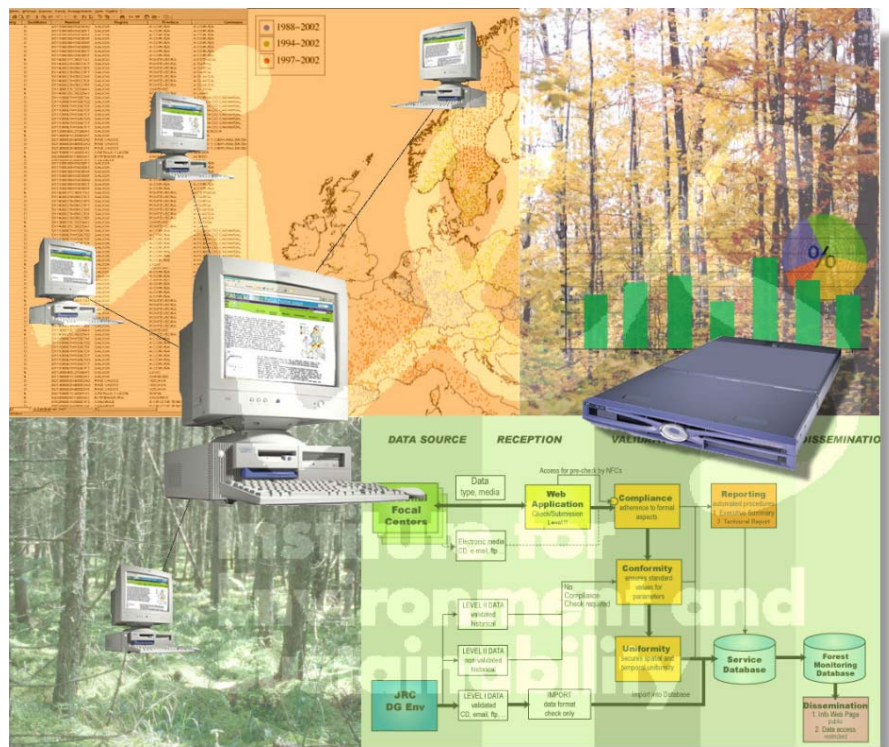# Forest Focus Monitoring Database System

VALIDATION METHODOLOGY 2008

Hiederer, R., T. Durrant, O. Granke, M. Lambotte, M. Lorenz, B. Mignon, V. Mues

MISSION OF THE INSTITUTE FOR ENVIRONMENT & SUSTAINABILITY

The mission of the Institute for Environment and Sustainability is to provide scientific-technical support to the European Union's Policies for the protection and sustainable development of the European and global environment.

European Commission Joint Research Centre
Institute for Environment and Sustainability (IES)
I-Ispra (VA)
Tel.: +39 0332 78 95 14
e-mail: forestfocus-data@jrc.it

This document was prepared under contract of

European Commission Joint Research Centre
Institute for Environment and Sustainability
via Fermi, 1
21020 Ispra (VA)
Italy

Under contract

Service Provision for
Forest Focus Data Management Action

Contract No.: 383547

I-MAGE Consult
rue de Gembloux 122
B-5002 St-Servais (Namur)
Belgium
http://www.i-mage.be/
Contact: Michel Lambotte, m.lambotte@i-mage.be

Nouvelles Solutions Informatiques s.a. (NSI)
Chaussée de Bruxelles, 174A
B-4340 Awans
Belgium
http://www.nsi-sa.be
Contact: Bertrand Mignon, B.Mignon@nsi-sa.be

Johann Heinrich von Thünen-Institute (vTI)
Federal Research Institute for Rural Areas, Forestry and Fisheries
Leuschnerstraße 91
D-21031 Hamburg
Germany
http://www.vti.bund.de/en/
Contact: Oliver Granke, oliver.granke@vti.bund.de

[1]European Commission Joint Research Centre
Institute for Environment and Sustainability
via Fermi, 1
21020 Ispra (VA)
Italy

[2]I-MAGE Consult
rue de Gembloux 122
B-5002 St-Servais (Namur)
Belgium

[3]Nouvelles Solutions Informatiques s.a. (NSI)
Chaussée de Bruxelles, 174A
B-4340 Awans
Belgium

[4]Johann Heinrich von Thünen-Institute (vTI)
Federal Research Institute for Rural Areas, Forestry and Fisheries
Leuschnerstraße 91
D-21031 Hamburg
Germany

Forest Focus Monitoring Database System

VALIDATION
METHODOLOGY 2008

Hiederer, R., T. Durrant, O. Granke, M. Lambotte,
M. Lorenz, B. Mignon, V. Mues

Table of Contents

List of Tables

List of Figures

List of Acronyms and Abbreviations

CODE	DESCRIPTION
ASCII	American Standard Code for Information Interchange
BFH	Federal Research Centre for Forestry and Forest Products Bundesanstalt für Forst- und Holzwirtschaft
BLOB	Binary large object
CLRTAP	Convention of the Long-Range Trans-boundary Air Pollution
dbh	Diameter at breast height
DAR	Data-Accompanying Report
DG AGRI	Agriculture Directorate General
DG ENV	Environment Directorate General
DSM	Data Submission Module
EC	European Commission
EMEP	Co-operative Programme for Monitoring and Evaluation of the Long Range Transmission of Air Pollutants in Europe
EU	European Union
FFMDb	Forest Focus Monitoring Database
FIMCI	Forest Intensive Monitoring Coordinating Institute
ICP Forests	International Cooperative Programme on Assessment and Monitoring of Air Pollution Effects on Forests
IES	Institute for Environment and Sustainability
JRC	European Commission Joint Research Centre
LM&NH	Land Management & Natural Hazards Unit
NFC	National Focal Centre
NSI	Nouvelles Solutions Informatiques s.a.
PCC	Programme Coordinating Centre
PDF	Portable Document Format
UN-ECE	United Nations Economic Commission for Europe
XML	Extended Mark-up Language

List of Survey Codes

Code	Survey Name
AQ	Air Quality
CC	Crown Condition
DP	Deposition
FO	Foliar Chemistry
GR	Growth and Yield
GV	Ground Vegetation
LF	Litterfall
MM	Meteorology
OZ	Ozone Injury
PH	Phenology
SI	System Instalment
SO	Soil Condition
SS	Soil Solution

1 GENERAL INFORMATION

1.1 Background

Forest Focus (Regulation (EC) No 2152/2003[1]) is a Community scheme for harmonised, broad-based, comprehensive and long-term monitoring of European forest ecosystems. It concentrates in particular on protecting forests against air pollution and fire. To supplement the monitoring system, Forest Focus stipulates the development of new instruments relating to soil monitoring, carbon sequestration, biodiversity, climate change and protective functions of forests.

Under this scheme the monitoring of air pollution effects on forests is carried out by participating countries on the basis of the systematic network of observation points (Level I) and of the network of observation plots for intensive and continuous monitoring (Level II). These monitoring activities under Forest Focus continue from the network and plots established and implemented under Council Regulation (EEC) No 3528/86[2] and Regulations (EEC) No 1696/87[3] and (EC) No 1091/94[4].

The monitoring programme of air pollution effects is linked to the *International Cooperative Programme on Assessment and Monitoring of Air Pollution Effects on Forest* (ICP Forests). ICP Forests reports to the working Group on Effects of the *Convention of the Long-Range Trans-boundary Air Pollution* (CLRTAP) of the *United Nations Economic Commission for Europe* (UN-ECE).

Forest Focus Article 15(1) stipulates that the Member States shall annually, through the designated authorities and agencies, forward to the Commission geo-referenced data gathered under the scheme, together with a report on them. For managing the data the European Commission Joint Research Centre (JRC) has implemented a Forest Focus Monitoring Database System. The system was developed and realized under contract by a Consortium, coordinated by I-MAGE Consult with Nouvelles Solutions Informatiques s.a. (NSI) as consortium partner and the Bundesforschungsanstalt für Forst- und Holzwirtschaft (BFH), now Johann Heinrich von Thünen-Institute (vTI), Federal Research Institute for Rural Areas, Forestry and Fisheries, as sub-contractor.

The designated authorities and agencies, the National Focal Centres, submitted annually to the JRC their observations made on Level II plots. Data are submitted via a Web-Module specifically designed for the task as part of the Forest Focus Monitoring Database System. The data are then validated in a process of three stages of checks of various aspects of the information submitted before entering the Forest Focus Monitoring Database (FFMDb).

---

[1] OJ L 324, 11.12.2003, p. 1-8
[2] OJ L 326, 21.11.1986, p. 2
[3] OJ L 161, 22.06.1987, p.1 - 22
[4] OJ L 125, 18.05.1994, p.1 - 44

1.2 Data Flow

An overview over the generic flow of data within the FFMDb System, referred to in subsequent chapters as the system, and the various stages of data processing is presented in form of a schematized standard data flow in Figure 1.
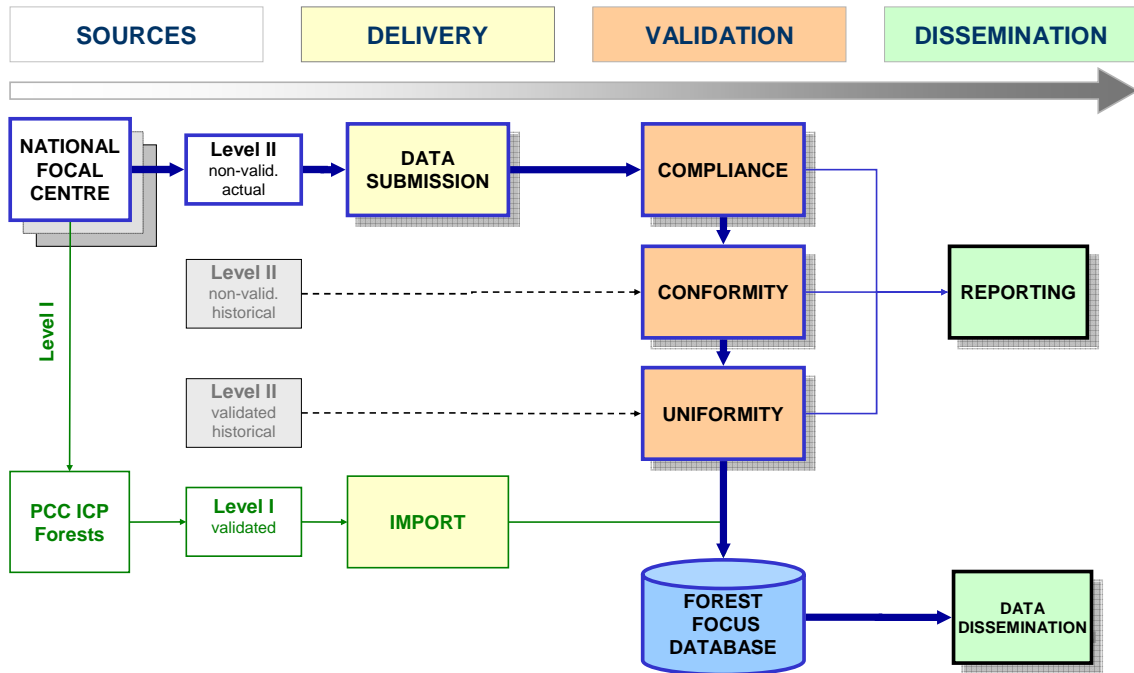
Figure 1: Schematized Standard Data Flow

Details on the various stages in the data flow are given in the sections hereafter.

1.2.1 Data Sources

Data are collected at the Level I (systematic) and Level II (intensive) monitoring plots by EU Member States and countries participating in the common monitoring scheme through bodies designated by the responsible national institutions. The data collected are forwarded by the designated authorities and agencies (National Focal Centres, NFCs) to the European Commission on an annual basis.

Data from Level II monitoring plots are provided by NFCs to the JRC and validated under the responsibility of the JRC. Data from Level I plots are managed and validated under the responsibility of the Programme Coordinating Centre (PCC) of ICP Forests. The validated data are provided by the PCC to the JRC once per year and are integrated into the system database. For both monitoring surveys only validated data enter the FFMDb.

1.2.2 Data Submission

Submitting data from monitoring surveys by the NFCs to the JRC is scheduled on an annual basis. Data for a given monitoring campaign should be submitted to the JRC by December of the year following the monitoring activity. However, some surveys are not performed annually and only submitted at more infrequent intervals.

1.2.3 Data Validation

Data validation concerns the check of the data submitted by NFCs to identify and exclude ambiguous, unlikely or incorrect values. Only fully validated data enter the FFMDb. The validation of Level II data is separated into three distinct stages. The first stage concerns the adherence of the data to the data format specifications stipulated in the *Technical Specifications* issues by the JRC for each monitoring year (Compliance Check). The check is performed on-line and a report on the results is generated when testing the data. The report allows NFCs to verify the adherence of the format of their data according to the specifications and to correct the data before submitting the forms.

Data that pass the Compliance Check are subjected to an evaluation of Conformity. Those tests concern the content of the data provided as opposed to the Compliance Check, which reported on formal aspects. The Conformity Check stage is followed by tests of data Uniformity. The tests are intended to establish the suitability of the data for further temporal and spatial analyses. Conformity and Uniformity Checks are performed off-line using the Service Database, because some of the tests require relatively intense processing and direct access to the FFMDb.

1.2.4 Reporting

The objective of the reporting task is to provide a comprehensive account on the data provided for a given monitoring year in form of standardized documents. The main documents produced are the *Data Submission Reports* and the *Technical Reports*. The latter are accompanied by *Executive Summary Reports*. Reports are prepared on an annual basis.

1.2.5 Dissemination

Level I and Level II data from all surveys and monitoring years can be made accessible to third parties for further analysis. Data can be disseminated by providing access to the FFMDb through a web-application for downloading the relevant parts of the database in form of an XML file. Access is restricted to authorized users, who can download part or all of the validated data.

2 PROCESS CONTROL

Data are processed by NFCs until they are submitted using the Data Submission Module (DSM). There are some principal differences in managing data before and after data submission. Before data are submitted they can be tested, deleted and re-loaded into an intermediate storage area as often as considered necessary by an NFC. Once submitted the data are no longer accessible to an NFC and cannot be modified or deleted. However, new versions can be submitted and take precedence over previous versions

2.1   Process Control before Data Submission

A graphical presentation of the process control for data submission is given in Figure 2.
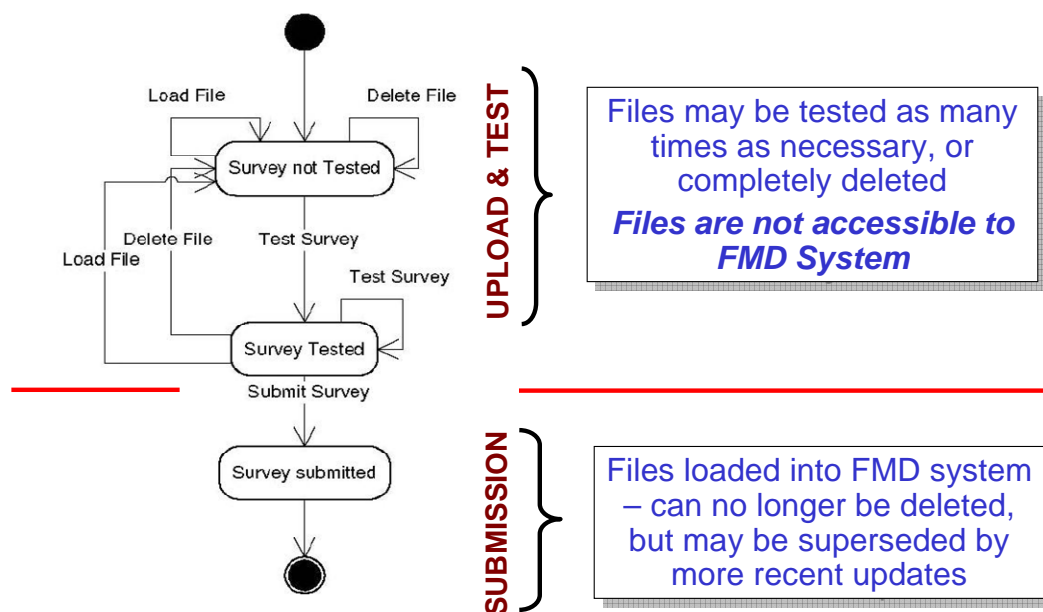
Figure 2: Process Control

For a given monitoring year the forms comprising a survey are selected and then uploaded into the intermediate storage area on the Web server. Once all forms comprising a survey are uploaded the survey is tested. Forms generating errors can be deleted, data corrected and reloaded by the NFC without any restriction. Once a survey is complete the data are tested for compliance. Testing a survey can be performed as required and the last results are stored in form of a report, which is available to the submitting NFC in PDF format. Once a survey has been tested it can be submitted. It

should be noted that a survey can be submitted containing warnings, but also errors. However, surveys containing errors cannot be processed.

2.2 Process Control after Data Submission

When a survey has been submitted, the files are passed on to a different storage location and are no longer available to the user for modifications. The user can still view the results of the Compliance Check and a submission summary, but the data from surveys submitted can no longer be deleted from the system. This data management policy has been adapted to allow generating a history of data submissions, which not only contains the dates of previous submissions, but also the data transferred.

In case a survey is submitted more than once the following rules apply:

1. Only one version of data will ever be processed and incorporated in the database.

2. When two survey types for the same year are submitted without errors, the more recent one will be processed. The NFC is encouraged to add an explanatory note to the files of the survey newly submitted.

3. For new submission made after the end of the submission period the new data can only be accepted and processed, if

   a. processing of a corresponding valid submission has not already been started or

   b. new data is requested due to inconsistencies in the format or value submitted, which were detected during subsequent processing of the data.

In all cases concerning data submission copies of the files are kept in the system for reasons of transparency.

Subsequent to the management of data in the data submission module a number of tasks are launched to transfer the values to the FFMDb for further processing:

- The files submitted via the JRC Internet server are copied to the system of the Service Provider. All submitted files, forms, Data DARs and other files must be loaded in the database. They will be kept in their original form as BLOB fields of the database, thus retaining the original file formats.

- The forms are loaded in corresponding database tables (staging area) for further processing. At the same time, the results from the Compliance Check performed during data submission are stored in the database in the same form as other test results. In this way, they will be available for reporting by querying the database.

- The data are tested for Conformity and Uniformity. Results from these tests are also recorded in the database.

- Some situations having generated a message can be marked as extreme events after confirmation by the NFC.

- Finally those data, which have passed the validation process, are transferred to the FFMDb.

2.3 Interpretation of Warnings and Errors

A sliding scale of warning and error messages was developed to label the results of the validation tests, because it is frequently not possible to identify without doubt that data are incorrect. The result of each validation test carries a message and associated severity code. The status "error" is only given when the code exceeds 50 and there is a clearly impossible situation. Some modification of the data will be required before further processing can take place. Warnings, however, simply draw attention to unusual events. In this case the NFC is asked to check each flagged value and either confirm its correctness or (if the value was erroneous after all) resubmit a corrected survey.

At the compliance stage, errors are fairly simple to detect and interpret. They are divided into three main types:

- Errors in the data submission procedure itself (missing mandatory form, not enough forms to complete the survey).

- Known "impossible" values within the files themselves, such as invalid dates, invalid characters and codes outside the given lists.

- Integrity checks within the survey to check that plots within the data file are also mentioned within the reduced plot file.

Warnings draw attention to missing optional forms (in case the NFC intended to submit the data but forgot), blank lines (in case this should have contained data) and comment lines (to confirm that the line should be there and is a genuine comment).

At this stage no consideration is given to the plausibility of a given value, only whether it fits the stated data formats.

At the conformity stage the actual data values are checked. As before, an error message confirms that something is wrong; however in this case it is not necessarily possible to ascertain precisely where the error lies. Most of these tests yield warning messages rather than errors as it becomes more difficult to detect values that are clearly erroneous.

Errors are divided into three main types according to the type of test applied:

- *Single parameter range tests* (e.g. values must be between 0 and 100 for percentage values).

- *Multiple parameter range tests* within a given survey (e.g. start date must be before end date).

- *Temporal consistency tests* (e.g. invariable parameters such as coordinates, altitude must not change).

Warnings are similarly divided. The single parameter range checks flag any data value that is outside an expected range for that parameter. Ranges were mostly derived from

the legacy data set and identify any value outside an approximate 95% level. Multiple parameter range checks note anomalous combinations of values, and the temporal consistency tests check for unusual increases/decreases in parameters (e.g. diameter values should increase over time, but not by more than a certain amount).

The validation system therefore identifies impossible values and also many unusual ones. However, there are limitations:

- The tests can detect an anomalous difference between two values but cannot compute which of them is erroneous.

- Submitted values that do not conform to the protocols (e.g. using different units) may not be detected unless the different units lead to data values outside the expected range. Similarly, elements submitted in the wrong order but within correct column widths will only generate errors if the normal ranges of the elements are different from each other.

- The range checks cannot pick up every implausible value. An average daily temperature of 30˚C in Spain in July will be flagged with a warning as an extreme event but 20˚C in Finland in January will not, because at present there are no seasonal/geographical constraints built into the system. To do so would introduce a significantly increased level of complexity into the tests; which may be out of proportion to the extra number of anomalous values actually detected.

The more complex the checks, the less clear-cut will be the results provided. The validation checks have to strike a balance between being too strict and thus incorrectly highlighting valid data or too broad to identify genuinely erroneous values.

3 DATA VALIDATION PROCESS

Data validation of data submitted by NFCs is the central task of data processing. Its purpose is to ensure that the information stored in the system can be used for an assessment of the state of a parameter sampled and in the evaluation of temporal and spatial trends between plots. It should also allow the integration of the data with other data sources in more extensive thematic analyses. This documentation details the validation of 2006 Level II data during the validation period of 2008.

Data are validated based on the principle that it is not possible to identify the correctness of data, but rather that it may be possible to identify the probability that data represent valid measurements. The degree of probability is expressed by grading data by severity codes from 0 to 100 using a sequential procedure, which assesses various characteristics and applies increasingly involved checks. The value attributed during validation represents a deviation from the expected value or range of values.

Severity codes below 50 generate warnings and are given in cases of non-standard situations, e.g. when an optional form is not submitted or when a line contains a comment. Warnings are reminders for the NFCs to re-examine their data and do not prevent the data from being further processed, once the values are confirmed by NFCs. For severity codes exceeding 50 the result of a test is an error. Surveys containing errors cannot be further processed or loaded into the database, and the NFC will have to submit new values.

3.1   Validation Stages

The validation of the data is achieved by subjecting the data to various test routines. The tests are arranged in three groups of checks:

   a) data format

   b) data content

   c) inter-comparability

The tests applied to verify data formats very much rely on the condition defined in the *Technical Specifications* for the arrangement of data within the survey forms. They assure that the individual characters of the ASCII file are assembled correctly into parameter values. The tests used to verify the actual values check for valid ranges of the values and consistency of the values recorded with previously submitted observations. The final group of tests compares the values from different plots.

The checks are applied in sequential order with increasing degree of complexity of the checks performed. A graphical overview of the validation tests is given in Figure 3.
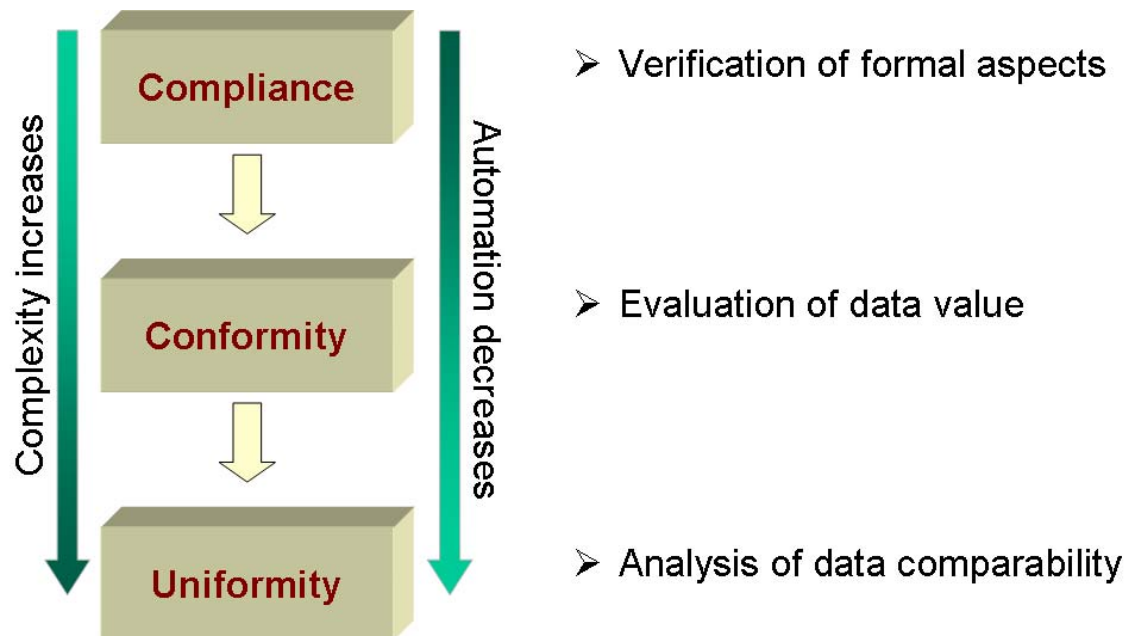
Figure 3: Sequential Arrangement of Data Validation Tests

The parameters used in the tests for Compliance and Conformity are aligned to several documents

- Technical Specifications as published annually by the JRC;

- the Manual of the International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests (ICP Forests) under the United Nations Economic Commission for Europe (UNECE);

- data verification routines developed by the former contractor of the European Commission, the Forest Intensive Monitoring Coordinating Institute (FIMCI), and ICP Forests;

- the experience in the verification and evaluation of data gained by the Programme Coordinating Centre (PCC) of ICP Forests at the Federal Research Centre for Forestry and Forest Products (BFH) in Hamburg, Germany, since 1985;

- measures on data quality assurance and evaluations laid down in annual Technical Reports on Forest Condition in Europe edited by PCC.

Conformity and Uniformity Checks are performed off-line using the Service Database, because some of the tests require relatively intense processing and direct access to the FFMDb. Details on the tests applied at the various stages of data validation are presented in the following section.

3.2 Compliance Check

The tests applied as part of the Compliance Check verify if the data in the submitted files of a survey comply with the specifications of the fixed formats ASCII files as stipulated in the JRC *Technical Specifications* documents. The documents are issued for each monitoring year. During compliance only syntactic checks are applied.

3.2.1 Survey Structure

The check of data Compliance is performed on the survey, the forms or files comprising a survey and the data within the survey. A general overview over the structure of a survey is given in Figure 4.
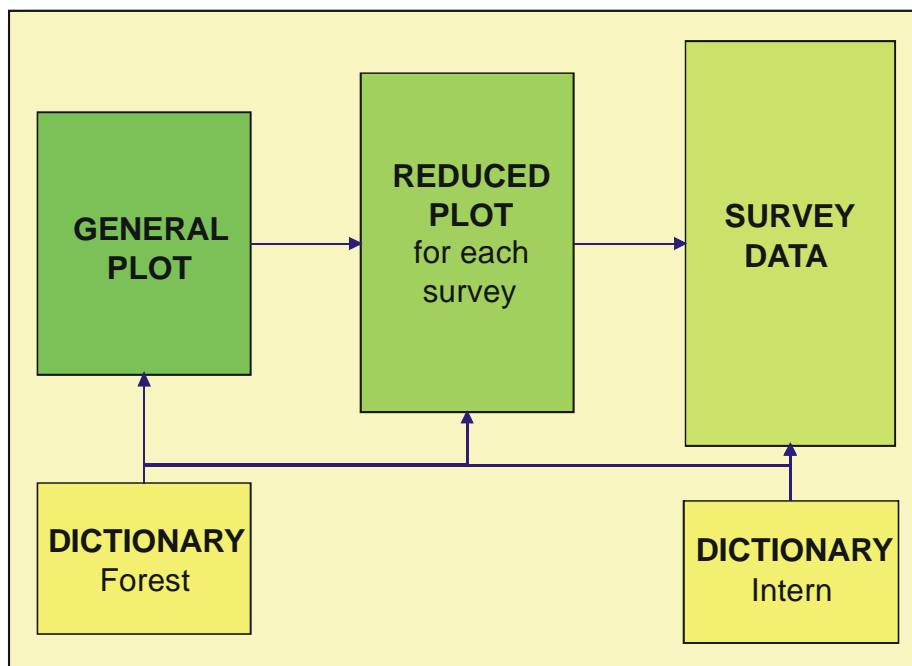
Figure 4 General Structure of Survey Forms

Each NFC compiles a file, which contains general information of the assemble of the Level II monitoring plot. This file is only provided when changes to the plot are made, e.g. when new plots are added. For each survey and monitoring year additional

information of the monitoring plots are given in a reduced plot file. The reduced plot file is linked to one or several files contain the data from the monitoring survey. Any coded information in the forms is linked to dictionary tables. The dictionary tables are either defined in the ICP Forests Manual or are FFMDb codes used internally.

The arrangements of the forms and the information stored contains several elements of data redundancy. This redundancy is one of the potential causes of data inconsistency af information stored within, but also between surveys.

In order to process the various survey types, some general information about the structure of the surveys has to be compiled and stored as part of the system. The survey type design is given in Figure 5.

The **Survey Type** table is used to store information about the different survey type:

- a two letter code for the survey type;

- survey name;

- survey frequency;

- two **validity years** for the survey;

- a flag indicating if the survey is a level II survey type. Only Level II surveys can be submitted by NFCs, but the FFMDb contains both Level I and Level II and both type can be disseminated.

The **Forms** table is used to store information about the various forms:

- the **survey type** of the form;

- a **3 letter code** (normally the file extension is used);

- the **Form Name;**

- a **comment;**

- a flag indicating whether the form is **mandatory or optional;**

- two **validity years** for the form;

- the **file extension** used for the form;

- a flag indicating if the form is a reduce plot file for the survey.

While designing the system, we noticed that the format of a form may change over time. For the Crown Condition survey, we have two versions of the PLT, TRM and TRM forms.

The **Forms Versions** table is used to stored information about the various versions of the forms:

- the form this version belongs to;

- a version code;

- a comment;

- the validity period in the form <Valid From>, <Valid To> year.

In the system design used the file extension used for a form must remain the same for all versions of the form.

Survey Type			
Survey type Code	<pi>	VA2	<M>
Survey type Name		VA50	<M>
Survey Frequency (years)		N2	<M>
Survey Frequency comment		VA30	
Is Level II		YN_VALUE	<M>
Valid From		YEAR	
Valid To		YEAR	
ST Processing order		N1	<M>
Survey ID <pi>			

Form Is part of survey

Forms			
Form Code	<pi>	VA3	<M>
Form Name		VA50	<M>
Comment		COMMENT	
Mandatory Optional		YN_VALUE	<M>
Valid From		YEAR	
Valid To		YEAR	
File extension		VA3	<M>
File Comment		VA100	
Processing order		N1	
Is reduced Plot File		YN_VALUE	
FORM_ID <pi>			

Form version to form

Forms Versions			
Form Version Code	<pi>	VA4	<M>
Comment		COMMENT	
Valid From		YEAR	
Valid To		YEAR	
FORM_ID <pi>			

Form Format to Form Version

Form Format Description			
Param Name	<pi>	COLUMN_NAME	<M>
Comment		COMMENT	
Start position		N3	<M>
End position		N3	
Format		VA10	
Param type		VA20	
Param length		N4	
Param precision		N4	
Is redondant		YN_VALUE	
Decode Function		VA50	<M>
Enforce Dict Constraint in Compliance		YN_VALUE	
Is Plot Param		YN_VALUE	
PK_ST_FORM_FORMAT <pi>			

Referenced dictionary

Dictionaries			
Code of the dictionary	<pi>	TABLE_NAME	<M>
Name of the dictionary		VA100	
Name of Id column		COLUMN_NAME	
Identifier_1 <pi>			

Figure 5: Survey Type Table

The general rule for accepting surveys is that at least 2 forms have to be submitted for a survey. An exception is the System survey, which only comprises the XXGENER.PLT form.

3.2.2   Submission File Format

Each survey to be submitted comprises at least two forms. Each form describes the content of a file. The file formats of all forms are based on the *Technical Specifications* issued by the JRC for each monitoring year from 2002 to 2006.

An example of the file format description is given in Figure 6.

Figure 6: File Format Description (example PLT Form of Crown Condition)

In the FFMDb the survey forms are stored in the Form Format table records:

- a parameter name;

- a description of the parameter;

- the start position of the parameter in the form;

- the end position of the parameter in the form;

- if the parameter is a coded value, the name of the dictionary where the list of values can be found;

- the type of the parameter (alpha-numeric, numeric, date, code);

- the parameter field length;

- the precision of the parameter if it is numeric.

Some supplementary technical attributes are also defined:

- A flag indicating if the verification that values are in the dictionary table must be enforced during compliance. This flag is normally always true.

- The decode function used to transfer the parameters to the staging area.

- The redundant flag is used for parameters that have to be duplicated for technical reason.

- The field <Is Plot Param> is used to indicate that a parameter is a plot number.

It should be noted that the length and precision of the parameter defined in the specifications can be different from the way it must be stored in the database. The format fro data submission can be more strictly defined than the ones used in the FFMDb which has to accommodate different version of a form between monitoring years and also includes legacy data from surveys preceding 2002.

The various fixed-format file specifications are summarized in tabular form in Section 1 of the *Technical Specifications* for each of the Level II surveys. The tables of Section 1 provide descriptions on:

- the parameter value start and end positions;

- description of the parameter to be reported;

- the field format (type, dimension, precision);

- in case of a coded field the name of the related dictionary.

Valid field codes are given in Section 2 of the *Technical Specifications* document.

3.2.3   Data Formats

Data are stored in the files in fixed ASCII format. The format is characterized by:

- fixed position for each character;

- fixed field dimension, sometimes fixes line length;

- no data type definition;

- no separator symbol;

- no character identifier symbol;

- requires field definition (dictionary).

The ASCII (American Standard Code for Information Interchange) defines a set of **control** and **character** codes in 7 bits. The code was first published by the American Standards Association (later ANSI) in 1963. The Standard was published in 1967, with amendments until ANSI X3.4-1986(R1997). Derivates of the standard are: ECMA-6, ISO/IEC 646, ITU-T, CCITT, DIN 66003, RCF 20, IBM version of ANSI X3.4-1986.

For Forest Focus data the following specifications apply:

- Use of 7-bit IBM PC / MS-DOS code page 437 (codes 0 – 127)

- 33 control characters, 0 – 31, 127

- 95 printable characters, 32 – 126

- End-of-line defined by [Carriage Return + Line Feed]

- Right-alignment in data fields

- Left-alignment in comment field

- + / - sign for co-ordinates

- Leading 0 for all value fields and co-ordinates and date values

- Point "." as decimal separator

- Space as field separator and empty positions

It should be noted that the extended 8-bit ASCII code (255 characters) is not used. As a consequence, non-US characters (e.g. accents)cannot be stored in the data files. Data lines have to end by a [CR] + [LF]. The code 26 [SUB] ("^Z") must not be used to indicate an end-default DOS/Windows environment. A graphical presentation of end-of-line characters is given in Figure 7.
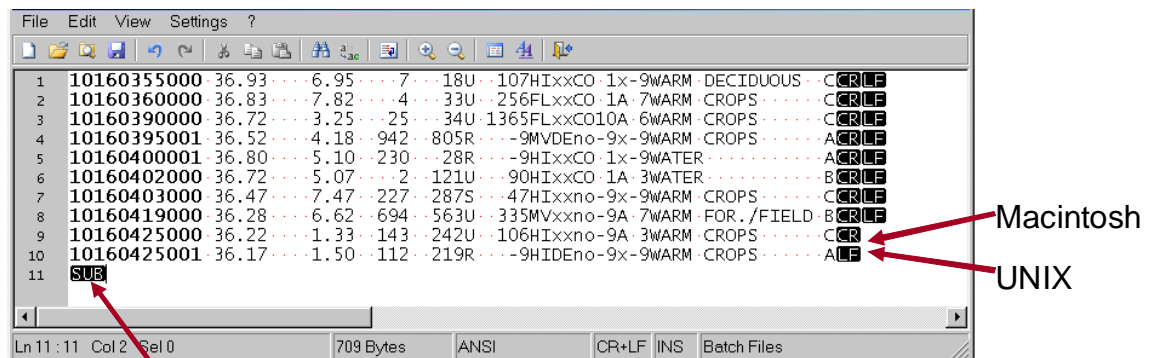
Figure 7: Use of End-of-Line Codes form Different Operating Systems or Processing Packages

The data in the ASCII files represent different field formats:

- Alpha-numeric:
  Text
  Codes

- Numeric:
  Integer
  Float

- Date

- Composites:
  Geographic co-ordinates

All fields are separated by one or several space characters (ASCII code 32).

3.2.4   Tests for Data Compliance

The tests for data Compliance verify the adherence of the data to the pre-defined positions in a survey form. Any deviation from the defined format will lead to a warning message and, in case of significant deviations, an error. Also validated by the Compliance Check is whether the symbolic values used for conditions are defined, e.g. the linked dictionary entries in case of categorical parameters (codes). If a file or data value fails a test applied for Compliance, i.e. an error condition could not be resolved, the survey cannot be further processed.

The tests cover the following aspects of the data forms:

- Completeness of forms (mandatory, optional)

- Reference for Plot No. in reduced plot file

- Comment lines

- No entries in no-data position

- Validity of date (entry and valid date)

- Validity of co-ordinate (entry and valid geographic unit)

- Validity of numeric type

- Reference to code values

- Length of field entry

Details of the tests applied for data Compliance is given in Table 1.

Table 1: Checks Applied for Data Compliance

CODE	MESSAGE	SEVERITY
MISSING_MAN_FORM	Some mandatory form is not present: %FORM_NAME%. The corresponding file should have this extension: %ENTENSION NAME	50
MISSING_OPT_FORM	WARNING: Some optional form is not present: %FORM_NAME%. The corresponding file should have this extension: %ENTENSION NAME	10
PLOT_NOT_IN_REDUCED_PLOT _FILE	The plot %PLOTNUMBER% is not in the reduced plot file	55
NO_VALUE_ALLOWED	There is a character: %CHAR% in a column that should not contain any data : %COLUMN_NUMBER%	60
CODE_NOT_IN_LIST	A coded parameter has a value %PARAM_VALUE% not in the list %DICTIONARY_NAME%	65
NOT_A_VALID_DATE	Parameter %PARAM_CODE% at position %START% to %END% : %PARAM_VALUE% is not a valid date. Format must be %FORMAT%	70
NOT_A_VALID_NUMBER	Parameter %PARAM_CODE% at position %START% to %END% : %PARAM_VALUE% is not a valid number.	75
VALUE TOO LONG*	Parameter %PARAM_CODE% at position %START% to %END% : %PARAM_VALUE% is not a valid number.	80
TOO_MUCH_DECIMAL*	Parameter %PARAM_CODE% at position %START% to %END% : %PARAM_VALUE% has too many decimals. Format must be %FORMAT%. The value will be interpreted as %ROUNDED_VALUE% in further processing	20
TOO_FEW_FORMS	Error, you must submit all forms, DARQ and other documents of a survey in one submission. Your submission contains only one form and a survey must contain at least two forms	90
INVALID_CHAR	Line contains invalid character	60
CODE_NOT_IN_LIST	A coded parameter has a value not in the corresponding dictionary	80
CODE_COUNTRY_NOT_CORRES PONDING	The country code does not correspond to the current country	80
NOT_A_VALID_COORDINATE	Not a valid coordinate	40
BLANK_LINE	Blank line	05
CMNT_LINE	Line was interpreted as a comment	05

* The VALUE_TOO_LONG and TOO_MUCH_DECIMAL errors should not occur following changes to the system in 2006, although the condition is still tested.

3.2.5 Interpretation of Field Formats

Over the time the interpretation of the filed formats had to undergo a process of adaptations. Originally the interpretation of the formats was exactly as given in the specifications. After the first submissions of data it became obvious that some field dimensions were insufficient to hold the data reported. The previously suggested procedure to deal with such cases, i.e. fill the field with the maximum value and report the actual measurement in the comment field, places the actually measured value outside the range of standard analysis tools. Correspondingly, measurements too small to be recorded in the dimension of the field were frequently rounded to 0 or to the smallest recordable value. Those practices carry the risk of generating spurious results when computing summary statistics for a parameter and can invalidate relationships between parameters.

Using a fixed-format to record the data does not allow enlarging the fields without having an effect on the position of all subsequent fields in the form. Changing the field dimensions would also have to be transferred to the ICP Manuals to remain consistent in the specifications. The process is rather lengthy and would not have helped to manage the situation already at hand. The solution applied was to apply a more tolerant interpretation of the field formats. The modifications concern the position of the decimal point in float fields and the definition of some integer fields to allow float values to be stored in the fields.

- **Floating Decimal Point**

  The interpretation of the format for numerical values has been changed in July 2006 to allow more flexibility. In the initial tests the position of the decimal within the format specified was fixed. For example, a format of 99.9 could only hold values between 0.1 and 99.9. For some parameters it was found that the formats specified did not allow storing the measured value for certain parameters. As a consequence of using a fixed-format file definition a change in one area would affect all subsequent field positions. This problem was avoided by not controlling the position of the decimal point. This interpretation increases the storage capacity of a field by several orders of magnitude, but provides less intrinsic control over the values submitted. The VALUE_TOO_LONG and TOO_MUCH_DECIMAL errors should not occur, although the condition is still tested.

- **Integer Field with Float Option**

  The rules for the interpretation of integer values are:

  a) Discrete units (any "No. of...") are tested as integer values.

  b) Numeric fields linked to a dictionary associated as integer values.

  c) All fields dimensioned as [99] remain integer values.

  d) All integer fields dimensioned >[99] are tested as float value, if not a) or b).

For most fields defining a measured or observed parameter, the position of the decimal separator is indicative. As a consequence a filed defined as [99.99] can contain up to 5 digits. The range of values stretches from 0.001 to 99999.

Should a value exceed the range of values set by the format specifier for a given field it is advised to verify the validity of the value before changing the specified position of the decimal separator. Values not conforming to the format specifications generally indicate a problem with the measurement units and only in rare cases the occurrence of an extreme event.

The interpretation of some integer fields as float was noticed also in the legacy data. When importing the legacy data the previous formats were maintained generally to 7 decimal places. No information was lost due to rounding or truncation during the transfer of the data to the FFMDb.

3.2.6   Reporting on Compliance Check Results

A summary report on the results obtained during the compliance test containing all useful comments about the checks performed and the outcome can be viewed from the Data Submission Module. Two reports area available in PDF format, the Data Submission report and the Compliance report. The two documents differ mainly in the level of detail provided. The Data Submission report provides information of the surveys and forms submitted by an NFC. Also included are any additional files submitted by an NFC, such as DARs or complementary documents. For each form the information given comprises:

- the processing date and

- the list of the processed files

- the number of lines per file

- the number of comment lines per file

- the number of data lines per file

- the number of valid lines per file

The general content of the Data Compliance summary Report is given in Figure 8.

Data Compliance Summary

NFC	SURVEY
NFC_NAME	Crown Condition

1. Test date: *15-Dec-2005 11:27*

2. Completeness:
Number of Plots expected: *15*
Number of Plots submitted: *15*

3. File tested:

Type	Name
Form	XX2002.PLT
Form	XX2002.TRM

4. Form inclusion remarks:

Warning, some optional form is not present, Crown Condition assessments optional parameters Level II, the corresponding file should be named XX2002.TRO or nn2002.TRO

Figure 8: General Information Presented in Data Compliance Summary Report

The Data Compliance Summary report shows the status of the survey forms in Section 5 of the report. An example is given in Figure 9.

5. Form compliance:

The survey has forms with errors that will prevent it to be processed by the JRC.

5.1 Crown Condition assessments Contents of file with the information on Plot level: XX2002.PLT

Number of lines in file : *20*
Number of data lines in file : *16*
Number of valid lines in file: *13*
Number of plots in file : *15*
General compliance of the form : *Bad*

Compliance results summary:

Parameter	COMPLIANCE	Number of errors	Number of warnings
Sequence number of plots (1 to 9999)	OK	0	0
Country Code (France = 01, Belgium = 02, etc.)	OK	0	0
Observation plot number (max. 9999)	OK	0	0
Date of assessment (DDMMYY)	Errors	3	0
Latitude in +DDMMSS (e.g.+505852)	OK	0	0
Longitude in (+or-)DDMMSS (e.g.+035531)	OK	0	0
Altitude (in 50 meter classes from 1 to 51)	OK	0	0
Exposure (1,2,3)	OK	0	0
Other observations (word)	OK	0	0

Figure 9: Example of Section 5 of Data Compliance Summary Report

The most important section of the reports concerns the final status of the survey forms. Only surveys with status "OK" can be advanced to the next processing stage. All other status comments require additional information to be provided by the NFC.

Details of the conditions leading to a warning or error during the check are lists in the Data Compliance report together with the values which caused a message to be generated. The number of messages is limited to avoid unwieldy file sizes. The structure of the results presented in the Data Compliance Report is given in Figure 10.

Line No. in original ASCII file

Parameter name and position from Technical Specifications

Value found and context of non-compliance

Compliance results details:

Line Nr	Parameter	P1	P2	Value	Message
1				!Seq Cou...	This line is interpreted as a comment
2				!number ...	This line is interpreted as a comment
3				! ...	This line is interpreted as a comment
8	Date of assessment (DDMMYY)	14	19	......	Parameter DATE_ASSESSMENT at position 14 to 19 : ...... is not a valid date. Format must be DDMMYY
18	Date of assessment (DDMMYY)	14	19	......	Parameter DATE_ASSESSMENT at position 14 to 19 : ...... is not a valid date. Format must be DDMMYY
19	Date of assessment (DDMMYY)	14	19	......	Parameter DATE_ASSESSMENT at position 14 to 19 : ...... is not a valid date. Format must be DDMMYY
20					Blank line

3

Details on three errors from previous table

Figure 10: Structure of Detailed Message List in Data Compliance Report

3.3   Conformity Checks

The Conformity Check comprises a number of tests that are applied after the submitted data have been subjected to the Compliance Check. The tests are not performed in the temporary storage area of the Web-server, but in the staging area of the database.

The principle of the Conformity Check is to evaluate the probability that a data value is an actual observation. The condition is evaluated with the aid of single parameter range tests, including test of boundaries for geographic coordinates. The tests can also detect impossible values, e.g. pH = 0. Data consistency is also tested via cross-checking for the continuity of static values, e.g. individual tree species, altitude, or logical continuity of the change of variable values, e.g. tree diameter according to temporal consistency. All these tests aim at assessing plot-specific conditions. Information from other plots is not taken into account at this stage.

By design the checking routines cannot detect unlikely values for a defined data range (approximately at the 95% level), which was mostly derived from the Level II legacy data validated by the Forest Intensive Monitoring Coordinating Institute (FIMCI) or from expert knowledge. It does not necessarily mean that a value generating a message is actually wrong. The NFCs are asked to pay attention to those values and state if the values are correct but outliers, or if the data need corrections and have to be re-submitted.

3.3.1   Single-Parameter Tests

The range tests are conducted by doing simple SELECT queries on the data. All values that do not fall within a specified range will be flagged with 'err' or 'warning', respectively. Because it is possible to vary these values the minimum and maximum parameters used during the checks are stored directly in the database. They are documented and reported together with the check results. When an NFC verifies the correctness of a value flagged during the range test this condition can be stored in the database by marking it as "extreme value".

To assess the position of plots within the administrative area of an NFC a simple test of minimum-maximum coordinates in latitude and longitude in forma of a box is performed. Examples of the boxes used are given Figure 11.
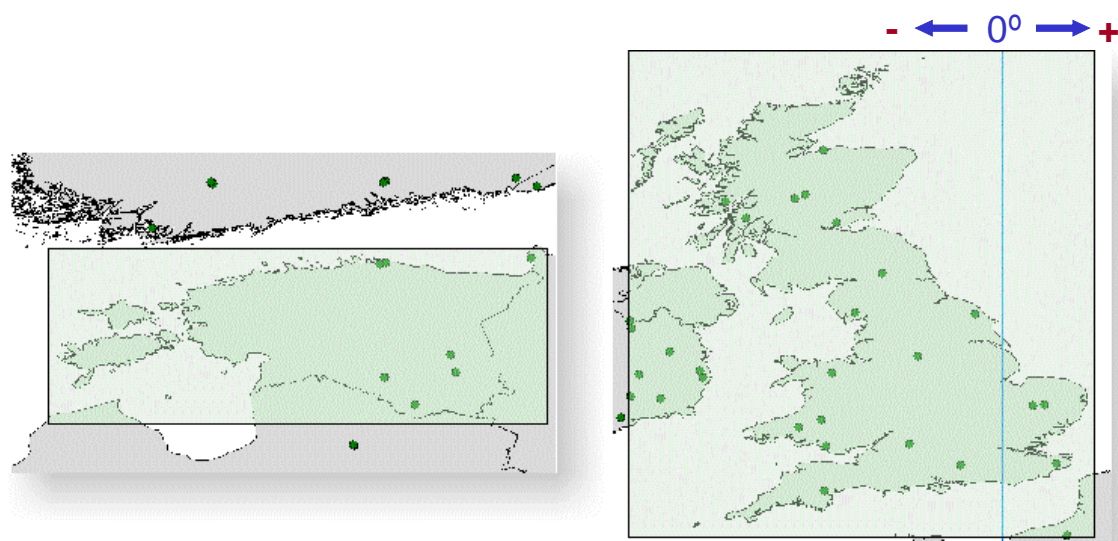
Figure 11: Assessment of Plot Position within Administrative Area of NFC

The tests can identify gross outliers or plots in the sea, but is limited in particular for NFCs, which have plots to the East and West of the Meridian.

A complete list of the single-parameter tests for Conformity is given in Table 2. The table columns contain the following information:

Column 1, number of proposed check:

> This number is used in order to allow for a short and concise reference to a particular check

Column 2, survey type:

> The survey type is one of the survey types described and regulated by the particular parts of the ICP Forests Manual (Crown Condition, Atmospheric Deposition, Phenology, etc.).

Column 3, table form:

> The name of the table or form, respectively, allows for a direct linkage of the check to be performed to a particular table in the database (e.g. DEM = atmospheric deposition, mandatory data). This table relies on the form for data submission as it is described by the ICP Forests manual and the R1sys document.

Column 4, field:

> The parameter to be checked, the name of the particular field in the table named in column 3.

Column 5, source of rule:

In this column the source of information is stored which the check is based on.

Column 6, severity:

The severity level of the test will indicate if a submission is possible but the warning should lead to a call for clarification to the respective NFC or if the value for the checked parameter clearly is out of any plausible range (error).

Column 7 and 8, minimum and maximum value:

The check routine which is proposed to be applied on the data of the particular field (column 4) of the particular table/form (column 3) (e.g. "0.3=<value=<10" for phosphorus content in the foliage, table FOM is proposed with check #64).

Table 2: Single-Parameter Range Tests for Conformity

#	Survey Type	Table Form	Field	Source of Rule	Severity	Min.	Max.
1	DP	DEM	Sample quantity	Derived from legacy data	warning	1	1845
2	DP	DEM	Alkalinity	Dep_analysis_def.asc	warning	-50	10000
3	DP	DEM	Conductivity	Dep_analysis_def.asc	warning	1	10000
4	DP	DEM	K	Dep_analysis_def.asc	warning	0.002	250
5	DP	DEM	Ca	Dep_analysis_def.asc	warning	0.001	275
6	DP	DEM	Mg	Dep_analysis_def.asc	warning	0.0025	100
7	DP	DEM	Na	Dep_analysis_def.asc, WG on QA/QC	Warning	0.003	500
8	DP	DEM	N-NH4	Dep_analysis_def.asc	Warning	0.002	175
9	DP	DEM	Cl	Dep_analysis_def.asc	warning	0.002	800
10	DP	DEM	N-NO3	Dep_analysis_def.asc	warning	0.002	175
11	DP	DEM	S-SO4	Dep_analysis_def.asc	warning	0.01	500
12	DP	DEM	pH	Dep_analysis_def.asc	warning	2.5	9.4
13	DP	DEM	N(total)	Dep_analysis_def.asc, WG on QA/QC	warning	0.03	350
14	DP	DEO	Sample quantity	Derived from legacy data	warning	1	1845
15	DP	DEO	Al^{3+}	Dep_analysis_def.asc	warning	1	8000
16	DP	DEO	Mn^{2+}	Dep_analysis_def.asc	warning	0.0001	15500
17	DP	DEO	Fe^{3+}	Dep_analysis_def.asc	warning	0.015	25000
18	DP	DEO	PO_4^{2-}	Dep_analysis_def.asc, WG on QA/QC	warning	0.0017	1000
19	DP	DEO	Cu	Dep_analysis_def.asc	warning	0.06	850
20	DP	DEO	Zn	Dep_analysis_def.asc	warning	0.005	4500
21	DP	DEO	Hg	Dep_analysis_def.asc	warning	0.02	100
22	DP	DEO	Pb	Dep_analysis_def.asc	warning	0.012	200
23	DP	DEO	Co	Dep_analysis_def.asc	warning	0.008	100
24	DP	DEO	Mo	Dep_analysis_def.asc	warning	0.008	100

#	Survey Type	Table Form	Field	Source of Rule	Severity	Min.	Max.
25	DP	DEO	S_{total}	Dep_analysis_def.asc, WG on QA/QC	warning	0.17	500
26	DP	DEO	N_{org}	Dep_analysis_def.asc, WG on QA/QC	warning	0.0003	100
27	DP	DEA	O_3	Dep_analysis_def.asc	warning	1	500
28	DP	DEA	SO_2	Dep_analysis_def.asc	warning	0.05	425
29	DP	DEA	SO_4	Dep_analysis_def.asc	warning	0.1	100
30	DP	DEA	NO_2	Dep_analysis_def.asc	warning	0.01	350
31	DP	DEA	NO	Dep_analysis_def.asc	warning	0.1	300
32	DP	DEA	HNO_2	Dep_analysis_def.asc	warning	0.001	30
33	DP	DEA	HNO_3	Dep_analysis_def.asc	warning	0.001	75
34	DP	DEA	NH_4NO_3	Dep_analysis_def.asc	warning	0.001	50
35	DP	DEA	NH_3	Dep_analysis_def.asc	warning	0.1	300
36	DP	DEA	Volatile Organic Compounds (V.O.C.)	Dep_analysis_def.asc	warning	0.001	100
37	MM	PLM	Vertical_Position	Chap7, Annex 6: Explanatory items (44)	warning	-2.5	40
38	MM	PLM	Instrument scanning interval	Derived from field wide	warning	0	999
39	MM	PLM	Instrument storing interval	Derived from field wide	warning	0	999
40	MM	MEM	Completeness	Expert knowledge	warning	1	100
41	MM	MEO	Completeness	Expert knowledge	warning	1	100
42	MM	MEM	value; Met_varaible = AT	Derived from legacy data	warning	-12.1	24.4
43	MM	MEM	value; Met_varaible = PR	Derived from legacy data	warning	0	36.2
44	MM	MEM	value; Met_varaible = RH	Derived from legacy data	warning	34	100
45	MM	MEM	value; Met_varaible = SR	Derived from legacy data	warning	0.3	362.1
46	MM	MEM	value; Met_varaible = ST	Derived from legacy data	warning	-1.3	20.3
47	MM	MEM	value; Met_varaible = WD	Derived from legacy data	error	0*	360
48	MM	MEM	value; Met_varaible = WS	Derived from legacy data	warning	0	6.9
49	MM	MEM	minimum; Met_varaible = AT	Derived from legacy data	warning	-17.2	18.7
50	MM	MEM	minimum; Met_varaible = RH	Derived from legacy data	warning	16	100
51	MM	MEM	minimum; Met_varaible = SR	Derived from legacy data	warning	0	206,6
52	MM	MEM	minimum; Met_varaible = ST	Derived from legacy data	warning	-1.7	18.5
53	MM	MEM	minimum; Met_varaible = WD	Derived from legacy data	error	0*	360
54	MM	MEM	minimum; Met_varaible = WS	Derived from legacy data	warning	0	5
55	MM	MEM	maximum; Met_varaible = AT	Derived from legacy data	warning	-8.1	31.8

#	Survey Type	Table Form	Field	Source of Rule	Severity	Min.	Max.
56	MM	MEM	maximum; Met_varaible = RH	Derived from legacy data	warning	49.8	100
57	MM	MEM	maximum; Met_varaible = SR	Derived from legacy data	warning	13.8	1139.0
58	MM	MEM	maximum; Met_varaible = ST	Derived from legacy data	warning	-1.1	24.2
59	MM	MEM	maximum; Met_varaible = WD	Derived from legacy data	error	0	360
60	MM	MEM	maximum; Met_varaible = WS	Derived from legacy data	warning	0	20.0
61	MM	MEO	value; Met_varaible = AP	Derived from legacy data	warning	989.7	1039.3
62	MM	MEO	value; Met_varaible = AT	Derived from legacy data	warning	To be submitted in MEM	
63	MM	MEO	value; Met_varaible = MP	Derived from legacy data	warning	-18.4	738.5
64	MM	MEO	value; Met_varaible = NR	Derived from legacy data	warning	-24.6	200.8
65	MM	MEO	value; Met_varaible = PR	Derived from legacy data	warning	To be submitted in MEM	
66	MM	MEO	value; Met_varaible = RH	Derived from legacy data	warning	To be submitted in MEM	
67	MM	MEO	value; Met_varaible = SD	Derived from legacy data	warning	0	32.5
68	MM	MEO	value; Met_varaible = SF	Derived from legacy data	warning	0	5.2
69	MM	MEO	value; Met_varaible = SR	Derived from legacy data	warning	To be submitted in MEM	
70	MM	MEO	value; Met_varaible = ST	Derived from legacy data	warning	-1.3	20.3
71	MM	MEO	value; Met_varaible = TF	Derived from legacy data	warning	0	91.7
72	MM	MEO	value; Met_varaible = UR	Derived from legacy data	warning	0	0.9
73	MM	MEO	value; Met_varaible = WC	Derived from legacy data	error	0	99
74	MM	MEO	value; Met_varaible = WS	Derived from legacy data	warning	To be submitted in MEM	
75	MM	MEO	minimum; Met_varaible = AP	Derived from legacy data	warning	983	1036
76	MM	MEO	minimum; Met_varaible = AT	Derived from legacy data	warning	To be submitted in MEM	
77	MM	MEO	minimum; Met_varaible = MP	Derived from legacy data	warning	-23.3	694
78	MM	MEO	minimum; Met_varaible = NR	Derived from legacy data	warning	-83.5	-5.2
79	MM	MEO	minimum; Met_varaible = RH	Derived from legacy data	warning	To be submitted in MEM	
80	MM	MEO	minimum; Met_varaible = ST	Derived from legacy data	warning	-1.7	18.5
81	MM	MEO	minimum; Met_varaible = WC	Derived from legacy data	error	0	98.5

#	Survey Type	Table Form	Field	Source of Rule	Severity	Min.	Max.
82	MM	MEO	maximum; Met_varaible = AP	Derived from legacy data	warning	995	1044
83	MM	MEO	maximum; Met_varaible = AT	Derived from legacy data	warning	To be submitted in MEM	
84	MM	MEO	maximum; Met_varaible = MP	Derived from legacy data	warning	-15.4	737
85	MM	MEO	maximum; Met_varaible = NR	Derived from legacy data	warning	19.6	694
86	MM	MEO	maximum; Met_varaible = RH	Derived from legacy data	warning	76.8	100
87	MM	MEO	maximum; Met_varaible = ST	Derived from legacy data	warning	-1.1	24.2
88	MM	MEO	maximum; Met_varaible = WC	Derived from legacy data	warning	3.6	100
89	MM	MEO	maximum; Met_varaible = WS	Derived from legacy data	warning	To be submitted in MEM	
90	GV	PLV	Total area	Expert knowledge	warning	1	9999
91	GV	PLV	Ground_Veg Cover	Expert knowledge	warning	1	100
92	GV	PLV	Shrub layer cover	Expert knowledge	warning	1	100
93	GV	PLV	Herb layer height	Expert knowledge	warning	0.01	2
94	GV	PLV	Mosses cover	Expert knowledge	warning	1	100
95	GV	PLV	Shrub_layer_height	Derived from legacy data	warning	0.01	10
96	GV	PLV	Herb_layer_height	Derived from legacy data	warning	0.01	2
97	FO	FOM	Mass_100_leaves	Derived from legacy data	warning	0.01	185
98	FO	FOM	Mass_1000_needles	Derived from legacy data	warning	0.01	211
99	FO	FOM	N	Foliar_analysis_def.asc	warning	5	45
100	FO	FOM	S	Foliar_analysis_def.asc	warning	0.2	10
101	FO	FOM	P	Foliar_analysis_def.asc	warning	0.3	10
102	FO	FOM	Ca	Foliar_analysis_def.asc	warning	0.5	40
103	FO	FOM	Mg	Foliar_analysis_def.asc	warning	0.1	10
104	FO	FOM	K	Foliar_analysis_def.asc	warning	0.75	35
105	FO	FOO	Zn	Foliar_analysis_def.asc	warning	4	500
106	FO	FOO	Mn	Foliar_analysis_def.asc	warning	5	11000
107	FO	FOO	Fe	Foliar_analysis_def.asc	warning	10	1500
108	FO	FOO	Cu	Foliar_analysis_def.asc	warning	0.01	100
109	FO	FOO	Pb	Foliar_analysis_def.asc	warning	0.15	100
110	FO	FOO	B	Foliar_analysis_def.asc	warning	2.5	150
111	FO	FOO	Cd	Foliar_analysis_def.asc	warning	0.01	150
112	AQ	PAC	Inlet_height	Expert knowledge	warning	2	50
113	SO	SOM, SOO	pH(CaCl$_2$)	Soil_analysis_def.asc	warning	2	8.2
114	SO	SOM, SOO	pH(H$_2$O)	Soil_analysis_def.asc, WG on QA/QC	warning	2.5	9
115	SS	SSM	pH	Soil_analysis_def.asc, WG on QA/QC	warning	2	9
116	SS	SSM	Electrical_conductivity	sso_analysis_def.asc	warning	1	10000
117	SS	SSM	K	sso_analysis_def.asc	warning	0.001	100

#	Survey Type	Table Form	Field	Source of Rule	Severity	Min.	Max.
118	SS	SSM	Ca	sso_analysis_def.asc	warning	0.005	500
119	SS	SSM	Mg	sso_analysis_def.asc	warning	0.01	175
120	SS	SSM	N-NO$_3$	sso_analysis_def.asc	warning	0.001	300
121	SS	SSM	S-SO$_4$	sso_analysis_def.asc	warning	0.01	1250
122	SS	SSM	Al	sso_analysis_def.asc	warning	0.001	500
123	SS	SSM	DOC	sso_analysis_def.asc	warning	0.001	600
124	SS	SSO	Water_content	sso_analysis_def.asc	warning	0.06	1
125	SS	SSM	Na	sso_analysis_def.asc	warning	0.001	200
126	SS	SSM	Al_labile	sso_analysis_def.asc	warning	0.001	300
127	SS	SSM	Fe	sso_analysis_def.asc	warning	0.0003	500
128	SS	SSM	Mn	sso_analysis_def.asc	warning	0.001	300
129	SS	SSM	P	sso_analysis_def.asc	warning	0.001	12.5
130	SS	SSM	N-NH$_4$	sso_analysis_def.asc	warning	0.0008	75
131	SS	SSM	Cl	sso_analysis_def.asc	warning	0.01	400
132	SS	SSM	Cr	sso_analysis_def.asc, WG on QA/QC	warning	0.04	100
133	SS	SSM	Ni	sso_analysis_def.asc	warning	0.001	1500
134 134	SS	SSM	Zn	sso_analysis_def.asc	warning	0.01	1000
135	SS	SSM	Cu	sso_analysis_def.asc	warning	0.05	1000
136	SS	SSM	Pb	sso_analysis_def.asc	warning	0.02	5000
137	SS	SSM	Si	sso_analysis_def.asc	warning	0.005	200
138	SS	SSM	Alkalinity	sso_analysis_def.asc	warning	0.1	27500
139	GR	PLI	Total_plot_size	Chap5.Annex 6, explanatory items (11)	warning	0.1	9.9999
140	GR	PLI	Sample_plot_size	Chap5.Annex 6, explanatory items (11)	warning	0.01	2
141	GR	IPM	Bark	Chap5.Annex 6, explanatory items (28)	warning	0.1	9.9
142	GR	IPM	Height	Chap5.Annex 6, explanatory items (29)	warning	1	50
143	GR	IPM	Tree_volume	Chap5.Annex 6, explanatory items (30)	warnig	0.001	15
144	GR	IPM	Crownlength	Chap5.Annex 6, explanatory items (31)	warning	0.1	40
145	GR	IPM	Crownwidth	Chap5.Annex 6, explanatory items (32)	warning	0.1	30
146	GR	INV	Stemwood volume remaining	Expert knowledge	warning	1	999
147	GR	INV	Stemwood volume Dead	Expert knowledge	warning	1	999
148	GR	INV	Stemwood volume removed	Expert knowledge	warning	1	999
149	LF	LFM	Mass of 100 leaves or 1000 needles	Derived from Foliage survey	warning	9	211
150	LF	LFM	Area of 100 leaves or 1000 needles	Derived from field wide	warning	1	9999
151	LF	LFO	Na	Expert knowledge	warning	0.005	99
152	LF	LFO	Zn	Derived from Foliage survey	warning	0.004	0.5
153	LF	LFO	Mn	Derived from Foliage survey	warning	0.005	11

#	Survey Type	Table Form	Field	Source of Rule	Severity	Min.	Max.
154	LF	LFO	Fe	Derived from Foliage survey	warning	0.01	1.5
155	LF	LFO	Al	Expert knowledge	warning	0.001	1.0
156	SO	SOM	Bulk density	Derived from field wide	warning	1	9999
157	SO	SOO	Bulk density	Derived from field wide	warning	1	9999
158	SO	SOM	Carbonates	Derived from field wide	warning	1	999
159	SO	SOO	Carbonates	Derived from field wide	warning	1	999
160	AQ	AQM	variable_c / O3_MC	Derived from stored data	warning	6.72	113.5
161	AQ	AQM	variable_c / O3_MCd	Derived from stored data	warning	8.65	67.68
162	AQ	AQM	variable_c / O3_MCn	Derived from stored data	warning	8.24	66.29
163	AQ	AQM	variable_c / O3_Max	Derived from stored data	warning	10.22	78.8
164	FO	FOO	C	Derived from stored data	warning	42	58

* "0" and "360" indicating "North"

3.3.2   Multiple Parameter Tests

Some tests check the consistency of a parameter with values of other parameters or fields reported. In some cases these rules imply specific conditions for the application of the check. For example, Check # 138 has to be applied only on those values submitted for mineral layers of the horizons M01, M12, M24, or M48. Other checks are related to parameters in the same table as the field that is checked (e.g. Check # 155) or in other tables (e.g. Check # 137). All the multiple parameter checks are performed using "SELECT …. WHERE …" queries. These checks, which are performed on more than one table, include a JOIN statement. The tests applied involving multiple parameters are summarized in Table 3.

Table 3: Multi-Parameter Tests for Conformity

#	Survey Type	Table Form	Field	Source of Rule	Severity	Rule
165	PH	PHE	Symptom	chap8, Annex2, Form11b(PHE)	error	only if event = 4 or 5
166	PH	PHE	Cause	chap8, Annex2, Form11b(PHE)	error	only if event = 4 or 5
167	PH	PHE	Scientific_name_cause	chap8, Annex2, Form11b(PHE)	error	only if event = 4 or 5
170	GV	PLV	Total_sampled_area	chap8, Annex2, Form10a	warning	value =< area of plot (general_plot information)
171	SO	SOM & SOO	Particle_size_clay	implicit (%)	error	0=<value=<100 (layers M01, M12, M24, M48)
172	SO	SOM & SOO	Particle_size_sand	implicit (%)	error	0=<value=<100 (layers M01, M12, M24, M48)
173	SO	SOM & SOO	Particle_size_silt	implicit (%)	error	0=<value=<100 (layers M01, M12, M24, M48)
174	SO	SOM & SOO	Coarse_fragments	implicit (%)	error	0=<value=<100 (layer M01 mandatory, M12, M24, M48 optional)
175	SO	SOM & SOO	Organic_layer_weight	soil_analysis_def.asc	warning	0.1=<value=<150 (OFH and HFH layers only)
176	SO	SOM & SOO	Organic_carbon	soil_analysis_def.asc	warning	layer O: 10=<value=<800, layer M: 0.4=<value=<500
177	SO	SOM & SOO	Total_N	soil_analysis_def.asc	warning	layer O: 0.5=<value=<35, layer M: 0.01=<value=<20
178	SO	SOM & SOO	Carbonates	soil_analysis_def.asc	warning	layer O: 0=<value=<50, layer layer M: 0=<value=<1000
179	SO	SOM & SOO	Exchangeable_acidity	soil_analysis_def.asc	warning	0.05=<value=<150 (both layers), mandatory
180	SO	SOM & SOO	Total_Al	soil_analysis_def.asc	warning	layer O: 100=<value=<50000, layer M: 100=<value=<100000, optional M01/M12/M24/M48
181	SO	SOM & SOO	Total_Ca	soil_analysis_def.asc	warning	layer O:150=<value=<150000, layer M: 20=<value=<350000 optional M01/M12/M24/M48
182	SO	SOM & SOO	Total_Fe	soil_analysis_def.asc	warning	layer O:1=<value=<50000, layer M: 1=<value=<150000 optional M01/M12/M24/M48

#	Survey Type	Table Form	Field	Source of Rule	Severity	Rule
183	SO	SOM & SOO	Total_K	soil_analysis_def.asc	warning	layer O:50=<value=<10000, layer M: 50=<value=<40000 optional M01/M12/M24/M48
184	SO	SOM & SOO	Total_Mg	soil_analysis_def.asc	warning	layer O:100=<value=<150000, layer M: 15=<value=<100000 optional M01/M12/M24/M48
185	SO	SOM & SOO	Total_Mn	soil_analysis_def.asc	warning	layer M: 1=<value=<3500 optional M01/M12/M24/M48
186	SO	SOM & SOO	Total_Na	soil_analysis_def.asc	warning	layer O:1=<value=<5000, layer M: 1=<value=<10000 optional M01/M12/M24/M48
187	GR	IRA	Diameter*t	Chap5.Annex 6, explanatory items (33)	warning	$Diameter*t < DBH_t$
188	GR	IRA	Diameter*t-5	Chap5.Annex 6, explanatory items (33)	error	Diameter*t-5 < Diameter*t
189	GR	IRA	Diameter*t-10	Chap5.Annex 6, explanatory items (33)	error	Diameter*t-10 < Diameter*t-5
190	GR	IRA	Diameter*t-15	Chap5.Annex 6, explanatory items (33)	error	Diameter*t-15 < Diameter*t-10
191	GR	IRA	Diameter*t-20	Chap5.Annex 6, explanatory items (33)	error	Diameter*t-20 < Diameter*t-15
192	GR	IRA	Diameter*t-25	Chap5.Annex 6, explanatory items (33)	error	Diameter*t-25 < Diameter*t-20
193	GR	IRA	Diameter*t-30	Chap5.Annex 6, explanatory items (33)	error	Diameter*t-30 < Diameter*t-25
194	GR	IRA	Diameter*t-35	Chap5.Annex 6, explanatory items (33)	error	Diameter*t-35 < Diameter*t-30
195	GR	IRA	Diameter*t-40	Chap5.Annex 6, explanatory items (33)	error	Diameter*t-40 < Diameter*t-35
196	GR	IRA	Diameter*t-45	Chap5.Annex 6, explanatory items (33)	error	Diameter*t-45 < Diameter*t-40
197	GR	IRA	Diameter*t-50	Chap5.Annex 6, explanatory items (33)	error	Diameter*t-50 < Diameter*t-45
198	GR	IEV	Basal_area_plot$_{t-5}$	Chap5.Annex 6, explanatory items (34)	warning	Basal_area_plot$_{t-5}$ < Basal_area_plot$_t$
199	GR	IEV	Volume/plot$_{t-5}$	Chap5.Annex 6, explanatory items (35)	warning	Volume/plot$_{t-5}$ < Volume/plot$_t$

#	Survey Type	Table Form	Field	Source of Rule	Severity	Rule
200	GR	IEV	Basal_area_plot$_{t-10}$	Chap5.Annex 6, explanatory items (34)	warning	Basal_area_plot$_{t-10}$ < Basal_area_plot$_{t-5}$
201	GR	IEV	Volume/plot$_{t-10}$	Chap5.Annex 6, explanatory items (35)	warning	Volume/plot$_{t-10}$ < Volume/plot$_{t-5}$
202	GR	IEV	Basal_area_plot$_{t-15}$	Chap5.Annex 6, explanatory items (34)	warning	Basal_area_plot$_{t-15}$ < Basal_area_plot$_{t-10}$
203	GR	IEV	Volume/plot$_{t-15}$	Chap5.Annex 6, explanatory items (35)	warning	Volume/plot$_{t-15}$ < Volume/plot$_{t-10}$
204	GR	IEV	Basal_area_plot$_{t-20}$	Chap5.Annex 6, explanatory items (34)	warning	Basal_area_plot$_{t-20}$ < Basal_area_plot$_{t-15}$
205	GR	IEV	Volume/plot$_{t-20}$	Chap5.Annex 6, explanatory items (35)	warning	Volume/plot$_{t-20}$ < Volume/plot$_{t-15}$
206	GR	IEV	Basal_area_plot$_{t-25}$	Chap5.Annex 6, explanatory items (34)	warning	Basal_area_plot$_{t-25}$ < Basal_area_plot$_{t-20}$
207	GR	IEV	Volume/plot$_{t-25}$	Chap5.Annex 6, explanatory items (35)	warning	Volume/plot$_{t-25}$ < Volume/plot$_{t-20}$
208	PH	PHE	Tree species	Chap 11, Annex2, Form (PHE)	error	Tree species must exist in PLP form
209	PH	PHE	Estimated_starting_date	implicitly in variables	error	before_date < estimated_starting_date< estimated_completion_d ate<date_after
210	PH	PHE	Estimated_completion_d ate	implicitly in variables	error	before_date < estimated_starting_date< estimated_completion_d ate<date_after
211	PH	PHI	Starting_date	Chap8, Annex 3: Explanatory items (3)	error	starting_date < completion_date
212	PH	PHI	Tree number	Chap 11, Annex2, Form (PHI)	error	Tree number must exist in PLP form
213	PH	PHI	Completion_date	Chap8, Annex 3: Explanatory items (3)	error	starting_date < completion_date
214	AQ	AQM	Sampler_Nr.	Chap10, Annex3, Form (AQM)	error	Passive_Sampler nr must exist in the passive sampler PPS form
215	AQ	AQM	Sampler_Nr.	Chap10, Annex3, Form (AQM)	error	Active_Sampler nr must exist in the passive sampler PAC form
216	LF	LFP	Active_sampling_period _from	implicitly in variables	error	active_sampling_period_ from<active_sampling_p eriod_till
217	LF	LFP	Active_sampling_period _till	implicitly in variables	error	active_sampling_period_ from<active_sampling_p eriod_till
218	LF	LFM	Collection_period_from	implicitly in variables	error	collection_period_from < collection_period_till
219	LF	LFM	Collection_period_till	implicitly in variables	error	collection_period_from < collection_period_till
220	LF	LFO	Collection_period_from	implicitly in variables	error	collection_period_from < collection_period_till

#	Survey Type	Table Form	Field	Source of Rule	Severity	Rule
221	LF	LFO	Collection_period_till	implicitly in variables	error	collection_period_from < collection_period_till

3.3.3 Temporal Consistency Tests

Temporal consistency is checked by comparing the values of the monitoring year with values which were submitted for the same parameter and plot in former years. The temporal consistency checks aim at assessing the continuity of those parameters which should not change over time, like the site co-ordinates. Any deviation from the previously validated values will result in an 'error'. For values that can vary over time, but which are expected to change in a certain direction or by a particular amount, a 'warning' is given. An example for this type of parameters is growth values. The tests applied to verify temporal consistency are given in Table 4.

Table 4: Temporal Consistency Tests for Conformity

#	Survey Type	Table Form	Field	Source of Rule	Severity	Rule
222	GEN.PLT	GENER.PLT.	Latitude, longitude	Expert knowledge	error	Coordinates must not change
223	CC	PLT				
224	SO	PLS				
225	SS	PSS				
226	FO	PLF				
227	GR	PLI				
228	DP	PLD				
229	MM	PLM				
230	GV	PLV				
231	AQ	PAC, PLL				
232	LF	LFP				
233	GEN.PLT	GENER.PLT	Latitude, longitude	Expert knowledge	error	Coordinates must be in country; latitude and longitude are within range of coordinates of the respective country
234	CC	PLT				
235	SO	PLS				
236	SS	PSS				
237	FO	PLF				
238	GR	PLI				
239	DP	PLD				
240	MM	PLM				
241	GV	PLV				
242	AQ	PAC, PPS				
243	OZ	PLL				
244	LF	LFP				
245	GEN.PLT	GENER.PLT	Altitude	Expert knowledge	error	Altitude must not change
246	CC	PLT				
247	SO	PLS				
248	SS	PSS				
249	FO	PLF				
250	DP	PLD				
251	MM	PLM				
252	GV	PLV				
253	AQ	PAC, PPS				

#	Survey Type	Table Form	Field	Source of Rule	Severity	Rule
254	OZ	PLL				
255	LF	LFP				
256	SO	PLS	Soil type / soil unit codes	Expert knowledge	error	Soil type must not change (except change from FAO to new classification)
257	GEN.PLT	GENER.PLT	Mean age of dominant storey	Expert knowledge	warning	Stand age must be (a certain number of years) higher than at previous observation (same class or higher as before may be used, if lower class submitted trees of main storage must have been harvested);
258	CC FO PH GR	TRM, TRO FOM PLP, PHI IPM	Tree number / sequence number of trees; *number of first to fifth tree in sample*; tree number; *tree number*	Expert knowledge	error	sequence tree numbers must be the same as at previous observation (unless new trees were intentiously involved - e.g. due to ingrowth - or existing trees were felled or wind-thrown)
259	CC FO PH GR	TRM FOM, FOO PLP, PHI IPM	Species; *tree species*; tree species code; *species*	Expert knowledge	error	Tree species must not change for particular tree
267	SO	SOM	Particle size clay	Expert knowledge	warning	Particle size (clay) must not change
268	SO	SOM	Particle size sand	Expert knowledge	warning	Particle size (sand) must not change
269	SO	SOM	Particle size silt	Expert knowledge	warning	Particle size (silt) must not change
270	SO	SOM	Texture class	Expert knowledge	warning	texture class must not change
271	GR	IPM	Diameter1, Diameter2	Expert knowledge	warning	diameter at breast height must not be smaller than at previous observation
272	GR	IPM	Diameter1, Diameter2	Expert knowledge	warning	relative increase in diameter at breast height must be smaller or equal 15%
273	GR	IPM	Height	Expert knowledge	warning	height must be equal or higher than (height of last observation - 1m); tolerance of 1 m as depending on tree species and age measurement accuracy is variable

#	Survey Type	Table Form	Field	Source of Rule	Severity	Rule
274	GR	IPM	Height	Expert knowledge	warning	relative increase in height must be higher than - 5%; (decrease of 5% allowed due to possible inaccuracies)
275	GR	INV	Stemwood volume per plot (remaining, newly dead, and removed, respectively)	Expert knowledge	warning	stemwood volume (remaining + dead wood + removed) must be lower than 115% of remaining stemwood at previous observation
276	GR	IEV	Volume per plot	Expert knowledge	warning	volume per plot = +/- 15% of this value at previous observation; higher deviation must be accompanied by respective comments / data (=> wind throw, snow brake, insects)

3.3.4   Treatment of Missing Measurement Values

The representation of missing data should be addressed by the Expert Panels and specific guidelines should be adopted and included in the ICP Manual. In the absence of such guidelines the JRC has developed specific rules for treating zero values in data submitted by NFCs for monitoring periods from 2002 onwards.

- **Classification of Missing Data**

   For the purpose of the data validation procedure, missing data are entries recorded in the data files in the reporting forms, which do not represent valid measurements or observations for a given parameter. Missing data can occur due to a given parameter not collected, not usable or lost. The validation process is not concerned with missing data, which are not recorded in the data files, e.g. the completeness of periodic measurements. Furthermore, issues of randomly or systematically missing data are not treated.

   The ICP Forests Manual mentions the coding of "missing data" in several places, for example for the data recorded in the forms SOM, SOO, SSM, SSO, FOO, DEM, DEO, DEA, LFM, LFO. The ICP Forests Manual identifies two cases of data being measured / observed, but at levels which cannot be represented in the field formats. Depending on the condition, recording the data in the forms is treated differently. A valid measured value may be either too small or too large to fit the field format. Both conditions frequently occur for several parameters.

- **Recommendations**

  The general approach to treating "missing data" in the validation process of the Forest Focus Monitoring Database has to take the properties of the legacy data into account as well as the variety of treatment of "missing data" by NFCs. The validation process is therefore based on the identification of valid values for measured or observed parameters. In this the approach differs profoundly from the identification of codes signifying missing data.

  The recommendations presented are given below, separated by the situations to which they apply:

  a. *Measured, but outside field specifications*

  - **Value too small for format specified for field**

    A measurement of a value should be recorded as measured, shifting the decimal point as needed. Data should not be rounded except where shifting the decimal point is still insufficient to record the measured value. For example, the format for recording $N-NO_3$ in the Soil Solution survey specified as 999.9. A measured value of 0.03 should be recorded as such. In the example given rounding should only be applied for values <0.001.

  - **Value too large for field format**

    A measurement of a value should be recorded as measured without the decimal part. For example, alkalinity in the soil solution at times exceeded 999.9μmolc/l. A value of 1500 should be recorded as such in the field. Data should not be entered into the field "Other observations".

  b. *Measured, but below limits of detection for instrument*

      The use of -1 for a measurement is defined to code a value below the detection limits of the instrument used. This condition occurs frequently in soil solution data. The values should not be rounded, interpolated or marked by a zero entry.

  c. *No Measurement*

      The field should be left empty. The condition should **not** be coded by using a zero entry, although this is sometimes recommended.

Cases *a*. and *b*. have been largely eliminated. The decimal point in the format is no longer tied to a fixed position. A format specified as 999.9 can hold values from 0.001 to 99999. It would have been preferable to adjust the field dimension in the format specifications. However, the process of modifying the specifications is lengthy and would not solve actual problems.

3.3.5   Reporting on Consistency Check Results

The results of the tests are compiled as a lists of flagged values, which indicate either an error for values indicating potentially unusual conditions or a warning for values outside a pre-set range. All flagged values are listed and described with an explanatory legend in a report. The report consists of two documents:

- **Table TR1B**

    The table is a text document which presents a summary for each survey and form the number of parameters tested, the number of tests resulting in a message and the status of the forms at the end of the Conformity Check. A survey can only be accepted if all forms belonging to a survey have the status OK.

- **Table TR3**

    The table contains a detailed list of all messages raised with an explanatory note and the result of the severity coding. To aid the identification of the record causing a warning the line number in the submitted files are also indicated.

An example of the results reported in the TR3 table is given in Figure 12.

NFC : Name	Survey : Deposition	Year : 2005	
Validation tests status : Finished	Submission Date : 29-DEC-06		
Form :	Line Sequ. Nr.:	Message :	Severity :
.DEM	8	K 0 NOT IN RANGE .002 TO 250 FOR PLOT 102	warning
.DEM	10	K 0 NOT IN RANGE .002 TO 250 FOR PLOT 101	warning
.DEM	12	K 0 NOT IN RANGE .002 TO 250 FOR PLOT 102	warning
.DEM	17	QUANTITY 0 NOT IN RANGE 1 TO 1845 FOR PLOT 101	warning
.DEM	18	QUANTITY 0 NOT IN RANGE 1 TO 1845 FOR PLOT 101	warning
.DEM	25	QUANTITY 0 NOT IN RANGE 1 TO 1845 FOR PLOT 101	warning
.DEM	26	QUANTITY 0 NOT IN RANGE 1 TO 1845 FOR PLOT 101	warning
.DEM	29	QUANTITY 0 NOT IN RANGE 1 TO 1845 FOR PLOT 101	warning
.DEM	30	QUANTITY 0 NOT IN RANGE 1 TO 1845 FOR PLOT 101	warning
.DEM	33	QUANTITY 0 NOT IN RANGE 1 TO 1845 FOR PLOT 101	warning
.DEM	34	QUANTITY 0 NOT IN RANGE 1 TO 1845 FOR PLOT 101	warning
.DEM	35	N_NH4 0 NOT IN RANGE .002 TO 175 FOR PLOT 102	warning
.DEM	36	N_NH4 0 NOT IN RANGE .002 TO 175 FOR PLOT 102	warning
.DEM	37	QUANTITY 0 NOT IN RANGE 1 TO 1845 FOR PLOT 101	warning
.DEM	38	QUANTITY 0 NOT IN RANGE 1 TO 1845 FOR PLOT 101	warning
.DEM	39	QUANTITY 0 NOT IN RANGE 1 TO 1845 FOR PLOT 102	warning
.DEM	40	QUANTITY 0 NOT IN RANGE 1 TO 1845 FOR PLOT 102	warning
.DEM	51	N_NH4 0 NOT IN RANGE .002 TO 175 FOR PLOT 102	warning

Figure 12: Example of Results Reported in TR3 Table

If no answer was provided by the NFC before the deadline and/or errors are still identified, data cannot be fully validated and the complete survey cannot be loaded into the FFMDb.

In response to the reports NFCs have the opportunity to react in several ways:

- Where extreme values are confirmed by the NFCs, corresponding registry lines will be flagged as extreme event and the data is carried forward;

- In case of errors, the NFC has to correct the errors and re-submit the whole survey through the data submission module. The data then have to pass back into

the workflow and pass through the complete validation process (Compliance, Conformity and Uniformity) again.

3.4 Uniformity Checks

The tests applied to check data Uniformity are intended to identify temporal and spatial data inconsistencies which could not be found during any of the previous checks. The Uniformity Check consists of an interpretation of temporal and spatial development of parameters using data from all plots. Contrary to Conformity data Uniformity is verified by comparative tests using more than the information from a single plot. Uniformity tests are more qualitative and require the interpretation of the results by an expert in the field. The interpretation includes a comparison with external data as far as such information is available in a suitable form.

The check includes an automatic procedure for generating maps for various key parameters monitored. In general, the map depicts the status of a given parameter for the monitoring year. Where appropriate a status map is supplemented by a map showing changes over a previous monitoring year. While the compilation of the maps is relatively straightforward for continuous surveys the process is less apparent for surveys with longer monitoring intervals, such as Growth or Soil Condition. The main obstacle for non-annual surveys and data collected for comparing conditions at one plot with those from other plots or analysing changes over time is the lack of data for any given monitoring year. This is most extreme for Soil Condition with a repeat cycle of 10 years. On average one would expect data for 10% of all plots for a monitoring year, which is largely insufficient for a comparative analysis. Therefore, the tests for data of non-annual surveys use data from one or several previous surveys, which are not from an immediately preceding year.

3.4.1 Crown Condition

Maps C1 – C6

Mean plot defoliation is mapped for data from each monitoring year for the following main tree species found in Europe:

- *Pinus sylvestris*
- *Picea abies*
- *Fagus sylvatica*
- *Quercus robur* and *Q. petrea*
- *Quercus ilex* and *Q. rotundifolia*
- *Pinus pinaster*

Only plots are depicted on the map with at least 3 trees of the respective tree species assessed in the reporting year. For each plot, defoliation is classified into 6 classes (0-10%, 11-20%, 21-30%, 31-40%, 41-50%, 51-100% mean defoliation).

Forest Focus Monitoring Database System
Validation Methodology 2008

Expert interpretation of these maps focuses on possible systematic differences between the assessment methods of different countries. Differences between the countries are interpreted in the context of existing research results published in previous EU/ICP Forests reports.

3.4.2   Soil Condition

Map S1

Uniformity of the Soil Condition survey concentrates on the data reported for pH (CaCl$_2$) for the upper mineral layer per plot. The pH values are taken from the M01, alternatively from the M05, the M51, or from the M02 layer in this order. The map shows pH values for the latest available year for each plot, but specifically indicates plots for which data were submitted in the reporting year. The pH is classified into the classes 2.1-3; 3.1-4; 4.1-5; 5.1-6; 6.1-7; >7.

Existing maps from the soil survey carried out on Level I plots show lowest pH-values (around 3) on plots in central Europe. The majority of plots show pH-values between 3 and 4.5 with many of them located in Scandinavia. Most plots with high pH-values (around 6) are located in the Mediterranean. The interpretation of the measurements reported for Level II plots concentrates on significant deviations from this general pattern. For reporting the emphasis is made on a description of the spatial pattern which is shown in the map.

3.4.3   Soil Solution

Maps SS1 – SS3

For identifying the uniformity of the Soil Solution data changes in the values reported for previous monitoring years are assessed for three soil solution compounds (S-SO$_4$, N-NO$_3$ and N-NH$_4$). The difference between the time-weighted mean concentration in the reporting year and the average of the weighted mean concentration of the five preceding years is evaluated as part of the tests. It is necessary to weight the mean concentrations with the length of the period in which it was measured. This period can be calculated by dividing the period length by the number of periods in the respective observation in the reduced plot file for soil solution (form PSS). The differences are grouped into five equidistant classes. The comparison is based on the topsoil (0 > sampling depth >= -0.3m). Only those plots are taken into account for which data from an over all period of at least 300 days per year are available.

Alternatively to the difference the quotient of the actual year value and the mean of the five preceding years can be mapped in order to detect those values which represent relatively high changes.

Not all Soil Solution data stored in the FFMDb are necessarily mapped. For plots displayed on the map the following conditions apply:

- the sample has to be taken from the mineral soil layer;

- the layer depth must be at least 30 cm;

- the total sample period must be more than 300 days.

The analysis focuses on sudden changes in concentrations of the compounds assessed in the soil solution as depicted on the corresponding maps. Within the interpretation, precipitation of the respective year has to be taken into account as a major additional influence on the concentrations.

3.4.4   Foliage

Maps F1 – F12

The foliar element concentrations constitute important response parameters for air pollution effects. Their spatial variation can give hints on the completeness and correctness of measurements in the participating countries. Plots with concentrations clearly deviating from surrounding plots are in the focus of the interpreting expert. Concentrations of nitrogen and sulphur are mapped plot-wise for tree species coded in field „sample number":

F1: Foliar Nitrogen Concentrations for *Pinus sylvestris*

F2: Foliar Sulphur Concentrations for *Pinus sylvestris*

F3: Foliar Nitrogen Concentrations for *Picea abies*

F4: Foliar Sulphur Concentrations for *Picea abies*

F5: Foliar Nitrogen Concentrations for *Fagus sylvatica*

F6: Foliar Sulphur Concentrations for *Fagus sylvatica*

F7: Foliar Nitrogen Concentrations for *Quercus robur* and *Qu. petrae*a

F8: Foliar Sulphur Concentrations for *Quercus robur* and *Qu. petraea*

F9: Foliar Nitrogen Concentrations for *Quercus ilex* and *Qu. rotundifolia*

F10: Foliar Sulphur Concentrations for *Quercus ilex* and *Qu. rotundifolia*

F11: Foliar Nitrogen Concentrations for *Pinus pinaster*

F12: Foliar Sulphur Concentrations for *Pinus pinaster.*

For each reporting year, mean plot concentrations are firstly calculated species and plot-wise and are then classified into five classes of same relative frequency (pentiles). The minimum of the first class is the minimum of the depicted values, the maximum of fifth class is the maximum of depicted values.

For reporting the interpretation will focus on a description of regional differences among the observed concentrations. Any found spatial patterns can be linked to respective findings in the deposition data.

3.4.5   Growth and Yield

Maps FG1 and FG2

To assess the uniformity of tree dimensions and forest growth the mean basal area per plot is used. The temporal consistency is validated by using the mean annual increment of basal area per plot, which is calculated from repeated measurements.

- *Map FG1: Mean basal area* [m²] is mapped based on the most recent data for each plot (submitted with form IEV, first group of "basal area per plot" and "volume per plot"). Mean basal area is classified into five classes with 20% of relative frequency each (pentiles, with: minimum of first class = minimum of values, maximum of fifth class = maximum of values). The map for mean basal area shows, when appropriate, the data of the latest available year for each plot, but specifically indicates plots with data submission in the reporting year.

- *Map FG2: Mean basal area increment* [m²] is mapped per plot and year, based on the most recent (five years) measurement period. For each plot, mean annual basal area increment is classified into five classes with 20% of relative frequency each, as is mean basal area. Mapped is the mean annual increment of the latest available (five years) period for each plot with available data, but specifically indicates plots with data submission in the reporting year.

Data can be mapped for the following parameters:

- mean basal area per plot, based on increment information (IEV);

- 5-year mean basal area increment per plot, based on increment information (IEV);

- calculated basal area, based on periodic data (IPM);

- 5-year calculated basal area increment, based on periodic data (IPM).

Forest growth is further validated by an index comparable to basal area calculated from the values of diameter (at breast height, dbh) parameter as reported in the IPM form. Contrary to the mean basal area taken from the IEV form the derived index comprises a unitless value independent of the size of the plot. The calculation of the index first sums up the tree specific area from the dbh values, using the mean diameter of the two values given in the form:

$$BA = \frac{\sum dbh^2 \times \frac{\pi}{4}}{sample \quad plot \quad size}$$

The mean for the plot is then obtained by dividing the dbh area sum by the sample plot size. A restriction for this calculation is that either

- the number of trees in this calculation (number of observation in the IPM file for this plot and year) is equal to the number of trees on the plot which is submitted

in the form PLI (plot file for growth) AND the sample plot size is equal to the total plot size (both submitted with PLI) OR

- the number of trees in this calculation divided by total number of trees (PLI) is +/- equal to the quotient of sample plot size (PLI) and total plot size (PLI).

Restriction (1):

number of observations (IPM) per plot and year ≈ number of trees in total plot (PLI) AND sample plot size (PLI) ≈ total plot size (PLI); in both comparisons the deviation should be not more than 10% of the lower values in the equation.

Restriction (2):

number of observations (IPM) / number of trees in total plot (PLI) ≈ sample plot size (PLI) / total plot size (PLI); the deviation should be not more than 10% of the lower value in the equation.

In case that the number of trees, the scale of the values or any other basic parameter deviates between two subsequent data submissions for a particular plot the division by the corresponding (constant) sample size will lead to a high change in basal area, which will allow for a more detailed check of the respective data. As in case of the mean basal area the calculated basal area index is mapped for data of the monitoring year and as an increment for the increment over the most recent measurement period.

3.4.6 Deposition

Maps DEP1 - DEP3

Validating Uniformity for data of the Deposition survey is based on contrasting the values reported for S-SO$_4$, N-NO$_3$ and N-NH$_4$ in two series of maps. The first series shows the plot-wise quantity weighted (volume of sampled precipitation) mean concentration of bulk deposition for S-SO$_4$, N-NO$_3$ and N-NH$_4$ in mg/l for the particular reporting year. The value is calculated as:

$$Quantity-weighted \quad mean \quad concentration_{dep} = \frac{\sum concentration \times quantity_{dep}}{\sum quantity_{dep}}$$

The calculations of quantity weighted mean concentration is necessary, because various instances of periodic measurements are submitted for a particular year. The calculations are only applied to data of plots for which data were submitted for at least 300 days (plot specific sum of period lengths in the PLD form). The resulting mean concentrations are grouped into 5 classes with 20% of relative frequency (pentiles, minimum of first class = minimum of values, maximum of fifth class = maximum of values). Extreme values in relation to values of surrounding plots are in the focus of the validating expert.

Maps DEP4 – DEP7

Within the interpretation, precipitation of the respective year has to be taken into account as a major additional influence on the concentrations. The purpose of this

second series of maps is intended to reveal sudden changes in concentrations of the depositions related to the amount of water (quantity of precipitation) in the bulk deposition.

The difference between the quantity weighted mean concentration in the reporting year (maps DEP1 – DEP3) and the average of the weighted mean concentrations of five preceding years is presented for the reporting year. The differences are grouped into five equidistant classes; minimum of 1st class is {-1*[max(-1*min;max)]}, maximum of 5th class is [max(-1*min;max)]. The analysis focuses on the description of observed spatial patterns of high / low deposition and will compare the monitored deposition levels with those for external data (if available) and former years. The calculations will be done only on those plots for which data were submitted for at least 300 days (plot specific sum of period lengths in the PLD form).

The reporting focuses on the description of observed spatial patterns of high / low deposition and will compare the monitored deposition levels with those for external data (if available) and former years.

3.4.7   Meteorology

Maps M1 and M2

Of all environmental factors, temperature and precipitation probably have the largest influence on forest condition. Map M1 shows the total annual precipitation (mm) and map M2 the mean annual temperature (°C) to validate data uniformity. For display purposes the data are grouped into 5 pentiles with 20% of relative frequency. Data were plotted in the map under the following conditions:

- Sum of precipitation and mean daily air temperature had to be measured for at least 300 days (continuity during year);

- Precipitation and air temperature measurements of at least 90% per day (continuity during day).

The validation evaluates the occurrence of values significantly different in relation to values of surrounding plots. The same continuity criteria as for table M1 are applied for the selection of plots to be mapped.

Only those plots have are included which have values for sum of precipitation and mean daily air temperature for at least 300 days (continuity during year) with a completeness of measurements of at least 90% per day (continuity during day). The condition for plotting can be expressed as:

$$[Completeness >= 90\%] AND [Days\_with\_values >= 300]$$

The spatial patterns of both variables will be examined. As the values of both variables are highly influenced by the topography at the measurement station the interpretation will have to take this into account.

External Meteorological Information

The Global Precipitation Centre (GPCC) offers an online overview on mean precipitation across Europe. Maps of mean precipitation for Level II plots can be compared to these products. It has however to be taken into account, that plot specific measurements are subjected to site and stand characteristics and will mostly deviate from model values for larger grids.
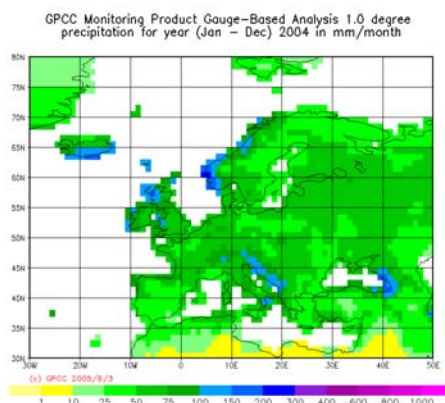

Figure 13: Mean monthly precipitation in 2004 in mm/month. From: GPCC

For reporting purposes found spatial patterns in the FMD data will be compared with those of the GPCC data. In case of significant deviations for single plots or groups of plots an explanation for this deviation will be searched for by examining the plot specific site and stand data.

Respective data on mean air temperature can be obtained from the US national climatic data centre at (http://www.ncdc.noaa.gov/oa/climate/climateresources.html) and is used to compare with the submitted data as described for precipitation above.

3.4.8   Ground Vegetation

Map GV1 and GV2

Ground Vegetation data are only sampled every 3 years. Consequently, the number of plots with data for 2002 is relatively low compared to other annual surveys. Data from the Ground Vegetation survey is shown on two maps.

- Map GV1 shows the plant species richness as the number of reported species over all layers (tree, shrubs, herbs and mosses) and surveys per plot in a specific reporting year. If a particular species code is submitted more than once per plot and year it is included only once. Resulting numbers are grouped and mapped using the following classes:

    <20, 20 to 40, 41 to 60, 61 to 80, >80 species.

- Map GV2 represents changes in species richness per plot compared to the most recent previous survey. Results are grouped into the following classes:

<-10, -10 to -3, -2 to +2, +3 to +10, > +10 species.

The classification of the groups allows distinguishing between plots and regions in which an increase of species numbers was observed and those where the number of species decreased.

The comparison between the number of species per plot in the reporting year with that observed in previous years should not yield extreme differences. Any changes in number or species composition of ground vegetation may indicate natural disturbances or management effects as well as errors in data submission. Extreme changes need to be followed by the validating expert.

3.4.9   Air Quality

Map AQ1

This map shows the time-weighted average concentration of O_3 per plot in a specific reporting year for all plots for which data were submitted from at least 180 days. Here, extreme values in relation to values of surrounding plots are in the focus of the validating expert, however taking into account the general increase in O_3 concentrations with decreasing geographical latitude. Concentrations are grouped into the following classes:

<30, 30-45, 46-60, >60 ppb.

In the interpretation of the result specific attention is given to extreme values in relation to values of surrounding plots, taking into account the general increase in O_3 concentrations with decreasing geographical latitude. Comparing plot data with external data could assist the analysis of the data.

External Air Quality Information

The Co-operative Programme for *Monitoring and Evaluation of the Long Range Transmission of Air Pollutants in Europe* (EMEP)[5] uses the unified EMEP model[6] for the modelling of deposition in 50*50 km grids across Europe. On the „data" section of the EMEP homepage[7] of main pollutants are published. The model results for sulphur and nitrogen concentrations can be compared to the precipitation weighted mean concentrations in bulk deposition from the Level II plots. It has however to be taken into account, that plot specific measurements are subjected to site and stand characteristics and will mostly deviate from model values for 50*50 km grids.

Whereas the aim of the validation procedure is to find questionable values submitted for single plots or countries for the reporting the validated FFMDb values will be compared
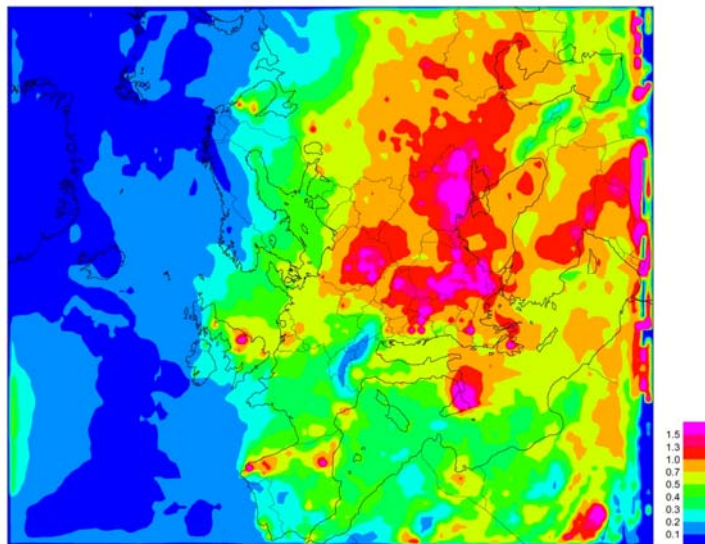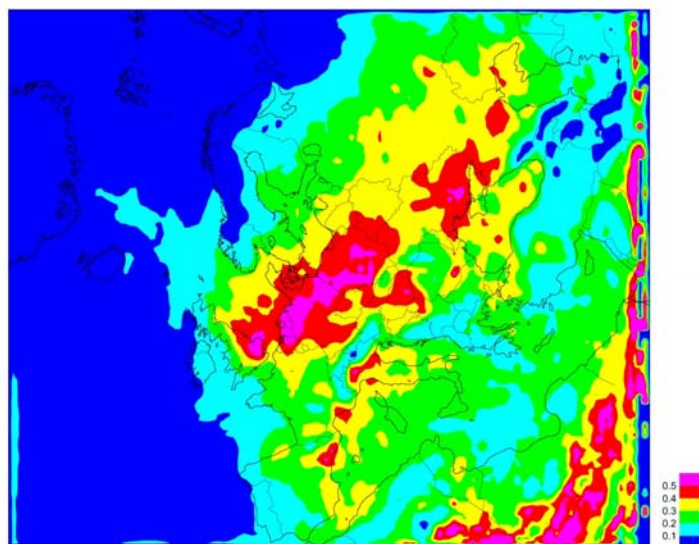
---

[5] http://www.emep.int/
[6] http://www.emep.int/UniDoc/index.html
[7] http://www.emep.int/Model_data/model_data.html

with the EMEP model values. It has to be taken into account, however, that the EMEP models cannot cover the site and stand specific situation at the FFMDb plots.

Figure 14: Comparison of model results for S in precipitation (mgS/l), 2002. from: EMEP Status Report 1/04

Figure 15: Comparison of model results for oxidized nitrogen in precipitation (mgN/l), 2002. from: EMEP Status Report 1/04

3.4.10   Visible Ozone-induced Injury

Table VOI2

Data from the survey of Visible Ozone Induced Injury are validated by means of a table rather than by a map. A map is not expected to show spatial patterns of injury because of the selective nature of positioning plots and because of the influence of local topographic conditions. In fact, the results given in the table confirm that a map would not have shown any spatial patterns. However, time series of observations should be established for identical plots in order to detect potential changes in visible ozone induced injury.

Table VOI2 shows the number of plots for which data for a particular tree species were submitted on "percentage of symptomatic leaves/needles" with form LTF and the number of plots on which this particular tree species showed injury ("percentage of symptomatic leaves/needles" with code 2 or 3 which means a share of more than 5% of the leaves/needles are affected by ozone damage).

Table 5: Number of plots with visible ozone injuries above 5% of the sampled leaves/needles at the most frequent tree species in the reporting year

Main Tree Species Prone to Ozone Injury	Total No. of Plots with Ozone Injury Assessment	No. of Plots with Ozone Injury Reported
Alnus glutinosa	x	z
Fagus sylvatica	x	z
Fraxinus excelsior	x	z
Picea abies	x	z
Pinus sylvestris	x	z
Quercus robur	x	Z

External Ozone Data

The Co-operative Programme for *Monitoring and Evaluation of the Long Range Transmission of Air Pollutants in Europe* (EMEP) has set up a so-called EMEP photo-oxidant model. The maps for the AOT40 (for forest) in ppb.h. can be compared to the AOT 40 per plot as evaluated in the Forest Focus data base. It has however to be taken into account, that plot specific measurements are subjected to site and stand characteristics and will mostly deviate from model values for larger grids.

For validation purposes the EMEP data on AOT40 can be used as a kind of reference. Regions in which high values of ozone concentrations can be expected can be identified and compared with the FMD values.

Concerning the reporting the concentrations of the FFMDb data will be interpreted with the background of the values from the EMEP models.
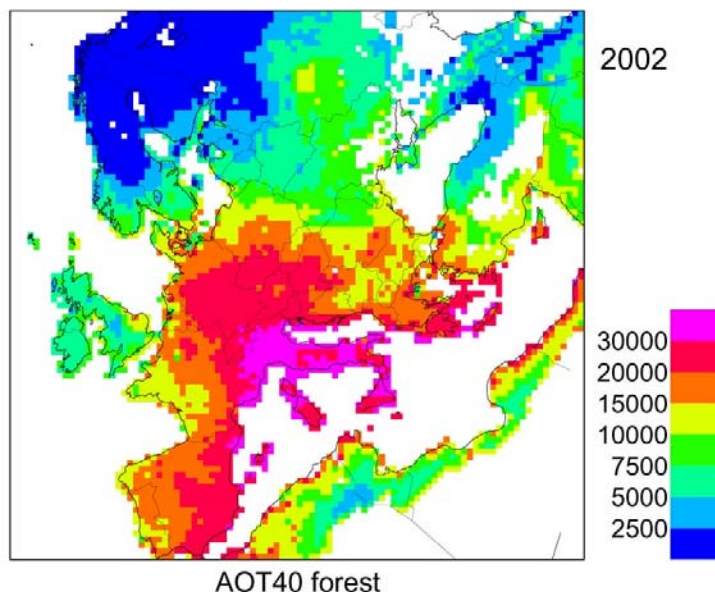
Figure 16: AOT40f (for forest) in ppb.h. 2002. from: EMEP Status Report 1/04

3.4.11   Phenology

Maps PH1 and PH2

Data from the Phenology survey are checked for uniformity by mapping the dates reported for the time of flushing (Event Code 1) and the dates reported for needle / leaf fall (Event Code 3). The dates are mapped when data for 50 or more plots are available.

3.4.12   Litterfall

Maps LT1, LT2 and LT3

For Litterfall the parameters of the dry weight (kg/m^2), the mean content of C (mg/g) and N (mg/N) are used, as reported in the LFM form. The dates are mapped when data for 50 or more plots are available.

4 VALIDATION LIMITS

Although the validation process is quite comprehensive and the tests are fairly complex the data stored in the FFMDb and made available for dissemination cannot necessarily be declared correct. According to the principle of the checks data are not tested for being correct, but for the probability that a value is outside of what could be expected as admissible. The limits of range tests are in most cases taken from the Level II legacy data and expert knowledge. For a given parameter the ranges are set globally and are not specific for countries or bio-geographic regions. This geographically unspecific method is low on maintenance overhead and straight forward to implement, but results in a higher probability of the oversight of outliers in countries with intermediate conditions. Whenever a parameter is similar in the range of observations to another parameter, e.g. for chemical elements, entering the parameter in the wrong column or even reporting the wrong parameter will also not be detected by the tests.

When data are recorded correctly in the forms there may still be differences in measurement methods between NFCs or laboratories. When differences in measurement methods lead to variations in the data reported those methods should be stored together with the data. This option is rarely available in the forms and the information is easily lost. In the absence of recording meta-data it is recommended to make use of the option of the system to include in the submission at least a document stating the methods and instruments used for collecting data at the plots as part of the DAR.

The option of allowing NFCs to declare their data correct in case a warning or message has been generated by the validation procedure allows accepting values outside the range, e.g. to record the results of extreme events. It also acts as an override option for changes in constant parameters specifying the plot, which happens frequently when the plot coordinates are re-assessed. The data may then enter the validated database although the actual values prompting the system to generate a warning or an error message have not been adjusted.

BIBLIOGRAPHY

Augustin, S., A. Bolte, M. Holzhausen and B. Wolff (2005) Exceedance of critical loads of nitrogen and sulphur and its relation to forest conditions. European Journal of Forest Research, 124 (4), p. 289-300.

De Vries, W., G.J. Reinds, M. Posch, M.J. Sanz, G.H.M. Krause, V. Calatayud, J.P. Renaud, J.L. Dupouey, H. Sterba, E.M. Vel, M. Dobbertin, P. Gundersen, J.C.H. Voogd. (2003) Intensive Monitoring of Forest Ecosystems in Europe. Technical Report 2003. UN/ECE and EC, Geneva and Brussels, 161pp.

European Commission (ed) (2007). Forest Focus Monitoring Database System – Technical Specifications 2006 Level II Data. EUR 22714 EN, Office for Official Publications of the European Communities, Luxembourg. 102pp.

Hiederer. R. T. Durrant, O. Granke, M. Lambotte, M. Lorenz, B. Mignon, V. Mues (2008) Forest Focus Monitoring Database System – Technical Report 2006 Level II Data. EUR 23578 EN. Office for Official Publications of the European Communities, Luxembourg. 161 pp.

Hiederer, R. T. Durrant, O. Granke, M. Lambotte, M. Lorenz, B. Mignon, V. Mues, V. (2007) Forest Focus Monitoring Database System – Validation Methodology. EUR 23020 EN. Office for Official Publications of the European Communities, Luxembourg. 56pp.

Lorenz, M., R. Fischer, G. Becher, V. Mues, W. Seidling, P. Kraft and H-D. Nagel (2006) Forest Condition in Europe. Technical Report 2006. UN/ECE, Geneva, 113 pp + Annexes.

Nigot, S., B. Mignon and R. Hiederer (2006) Forest Focus Monitoring Database System – Submission Module User Manual. EUR 22184 EN, Office for Official Publications of the European Communities, Luxembourg. 29pp.

Abstract

Forest Focus (Regulation (EC) No 2152/2003) is a Community scheme for harmonized, broad-based, comprehensive and long-term monitoring of European forest ecosystems. Under this scheme the monitoring of air pollution effects on forests is carried out by participating countries on the basis of the systematic network of observation points (Level I) and of the network of observation plots for intensive and continuous monitoring (Level II).

According to Article 15(1) of the Forest Focus Regulation Member States shall annually, through the designated authorities and agencies, forward to the Commission geo-referenced data gathered under the scheme, together with a report on them by means of computer telecommunications and/or electronic technology. For managing the data JRC has implemented a Forest Focus Monitoring Database System.

This report presents the methodology used to validate Level II data during the 2008 processing period. The procedures applied at the various stages of checking data Compliance, Conformity and Uniformity are described. The report also provides detailed information on the interpretation of data formats and the threshold values used in the Conformity range tests.

Mission of the JRC

The mission of the Joint Research Centre is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of European Union policies. As a service of the European Commission, the Joint Research Centre functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

JRC
EUROPEAN COMMISSION

Publications Office
Publications.eu.int