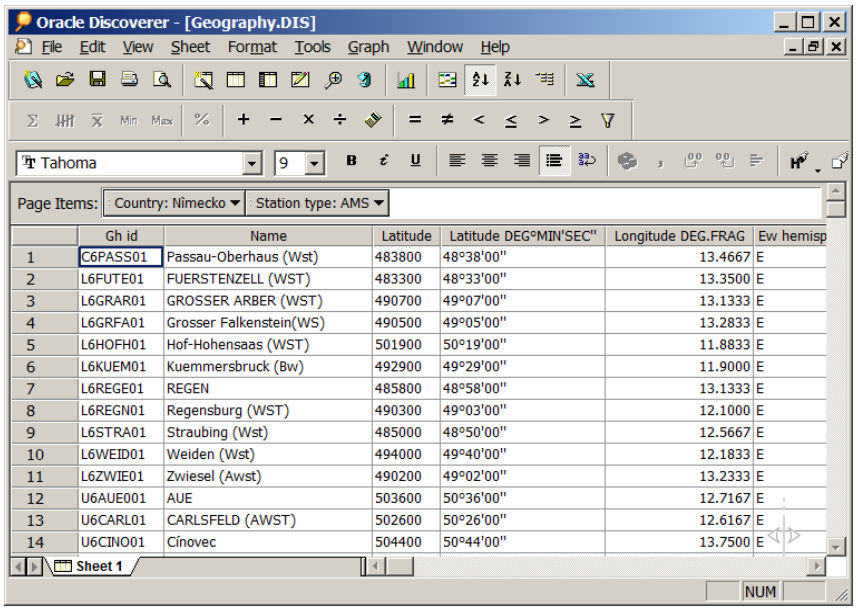
The Rich Java client manages all application parts. The main modules consists of:

- Stations – Geography and observation definition, definition of other station information
- Metadata – definition of metadata:
  - Countries
  - Elements
  - Units
  - Time schemes
  - Calculation and control formula
  - Etc
- Kef(Key entry forms) user defined key entry forms for data input or data corrections
- Discoverer – used for easy data extraction by means of user defined worksheets.
- Products – overview of main derived data
  - Yearly, Monthly, 10-days, 5-days data
  - Extremes
  - Normals
  - Wind Roses
  - Etc
- Administration – definition of users and rights, data imports and exports.

Oracle Discoverer

Is the data mining tool provided by Oracle. This tool is used for creation of the data worksheets like in MS Excel. Oracle Discoverer eases finding the desired data. You can view the data quickly, without waiting for the computer to search the entire database. Views of the data are easily comprehended, accessible and user friendly. You can analyse the data using a wide range of methods especially data meeting the required conditions or falling within a certain specified interval. You can sort and compare the data as well as share it with other users, or export it to other applications.



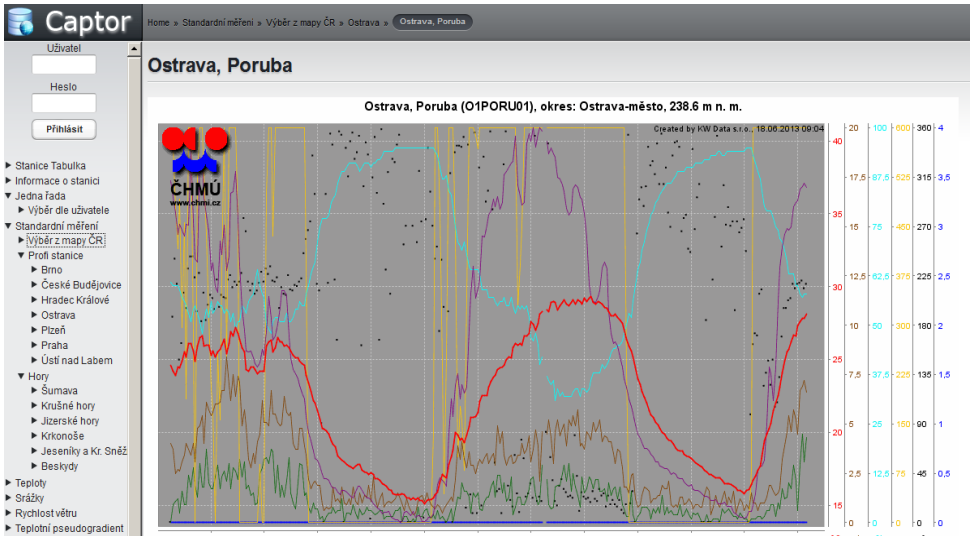


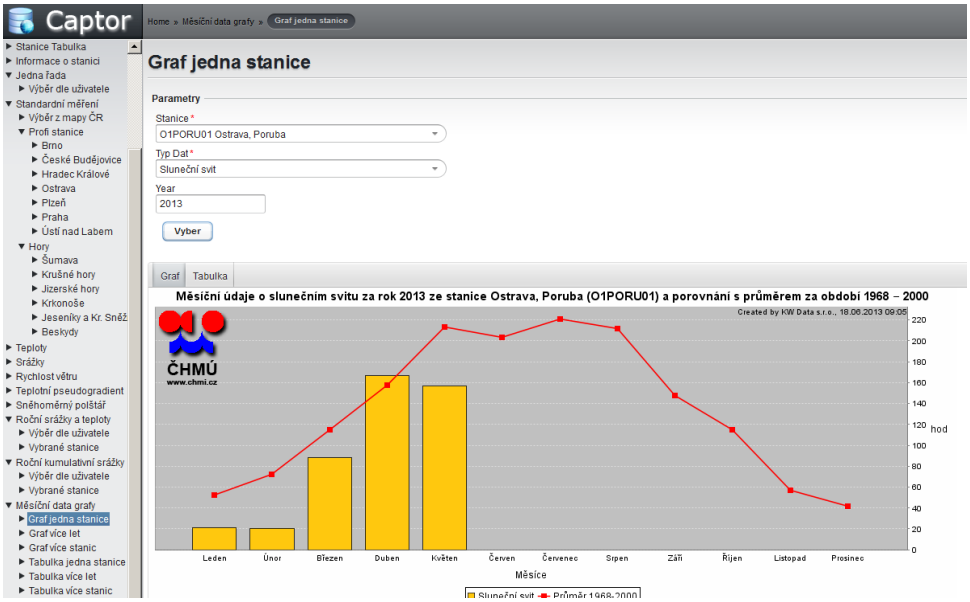
Web Client

Is the data reporting tool. This tool is used for creation of predefined charts and tables. This tool is developed in Java. Application server is necessary for running the Web Client. The Application Server can be any supporting Java Container. It can be installed on the same Hardware configuration as Clidata database or in the separate server. The data reports are then available using common Web Browser (Internet Explorer, Mozilla, Opera, Chrome, etc...).

Data are viewed easily an user friendly. Products available using Web Client are as follows:

- Basic Station Geography
- Actual Daily Data
- Historical Time Series
- Monthly data and charts
- Extremes
- Normals
- Cumulative precipitation charts
- Etc.

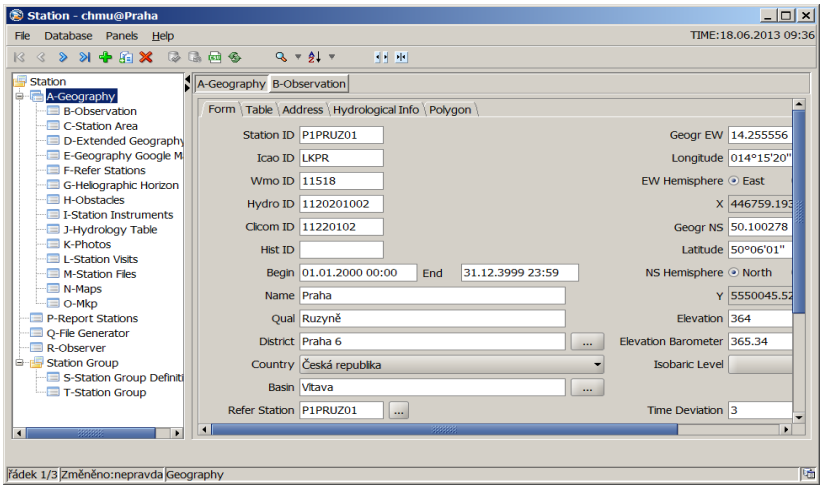


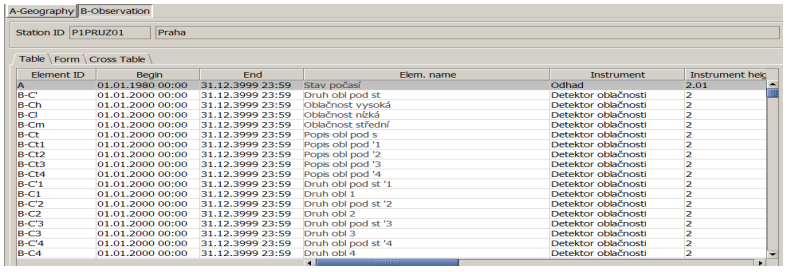


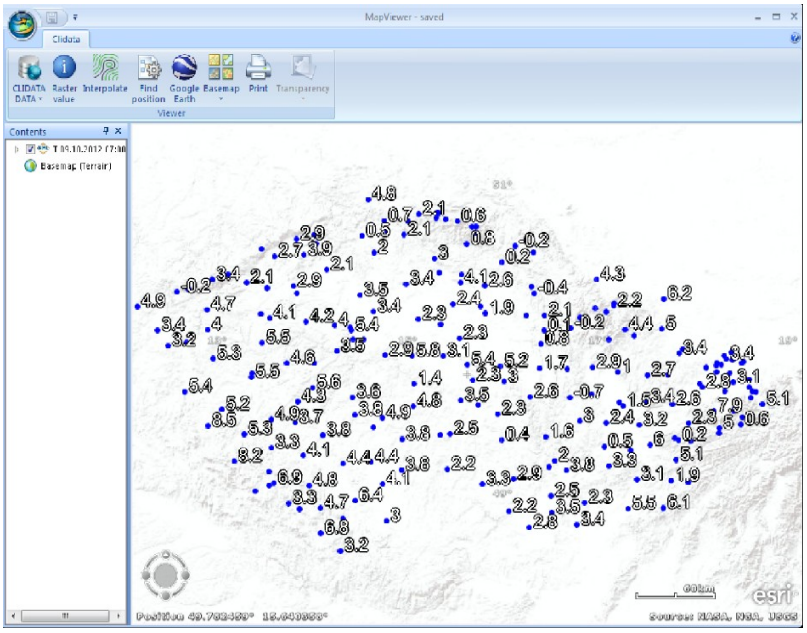
Map Viewer

Is the tool user for simple mapping. Available features:

- Simple station Maps
- Display data stored in Clidata or in local text Files
- Make surface maps from station data
- Can use WMS Layers
- Can use static climatological maps
- Export data in KML format(Google Earth)
- Using ArcGIS Explorer freeware from ESRI



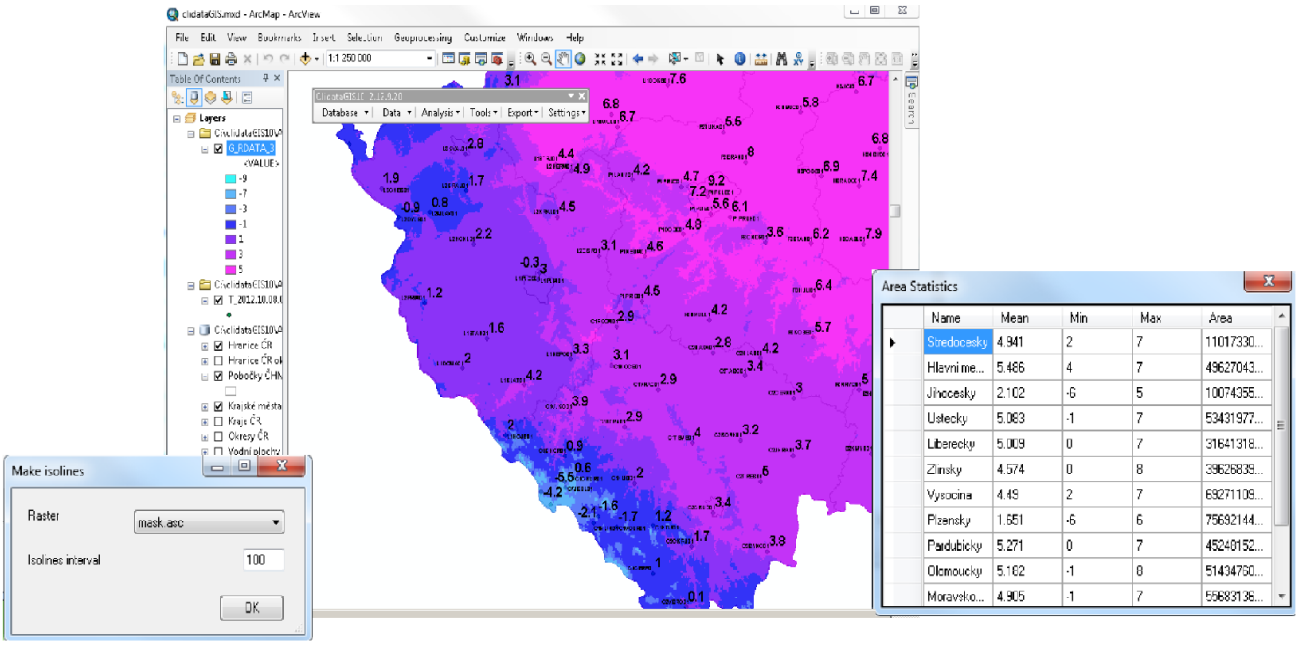


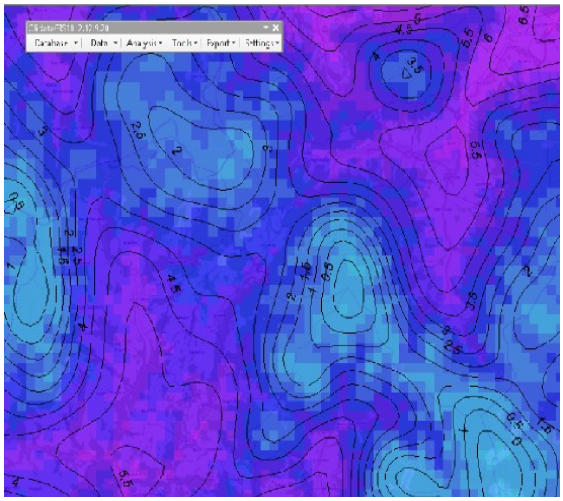

MapViewer - saved
Clidata
CLIDATA DATA
Raster value
Interpolate
Find position
Google Earth
Basemap
Print
Transparency
Viewer
Contents
Basemap (Terrain)
esri

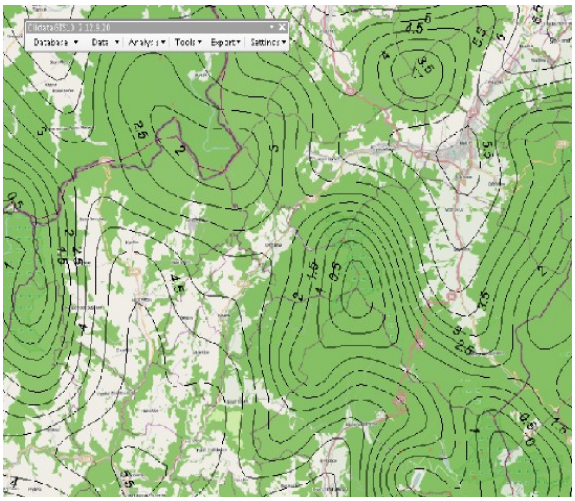
Clidata GIS and Spatial Analysis

Is the Advanced Clidata mapping Tool. With This tool you can:

- Create any map using Clidata application
- Make surfaces from station Geography
- Work with any data in Clidata including Daily, Monthly, Extremes, Normals, etc...
- Includes interpolation methods: IDW, LLR, Clidata DEM, Clidata DEM R2
- Smoothing surfaces
- Making isolines
- Calculate zone statistics
- Calculate rasters
- Simple user friendly export to many formats including PNG and PDF

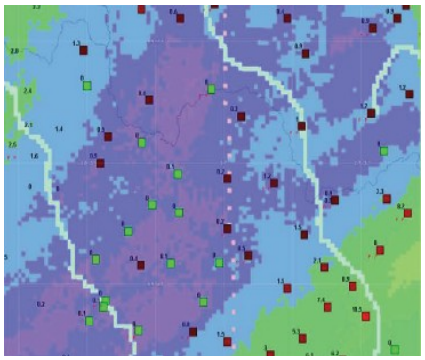


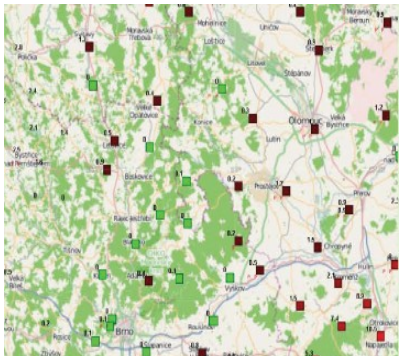


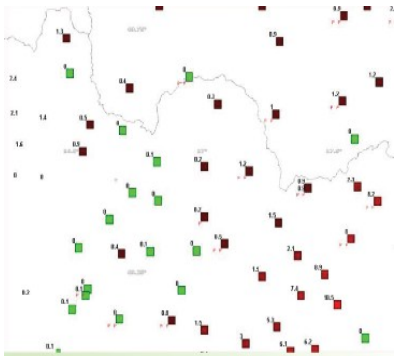


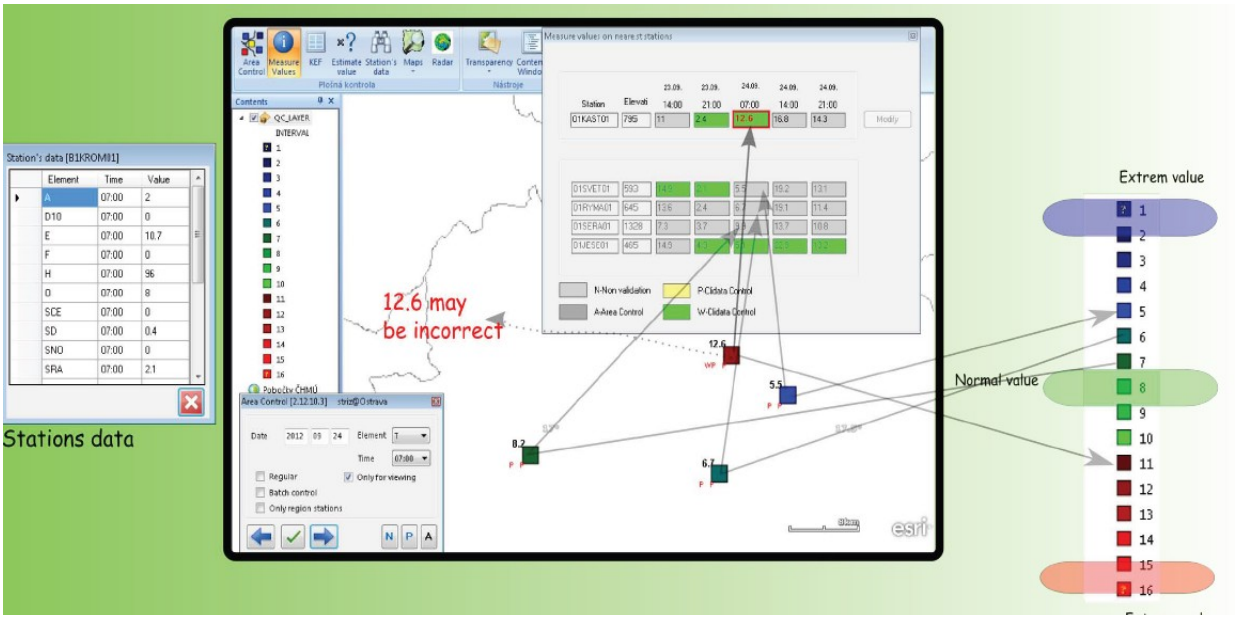
Clidata Area Control

Is the highest quality control level in Clidata. The values are crated as square fields meaning the interval of probability of exceeding. Quality control can be processed automatically or by human. Every station is compared with its surrounding stations. The value can be estimated by linear regression from its surrounding stations.









Station Geography

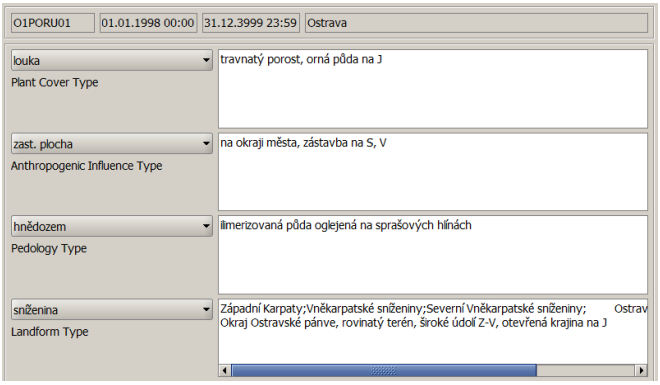
The information on station geography is available be means of well-arranged forms. The main features of this form have been adopted from the old Clicom system. The basic information has been extended and currently the users can maintain all the station documentation in one place. This definition of the station is the basis for data input checking. It is not possible to store the data for an undefined station.

The following additional information can be managed for station:

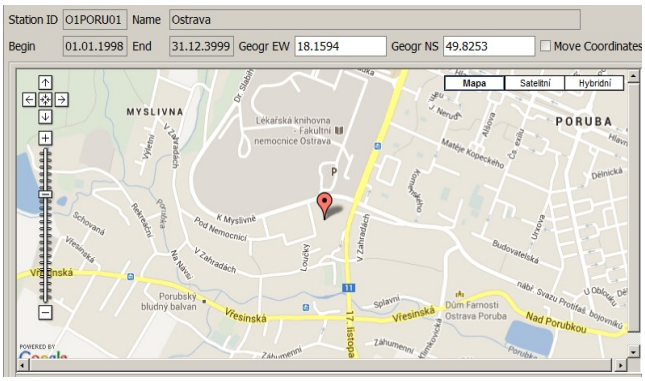
- Basic station geography
  - Station ID
  - Name
  - Period of measurement
  - Longitude
  - Latitude
  - Elevation
  - Country
  - Region
  - District
  - Other station Ids
  - Elevation
  - Address
  - Etc.
- Station Observation
  - Elements
  - Period of measurements
  - Time Scheme
  - Instrument
  - Calculations
  - Interpolation

- Maps – the map pictures. Additional information can be added to the map.
- Photos – complete digital photo documentation for each of the stations.
- Heliographic horizon – the horizontal shape of the surrounding landscape, natural and artificial obstacles are recorded in this part of the station geography.
- Extended station information – the additional information about pedology, anthropogenic influence, vegetation cover and relief.
- Hydrological information – additional information related to hydrology
- Observers – the list of current and ex-observers.
- Station documentation – the various types of documentation, e.g. MS Word or Excel files.
- Instruments – the list of all station instruments
- Measurement profile curves – relation between water flow and water height
- Polygon information – for station representing the polygon this part stores another information of polygon area data
- Reference stations – different reference stations related to this station
- Obstacles – the influence of surrounding especially used for the influence of the wind measurement
- Station groups – station can become member of some group. Many products can be run for the group of stations instead for one particular station.

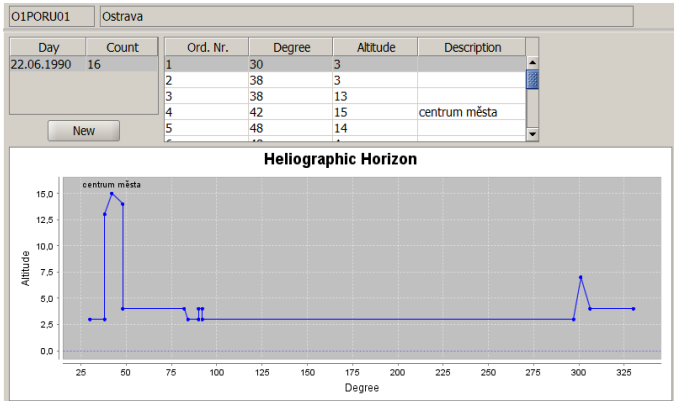
Additional station information



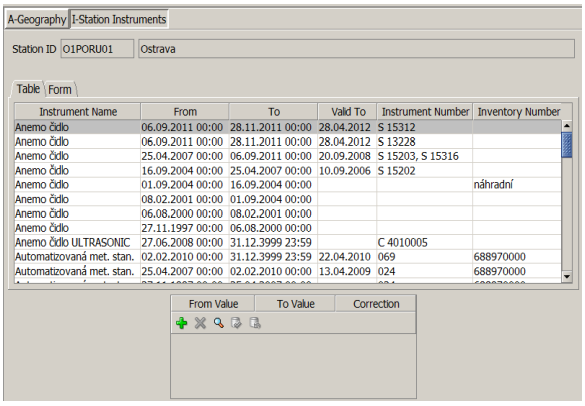
Google Maps



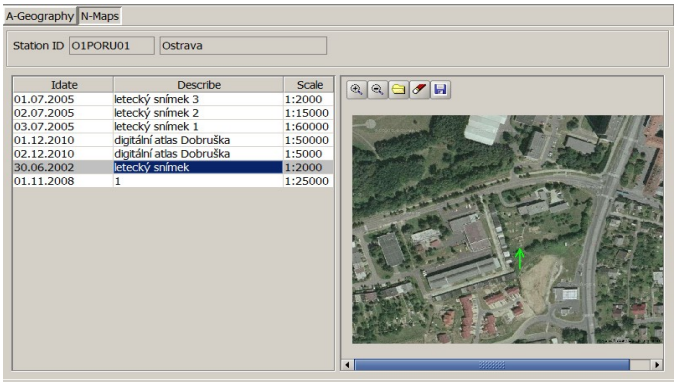
Heliographic horizon



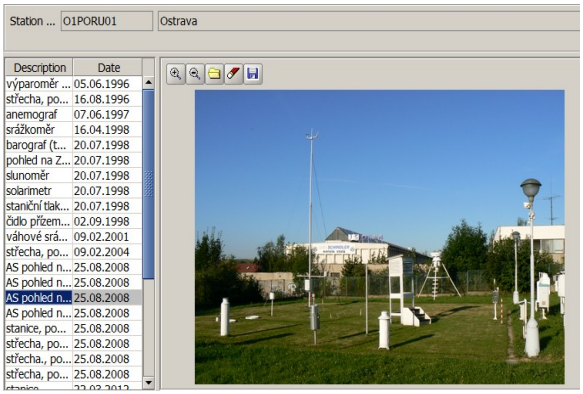
Station instruments



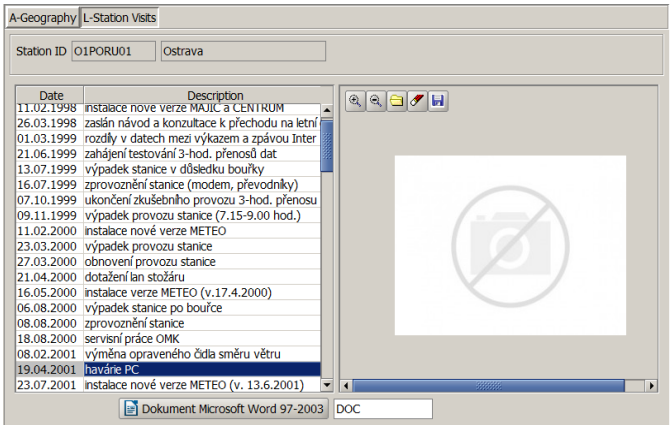
Maps



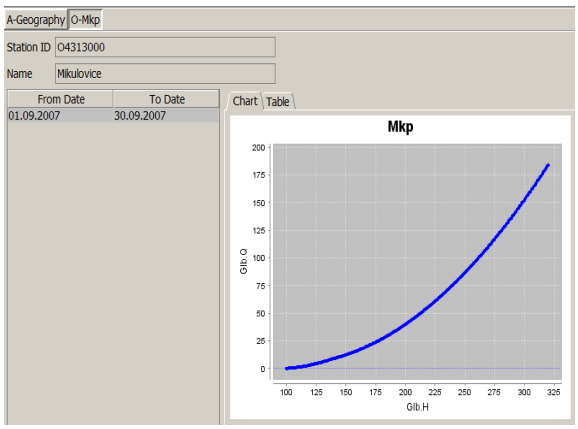
Photos



Station Files



Measurement Profile Curves



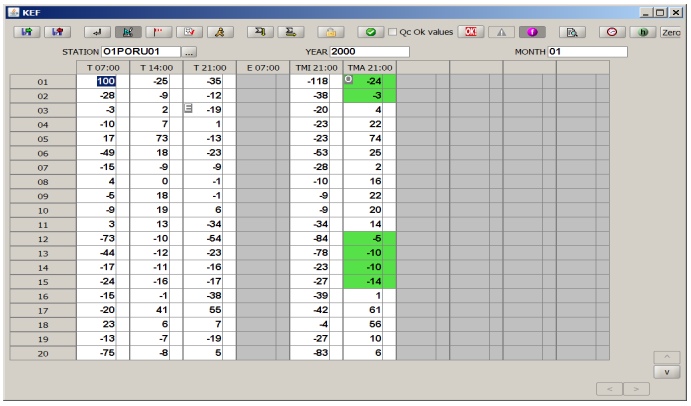
Key Entry Forms

In the Clidata system data can also be entered manually by transcribing data from paper forms filled by observers. The user can design his/her own key entry forms according to his/her needs and to precisely correspond to the original paper forms.

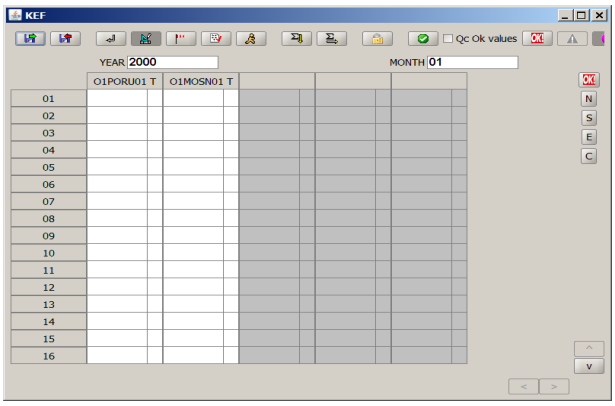
The data are keyed in well-arranged forms designed for user comfort. The keying system is based on the mature entry forms used in the Clicom system hence, users adapted to the new system easily.

In the key entry forms the users can take advantage of the Quality Control checksums in rows or in the columns.

Daily data layout: Day X Element+Time



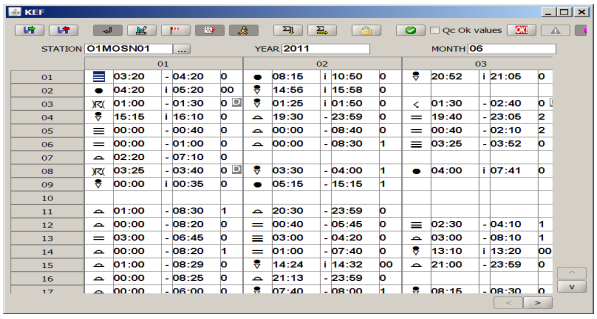
Daily data layout: Day X Station +Element



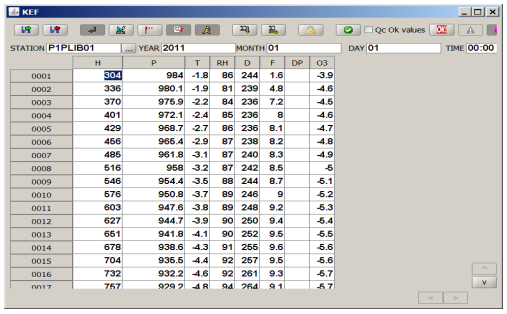
The arrangement of key entry form is up to user. Any layout of Station, Element, Year, Month, Day and time filed is possible.

Moreover specialized key entry forms for Rainfall Intensity, Meteorological Phenomena, Upper Air data is available.

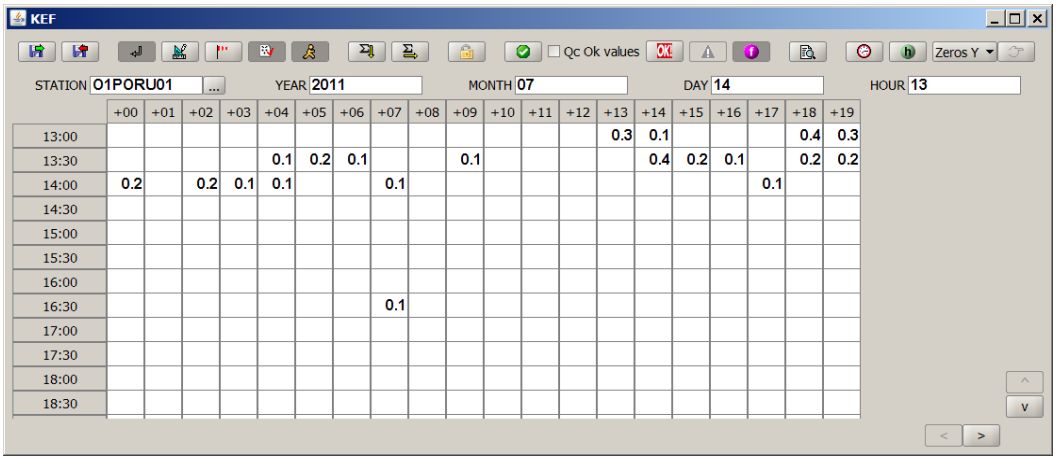
Meteorological phenomena



Upper Air Data



Intensity Rainfall

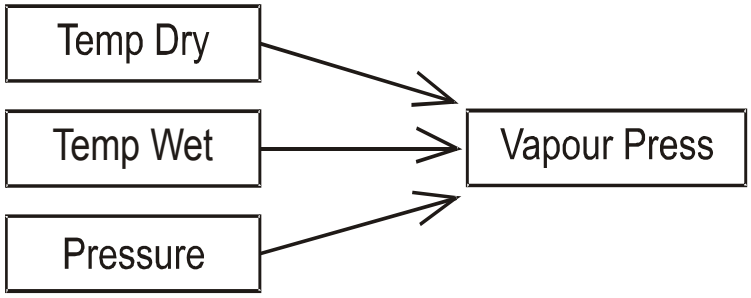


Automatic Derivative Calculation

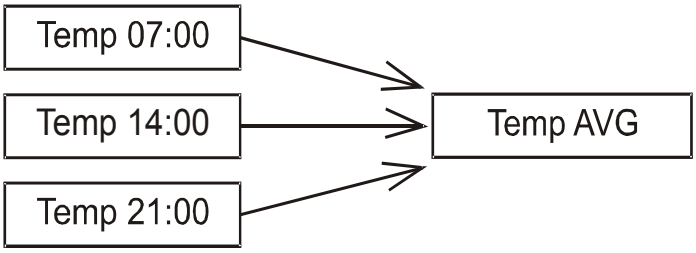
The Clidata system introduces user-defined automatic calculation of many derived elements from daily measured values. The derivatives system adopts a simple, user-friendly approach, such that the user defines only what he/she wants to calculate and the Clidata system carries out the calculations automatically without interaction. The definition of the formulas used for the calculation is very flexible, and the user is able to define very complex calculations.

The system is able to calculate:

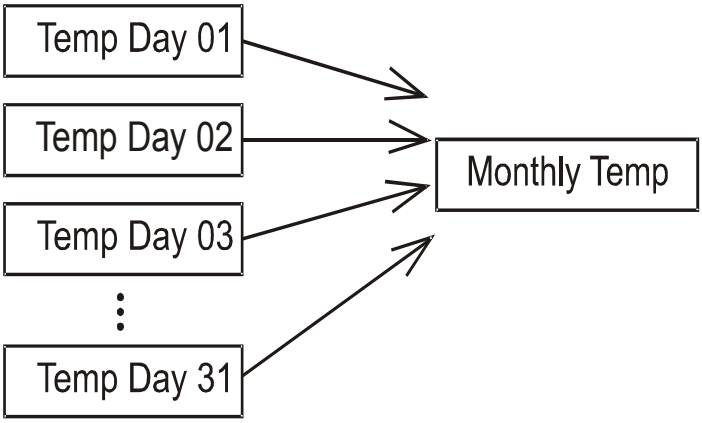
- Hourly values – the definition of the general formula related to a time during a day. For example, the calculation of relative humidity from the temperature, wet bulb temperature, pressure and wind speed, or the calculation of the absolute difference between two consecutive observation of the same element.



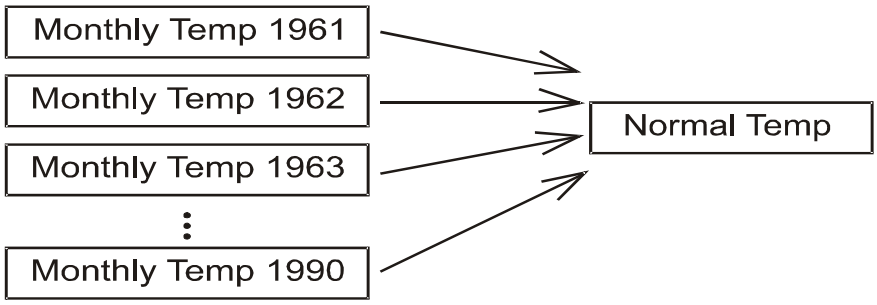
- Daily aggregated values – the definition of the general formula for daily maximum, minimum, sum or average.



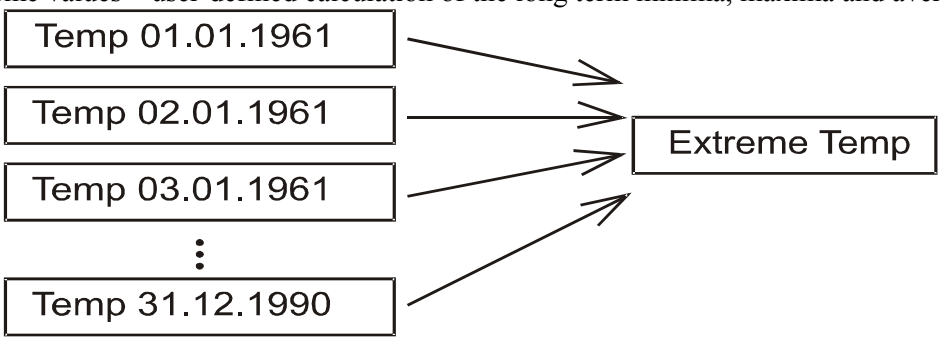
- Monthly/Yearly values – definition of the calculation, not only for the month and year, but also for the 1-3 decade or 1-6 pentade of the month. You can calculate the maximum, minimum, sum or average for these time periods. Alternatively, you can calculate the number of days which satisfy certain conditions over a period (e.g. the number of "summer days", where temperature for a "summer day" is above 25 °C).



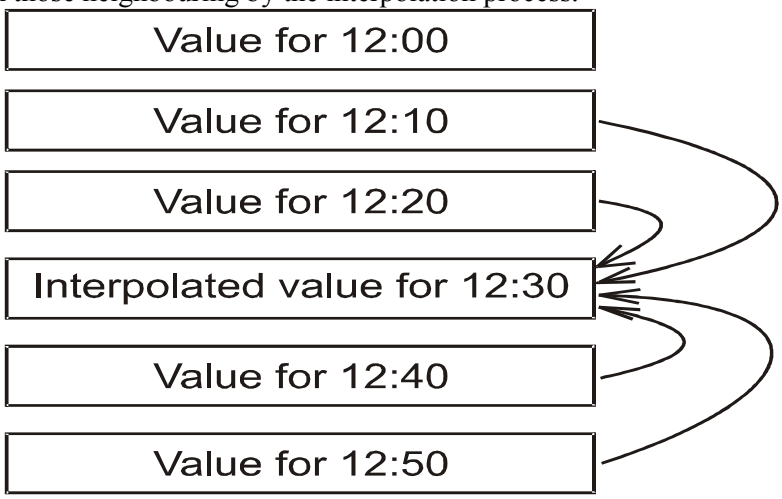
- Normals – definition of the long term calculation of the normal values; calculated for each month and for the year.



- Extreme values – user-defined calculation of the long term minima, maxima and averages.



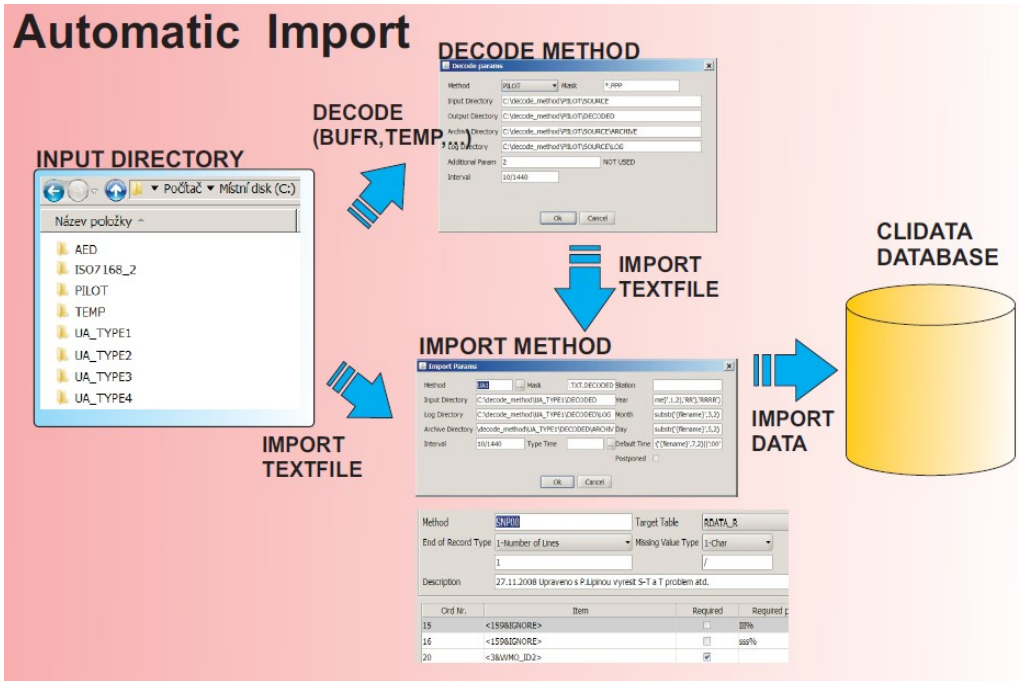
- Value interpolation – for the data measured regularly during a day the system can automatically infill missed values from those neighbouring by the interpolation process.



- Other automatic calculations
  - Days with phenomenon – the system automatically calculates the number of days in month and year in which the specific phenomenon has been observed. From this, the system also calculates (for long term period) the normal number of days for each of the months and for the year.
  - Normal number of phenomena – from monthly phenomena count the normal number is calculated
  - Standard isobaric levels for Upper Air Data – from raw data standard levels are calculated
  - Interpolated values for Upper Air Data – missing values are interpolated for all available pressure levels

Automatic Imports

The import of formatted text files is built into Clidata. The user defines their own method for each of the import text file types. The definition of the import method is very flexible and allows a wide range text file formats. After the creation of the import method the user sets up import job to carry out loading at regular intervals.



For standard WMO messages and not well formatted files the decode method is implemented in Clidata system. We can decode the following WMO messages:

- SYNOP
- CLIMAT
- CLIAMT TEMP
- BUFR
- Various Upper Air data formats (AED, Vaisala text files, …)

Decode method transfers the file into well formatted file suitable for Clidata import system. All structured files can be imported into Clidata. When reqired new decode methods can be implemented.

The import system is so flexible that all data in Czech Republic are imported by means of this system. Different kinds of the data are imported: daily data observed by observers and measured by automatic stations or data from the international exchange (the SYNOP messages), meteorological phenomena, one minute precipitation data etc.

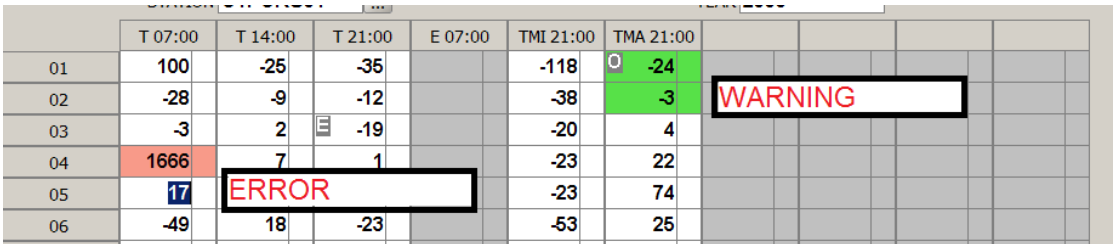
Data Quality Control

The data in Clidata are checked in the following levels. Basic levels are processed during key entry or during import of the data to the system. Advanced check are processed after the data are input into the system. Special application for Area control is the implemented as standalone module for data checking and is described in other part of this document.

Basic Checks

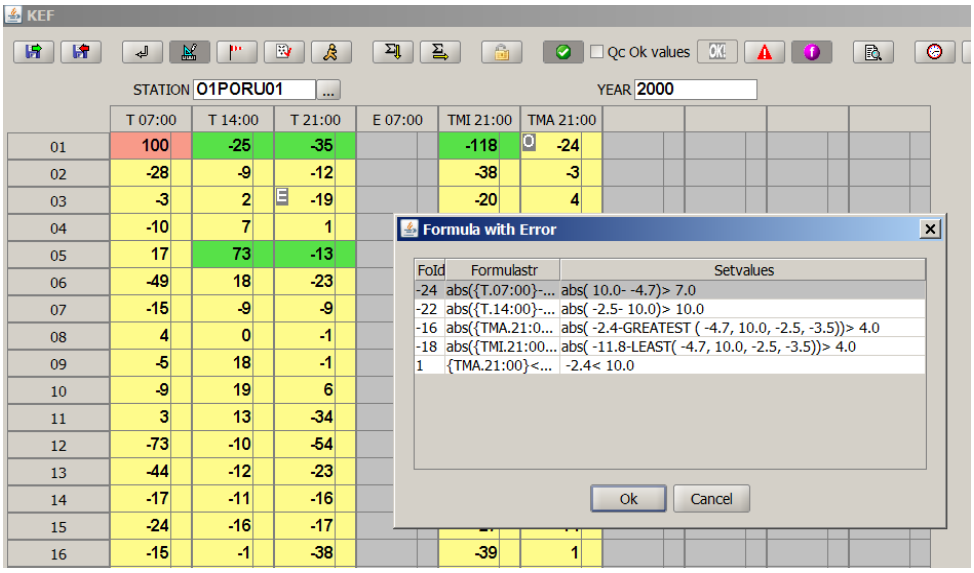
during key entry or Automatic import

- Integrity within the system – Station, element and observation definition must exists
- Error Limits – data outside this limits are not stored
- Warning Limits – data are stored, but are highlighted as suspicious
- Rounding and scale



Advanced Checks

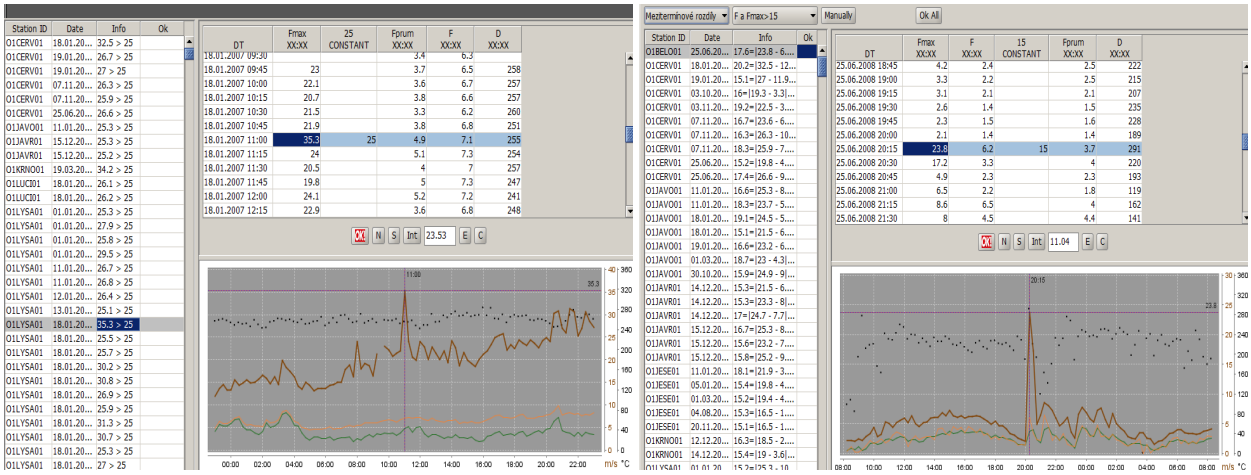
Data are checked by Quality Control Formula. User can define general formula with the specification of the relation between elements. The General formula can be very complex including any comparison and logical operators. The suspicious values are highlighted with different colours and errors or warnings are indicated. This kind of control is utilized after manual Key entry.



For many stations and long periods the Clidata system uses Batch Quality Control. This check are done by the same quality control formula. This utility is useful for checking historical and imported data.

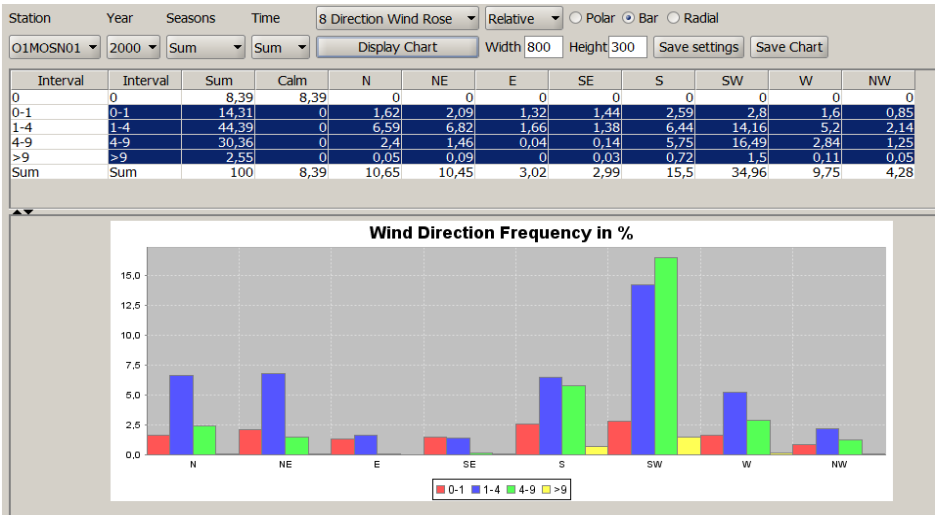
General quality control formula are not very effective for regular data (Data from Automatic weather stations).

Therefore Clidata introduces *Automatic checks* by simple formulas of actual regular data. Simple formulas are useful to find exceptional values above some predefined limits, peaks and unusual repeated values.

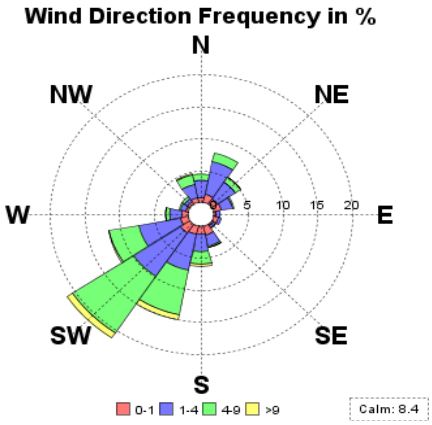


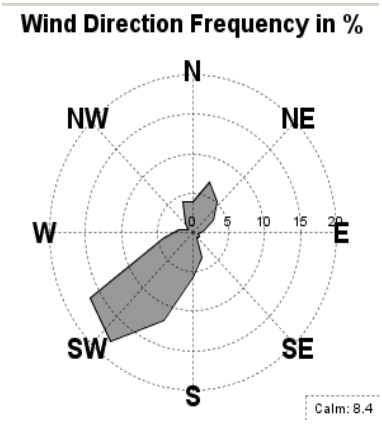
Wind Roses

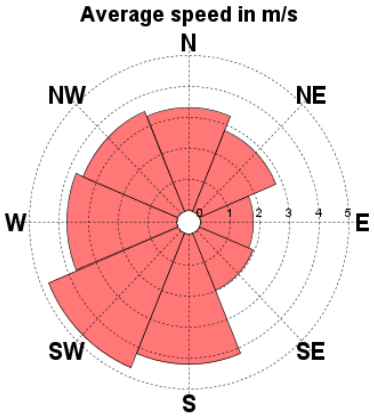
One of the often used product is wind rose creation. Clidata can create 8,12,16,18 and 36 direction wind rose. Wind rose is calculated for relative or absolute count of occurrences or for average wind speed.



Interval	Interval	Sum	Calm	N	NE	E	SE	S	SW	W	NW
0	0	8,39	8,39	0	0	0	0	0	0	0	0
0-1	0-1	14,31	0	1,62	2,09	1,32	1,44	2,59	2,8	1,6	0,85
1-4	1-4	44,39	0	6,59	6,82	1,66	1,38	6,44	14,16	5,2	2,14
4-9	4-9	30,36	0	2,4	1,46	0,04	0,14	5,75	16,49	2,84	1,25
>9	>9	2,55	0	0,05	0,09	0	0,03	0,72	1,5	0,11	0,05
Sum	Sum	100	8,39	10,65	10,45	3,02	2,99	15,5	34,96	9,75	4,28







Data protection

Data in Clidata are protected by several tools:

- Oracle protection - data are stored in database therefore all oracle data security is available
  - Data are protected by Oracle against corruption
  - Automatic Backup and Export enhance the data protection
  - Definition of User – each person interacting with the system connect be it own user
  - Database Rights – user is assigned by rights to tables or views. Without the right the user is not able to work with particular table or view
  - Restrictions – allows to limit user only on the part of the data (Restricted in space and time)
    - READ – user is able to read only part of the data
    - WRITE/MODIFY – user is able to modify only part of the data

    Restriction can be for example as follows:
    - Restrict user only to READ years 1990 till 2000 (see all data modify years 1990-2000)
    - Restrict user only to READ/WRITE station O1PORU01 (user does not see and modify other data than data belonging to station O1PORU01)
- Application protection – definition of the acces to only parts of the Clidata application.
  - Application protection can be for example as follows:
    - User can do Key Entry but is not able to do Quality Control.
    - User can read Station Geography and view products but is not able to do Key entry
    - User can only view products
- Roles – Each Application and Oracle protection can be grouped into more complex ROLES. Typical roles identified in the system are:
  - Key Entry Person
  - Quality Control Person
  - Person Responsible for Products
  - Administrator

  Roles can be more or less specific depending on requirements.

Standard DATA Exchange Exports

Clidata is able to create standard WMO Messages of the following Formats as the standard products:

- CLIMAT
- CLIMAT TEMP
- BUFR SYNOP
- BUFR TEMP
- WORLD WEATHER RECORDS

The creation can be manual or automatic.