Data FAQs

How to resubmit / correct data?

ICP Forests under the UNECE Convention
on Long-range Transboundary Air Pollution (Air Convention)

United Nations Economic Commission for Europe (UNECE)
Convention on Long-range Transboundary Air Pollution (Air Convention)
International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects
on Forests (ICP Forests)

http://icp-forests.net

CONTENTS

INTRODUCTION

This document tries to give examples in form of a "How-To-Guide" to correct and resubmit data in the ICP Forests database.

It does not cover a general description of how to use the data portal.

A user guide of the ICP Forests data portal can be found here:

https://www.icp-forests.org/pdf/icpf_data_portal_user_guide.pdf

This document as well as the data portal user guide are open documents which are adapted continuously. Please update your copies regularly and let us know if something is missing or unclear.

HOW TO CORRECT MY DATA

What should I do, if a mistake was found in my data?

- Go to the "CHECKOUT DATA" menu of the data portal (see user guide data portal).
- Checkout the submission which has to be corrected.
- You will get your data in the exact format needed for the resubmission.
- Make your correction in the downloaded data and resubmit it.

What should I do if a mistake covers **several years**?

If you think that the problem can easily be updated directly in the database, please get in contact with us.

Examples might be that a wrong unit has been used and the results can easily be recalculated or the IDs of samplers have to be updated.

If it is something more complex, then you have to resubmit all affected data year by year, starting with the earliest one.

In any case you can always use the checkout module in the data portal to get your data in a format ready for resubmission (see above).

If I want to **resubmit several years** of a survey, where should I start?

Start with the earliest of the years you want to resubmit!
Then go on year by year! For instance, start with year 2009, then go on with 2010 and so on.

User guide of ICP Forests data portal: https://www.icp-forests.org/pdf/icpf_data_portal_user_guide.pdf

Can I use the current forms to (**re-)submit all old survey years**?

Yes! The current forms work for all years.
Old forms which may have existed in the past, have been merged into the current forms.

The current forms documented under https://icp-forests.org/documentation are the only version of forms to be used. SUBMITTING data from the past IN OLD FORMS IS NOT POSSIBLE.

Some attributes were not assessed in the past, because they were introduced later to the programme. Other attributes may have changed from optional to mandatory in the meantime. Please try to submit all information available as complete as possible. Optional parameters which cannot be submitted can be left empty (NULL). Mandatory attributes can be submitted with a specific 'NULL' value for the time before they became mandatory or have been introduced. You can find the specific 'NULL' values in the online documentation in the respective dictionary or remark.
Some examples:

- In CC-TRC (crown condition parameters) – the social class can be reported as -99 (= Unknown) for years before 2007 (as documented in d_social_class)
- In C1-TRC (crown condition parameters) – removal_mortality can be reported as -99 (=No information provided) for years before 2011 (as documented in d_removal_mortality_ccgr)

TECHNICAL PROBLEMS WITH DATA PORTAL

When I would like to (re)submit data, why **can't** I **choose the survey I would like to submit**?

There can be only one survey year under construction at a time.

So if you are not able to open a submission for a specific survey, please check if there is another submission of the same survey not finalized yet and finish it first.

If you cannot finalize the open submission, you can delete all files already uploaded for this submission. This will open up the system again for another survey year.

See data portal user guide – 2.2 Open a new submission.

GETTING OVERVIEW OF (RE)SUBMISSIONS

Where do I get an **overview of what has been submitted** to the database?

There is an overview of all submissions and number of submitted plots on the website.

Please see: https://icp-forests.org/documentation/Overview_lists/Overview_of_submissions.html

For a specific survey you find an overview of what has been submitted per year in the "adds" folder when downloading your data. It is named: „*dar_plots_submitted_SURVEYCODE.csv*".

!!! Please read the README provided with the "adds" folder carefully !!!

Where do I get an **overview of data,** which is known to exist, but was **not yet submitted to the database**?

When downloading your data you find a report in the folder "adds" named: „*dar_plots_submitted_SURVEYCODE.csv*" (see above)

Submissions marked with "-9999" have been reported from the corresponding National Focal Centre to exist and will be (re)submitted in future.

!!! Please read the README provided with the "adds" folder carefully !!!

Where do I get an **overview about problems** existing in my data?

There are two options:

- When downloading your data there is a report in the "adds" folder named:

  „*known_problems_agg_survey_year_SURVEYCODE.csv*"

  This is an aggregated error report of the data of this survey.
  Every night the PCC tests all data in the database according to the latest error check routines that are currently executed during data submission. So this report summarizes all errors you would get if you resubmitted all your data today.

- In case that further problems have been reported by the EP of the corresponding survey, there might be a sub-folder in the "adds" folder named "*known_problems_ep*".

!!! Please read the README provided with the "adds" folder carefully !!!

User guide of ICP Forests data portal: https://www.icp-forests.org/pdf/icpf_data_portal_user_guide.pdf

FAQS - SURVEY DEPOSITION

Author: Peter Waldner (last change 24.5.2022)

Copyright for all pictures of this FAQ section: Peter Waldner

The last sampling period is overlapping into the next year. - Can we report it with the same SURVEY_YEAR?

Yes, it is not expected that sampling is performed exactly at the end of months or years, e.g. at midnight on the 31. December.
Periods overlapping the end of the year are assigned to a *SURVEY_YEAR* by the submitting country/partner, it just should be consistent across the forms and submissions.
The overlap into another year should of course not be very long (this is checked during submission).
Note: If you have already submitted a period with the previous *SURVEY_YEAR*, you do not need to submit it again.

We sampled continuously. - Shall we report the same DATE_END and DATE_START for two subsequent periods?

Yes, sampling can be carried out at any time during the day. The exact sampling time has been considered to be not relevant and is not reported.
Hence, the *DATE_END* and *DATE_START* of two subsequent sampling periods of continuous sampling should be the same (if sampled during the day) or the day after (if sampled at midnight).

Example: If sampler bottles were first exchanged on the 15. January 2021 this can be reported in the *DEM* form with:
First record having *DATE_END* = 210115.
Second record having *DATE_START* = 210115 (or *DATE_START* = 210116).

Once on a plot, we sampled soil solution but did not exchange the deposition sample bottles. - Should we now leave a period number out?

No, the period numbering does not need to be consistent across surveys.
Instead within a series in a survey, i.e. in deposition data having the same *SURVEY_YEAR*, *CODE_PLOT*, *CODE_SAMPLER* and *SAMPLER_ID*, the period number should preferably always start with 1 and continue without gaps, e.g. 1,2,3,4,5,6,7,8,9,10,11,12.
Hereby, the *DATE_START* of the first period should preferably be equal to the corresponding *DATE_MONITORING_FIRST* and the

DATE_END of the last period equal to the corresponding *DATE_MONITORING_LAST*, etc.

User guide of ICP Forests data portal: https://www.icp-forests.org/pdf/icpf_data_portal_user_guide.pdf

Remark: The period numbering was once introduced to
(i) reduce the file sizes of a survey by splitting into plot and time series information and
(ii) while providing a way to merge them and assign approximate sampling period dates (~ +/- 1 day).
Today, however, the sampling period dates can be provided directly in the DEM form and the period numbering is thus of less importance for merging the PLD and DEM information.

A storm destroyed the samplers. - How shall we report this?

Best is to clearly indicate it in the files if this situation occurred. We recommend to add a record for all periods, also for periods with missing or unreliable data.
The *CODE_VSAMPLING* 2 ('contamination'), 3 ('sampler destroyed') or 4 ('sampling not performed') can be used to classify the problem and variables that were not determined can be left empty.

There was not enough rain for chemical analyses. - Do I have to report anything then?

Best is to clearly indicate if there was a continues reliable sampling or not. Hence:
If you exchanged the bottles, measured the quantity but could not perform all or any analyses, then it is recommended to add a record to the *DEM* form with the quantity, leaving the variables that could not be determined empty and add an appropriate *CODE_VSAMPLING* (e.g. 7 ('too little volume for any analyses') or 8 ('too little volume for some parameters')).
If you exchanged the bottles, measured the quantity and kept the samples to pool it with the samples of the following period, then add one record for each of the two sampling periods, with the correct quantities, *CODE_VSAMPLING* 9 for the first record and *CODE_VSAMPLING* 1 for the second record and provide two times the same values for the parameters that could only be determined after pooling.
If you left the bottles in the field, the sampling period can be seen as being prolonged (not recommended, especially if the duration will be become longer than a month) and it can be reported as one period ending when you collect the bottles e.g. during the next visit.

User guide of ICP Forests data portal: https://www.icp-forests.org/pdf/icpf_data_portal_user_guide.pdf

We use winter and summer samplers. - Shall we use *CODE_SAMPLER_MODEL* to indicate this?

No, this variable was introduced for the project 'HarmonDepo' where one same sampler model (*CODE_SAMPLER_MODEL* 2 ('Harmonised')) has been compared on many plots across Europe to the 'National' samplers.

Having winter and summer sampler types can be reported by adding more than one record to the *PLD* form. This way the plotwise information (number of samplers, their surface area and height) can be reported in the *PLD* form for both, the winter and the summer samplers, separately. In the *DEM* file, the periods can be numbered continuously over the whole year.

Typically, three records are provided in the *PLD* files: first record for the sampling periods with the winter samplers at the beginning of the year, second for the sampling periods with the summer samplers and a third record for the sampling periods with winter samplers towards the end of the year.

We have two groups of samplers, one group below conifers another group below broadleaved trees that are analysed separately. - How can we report them individually?

In order to limit the complexity, the database was originally been designed to store plot-wise data. The standard procedure (recommended) is to report a volume weighted average for the whole plot.

However, later the variable *SAMPLER_ID* was introduced that allows a separate reporting of group of samplers used in parallel.

You may thus report records for one group of samplers with a first value for *SAMPLER_ID* (e.g. 1) and report records for the other groups with another values for *SAMPLER_ID* (e.g. 2). However, make sure that the necessary records are provided in all forms for the used *SAMPLER_ID's*.

This will facilitate the merging of the *PLD* and *DEM* records by including *SAMPLER_ID* as a further 'key'.

We have three groups of samplers running in parallel, one group for standard parameters, another group that is pre-acidified for heavy metals and a third group for mercury. - How can we report this?

This is the situation the *sampler(group)_ID* has been designed for.
You may add separate records to the *PLD*, *DEM* (and *LQA*) forms using one *sampler(group)_ID* value, e.g. 1 for standard samplers and another *sampler(group)_ID* value, e.g. 2, for the heavy

User guide of ICP Forests data portal: https://www.icp-forests.org/pdf/icpf_data_portal_user_guide.pdf

metal samplers and a third *sampler(group)_ID* value, e.g. 3 for the mercury sampler group. The three groups of samplers can, but do not necessary need to be collected in the same intervals.

We take an aliquot of the monthly samples, pool it to quarterly samples and analyse them for some additional parameters. - How can we best report those values?

Similarly to the a group of samplers, you may use another *sampler(group)_ID* to report the values and the corresponding quantities for the quarterly samples. Add the necessary records for this *sampler(group)_ID* to all forms (*PLD*, *DEM*, *LQA*).

We have developed a reliable method to reconstruct missing data, e.g. the quantity in case of overflow of the sampler. - Shall we report these corrected values?

Yes, we recommend to submit the corrected values and mark it with CODE_VSAMPLING 5 ('overflow, estimated precipitation') and/or in *OTHER_OBSERVATIONS*.
Best is to provide a description of the method used for estimating the values in a file that is added as DAR-Q file to the data (see related option on the data submission portal).

User guide of ICP Forests data portal: https://www.icp-forests.org/pdf/icpf_data_portal_user_guide.pdf