

EBAS Data format

Technical workshop on data quality and data reporting
to EBAS

October 26 - 28th 2016

Paul Eckhardt, ATMOS, NILU

pe@nilu.no



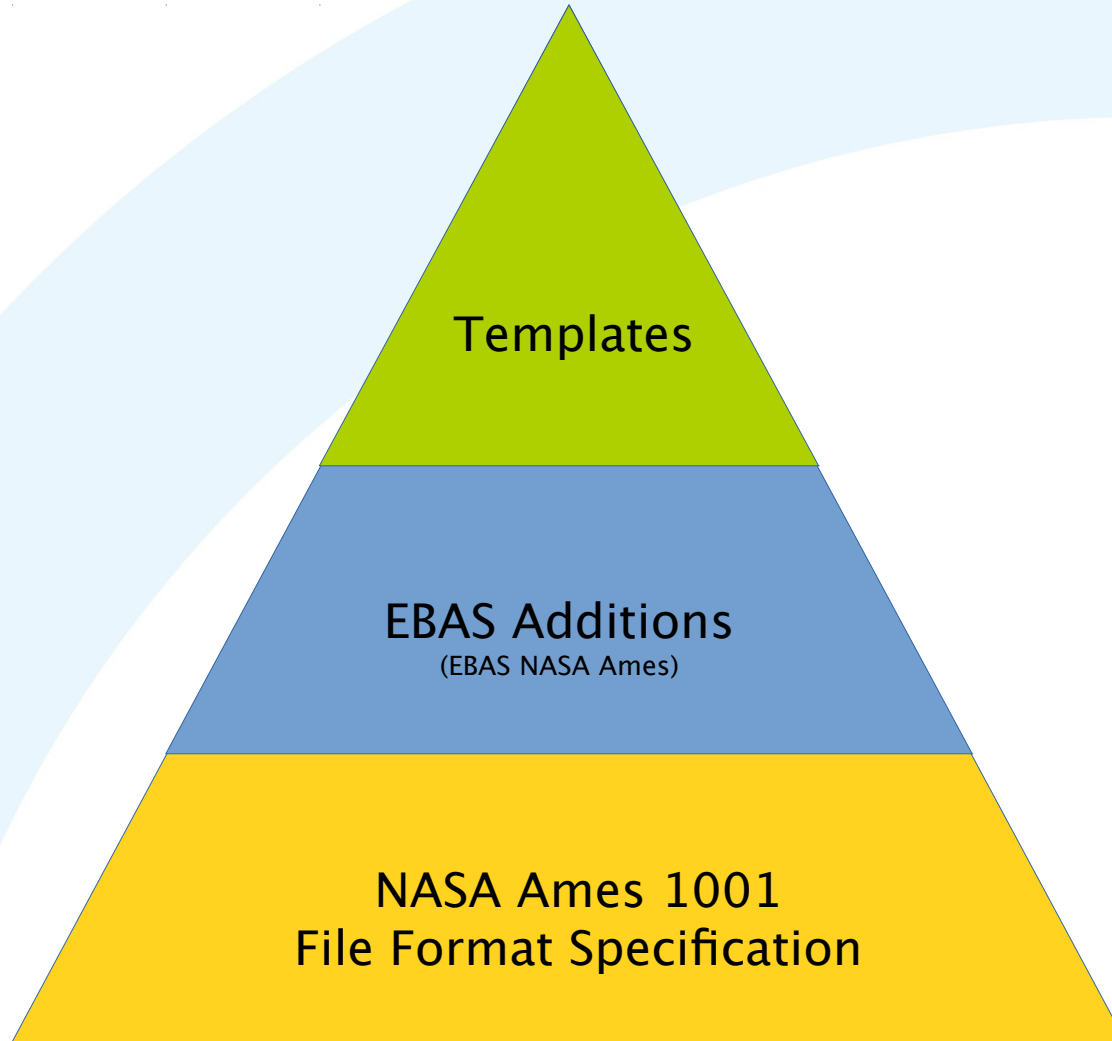
Reporting data to EBAS

A good starting point is the **EBAS Data Submission Manual**:

<http://ebas-submit.nilu.no>

- Templates for different types of data
- Line by line, metadata by metadata explanation
- Templates can be used as a starting point (copy and paste) if you submit for the first time, but:
 - Walk through all metadata and customize to your station and instrument
 - Check if all variables are needed (delete those not needed, please do not report all data missing if one parameter is not measured!)
 - After finishing the header, add your data
- Alternative: create files by software (this will be a separate session)

The EBAS NASA Ames format



The EBAS NASA Ames format

The EBAS data format is based on the NASA Ames file format specification, more specific on the File Format Index (=subtype) 1001, which is designed for:

- One unbounded independent variable
- Independent variable is a of data type real
- All dependent variables are real

The EBAS NASA Ames format

The EBAS data format is based on the NASA Ames file format specification, more specific on the File Format Index (=subtype) 1001, which is designed for:

- One unbounded independent variable
 - **EBAS uses time as the independent variable (*start time*).**
- Independent variable is a of data type real
- All dependent variables are real

The EBAS NASA Ames format

The EBAS data format is based on the NASA Ames file format specification, more specific on the File Format Index (=subtype) 1001, which is designed for:

- One unbounded independent variable
 - EBAS uses time as the independent variable (*start time*).
 - **This means, technically the *end time* is a dependent variable (we need to accept this)**
- Independent variable is a of data type real
- All dependent variables are real

The EBAS NASA Ames format

The EBAS data format is based on the NASA Ames file format specification, more specific on the File Format Index (=subtype) 1001, which is designed for:

- One unbounded independent variable
 - EBAS uses time as the independent variable (*start time*).
 - This means, technically the *end time* is a dependent variable (we need to accept this)
- Independent variable is a of data type real
 - **This implies: time has to be real (we use “days since” reference date).**
- All dependent variables are real

The EBAS NASA Ames format

The EBAS data format is based on the NASA Ames file format specification, more specific on the File Format Index (=subtype) 1001, which is designed for:

- One unbounded independent variable
 - EBAS uses time as the independent variable (*start time*).
 - This means, technically the *end time* is a dependent variable (we need to accept this)
- Independent variable is a of data type real
 - This implies: time has to be real (we use “days since” reference date).
- All dependent variables are real
 - **Dependent variables in an EBAS datafile are: *end time* and all *measured values, statistical variables* (std dev, percentiles, etc) and *flags*.**

NASA Ames 1001

NASA Ames File Format Specification can be found at

<https://cloud1.arc.nasa.gov/solve/archiv/archive.tutorial.html>

Another very good documentation by the British Atmospheric Data Centre can be found at:

<https://badc.nerc.ac.uk/help/formats/NASA-Ames/>

Format Specification for Data Exchange

By

Steven E. Gaines R. Stephen Hipskind

Version 1.3

18 June 1998

Version 1:
Version 1.1:
Version 1.2:

16 May 1990
6 February 1992
12 January 1998
R.S.H.

Voice: 650/604-5076
FAX: 650/604-3625
Email: hipskind@cloud1.arc.nasa.gov

S.E.G.

Voice: 650/604-4546
FAX: 650/604-3625
Email: gaines@cloud1.arc.nasa.gov



FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

```
NLHEAD      1001
ONAME
ORG
SNAME
MNAME
IVOL      NVOL
DATE      RDATE
DX(1)
XNAME(1)
NV
[ VSCAL(n), n=1,NV ]
[ VMISS(n), n=1,NV ]
[ VNAME(n) ] n=1,NV
NSCOML
[ SCOM(k) ] k=1,NSCOML
NNCOML
[ NCOM(k) ] k=1,NNCOML
[ X(m,1) ( V(m,n), n=1,NV ) ]
```

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

```
NLHEAD      1001
ONAME
ORG
SNAME
MNAME
IVOL      NVOL
DATE      RDATE
DX(1)
XNAME(1)
NV
[ VSCAL(n), n=1,NV ]
[ VMISS(n), n=1,NV ]
[ VNAME(n) ] n=1,NV
NSCOML
[ SCOM(k) ] k=1,NSCOML
NNCOML
[ NCOM(k) ] k=1,NNCOML
[ X(m,1) ( V(m,n), n=1,NV ) ]
```

Number of lines used
for the file header

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

```
NLHEAD      1001
ONAME
ORG
SNAME
MNAME
IVOL      NVOL
DATE      RDATE
DX(1)
XNAME(1)
NV
[ VSCAL(n), n=1,NV ]
[ VMISS(n), n=1,NV ]
[ VNAME(n) ] n=1,NV
NSCOML
[ SCOM(k) ] k=1,NSCOML
NNCOML
[ NCOM(k) ] k=1,NNCOML
[ X(m,1) ( V(m,n), n=1,NV ) ]
```

NASA Ames File Format
Index (subtype of the
NASA Ames specification)

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

```
NLHEAD      1001
ONAME
ORG
SNAME
MNAME
IVOL      NVOL
DATE      RDATE
DX(1)
XNAME(1)
NV
[ VSCAL(n), n=1,NV ]
[ VMISS(n), n=1,NV ]
[ VNAME(n) ] n=1,NV
NSCOML
[ SCOM(k) ] k=1,NSCOML
NNCOML
[ NCOM(k) ] k=1,NNCOML
[ X(m,1) ( V(m,n), n=1,NV ) ]
```

- NASA Ames: Originator Name
FREE TEXT
- EBAS: Data Originator Names
STRICT SYNTAX

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

```
NLHEAD      1001
ONAME
ORG
SNAME
MNAME
IVOL      NVOL
DATE      RDATE
DX(1)
XNAME(1)
NV
[ VSCAL(n), n=1,NV ]
[ VMISS(n), n=1,NV ]
[ VNAME(n) ] n=1,NV
NSCOML
[ SCOM(k) ] k=1,NSCOML
NNCOML
[ NCOM(k) ] k=1,NNCOML
[ X(m,1) ( V(m,n), n=1,NV ) ]
```

- NASA Ames: Organization Name
FREE TEXT
- EBAS: Organization Metadata
STRICT SYNTAX

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

```
NLHEAD      1001
ONAME
ORG
SNAME
MNAME
IVOL      NVOL
DATE      RDATE
DX(1)
XNAME(1)
NV
[ VSCAL(n), n=1,NV ]
[ VMISS(n), n=1,NV ]
[ VNAME(n) ] n=1,NV
NSCOML
[ SCOM(k) ] k=1,NSCOML
NNCOML
[ NCOM(k) ] k=1,NNCOML
[ X(m,1) ( V(m,n), n=1,NV ) ]
```

- NASA Ames: Source of the measurements
FREE TEXT
- EBAS: Data Submitter Names
STRICT SYNTAX

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

```
NLHEAD      1001
ONAME
ORG
SNAME
MNAME
IVOL      NVOL
DATE      RDATE
DX(1)
XNAME(1)
NV
[ VSCAL(n), n=1,NV ]
[ VMISS(n), n=1,NV ]
[ VNAME(n) ] n=1,NV
NSCOML
[ SCOM(k) ] k=1,NSCOML
NNCOML
[ NCOM(k) ] k=1,NNCOML
[ X(m,1) ( V(m,n), n=1,NV ) ]
```

- NASA Ames: Mission Name
FREE TEXT
- EBAS: Framework Associations
STRICT SYNTAX

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

```
NLHEAD      1001
ONAME
ORG
SNAME
MNAME
IVOL      NVOL
DATE      RDATE
DX(1)
XNAME(1)
NV
[ VSCAL(n), n=1,NV ]
[ VMISS(n), n=1,NV ]
[ VNAME(n) ] n=1,NV
NSCOML
[ SCOM(k) ] k=1,NSCOML
NNCOML
[ NCOM(k) ] k=1,NNCOML
[ X(m,1) ( V(m,n), n=1,NV ) ]
```

- NASA Ames: IVOL (volume number), NVOL (total number of volumes)
- EBAS: Not used, set to 1 1 to satisfy NASA Ames compatibility

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

NLHEAD

ONAME

ORG

SNAME

MNAME

IVOL NVOL

DATE RDATE

DX(1)

XNAME(1)

NV

[VSCAL(n), n=1,NV]

[VMISS(n), n=1,NV]

[VNAME(n)] n=1,NV

NSCOML

[SCOM(k)] k=1,NSCOML

NNCOML

[NCOM(k)] k=1,NNCOML

[X(m,1) (V(m,n), n=1,NV)]

- NASA Ames: UT date, start of the data
SYNTAX YYYY MM DD
- EBAS: Reference date, zero point for the
time axes in the file

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

```
NLHEAD      1001
ONAME
ORG
SNAME
MNAME
IVOL      NVOL
DATE      RDATE
DX(1)
XNAME(1)
NV
[ VSCAL(n), n=1,NV ]
[ VMISS(n), n=1,NV ]
[ VNAME(n) ] n=1,NV
NSCOML
[ SCOM(k) ] k=1,NSCOML
NNCOML
[ NCOM(k) ] k=1,NNCOML
[ X(m,1) ( V(m,n), n=1,NV ) ]
```

- NASA Ames: Revision date
SYNTAX YYYY MM DD
- EBAS: Revision date

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

```
NLHEAD      1001
ONAME
ORG
SNAME
MNAME
IVOL      NV
DATE
DX(1)
XNAME(1)
NV
[ VSCAL(n), n=1,NV ]
[ VMISS(n), n=1,NV ]
[ VNAME(n) ] n=1,NV
NSCOML
[ SCOM(k) ] k=1,NSCOML
NNCOML
[ NCOM(k) ] k=1,NNCOML
[ X(m,1) ( V(m,n), n=1,NV ) ]
```

- NASA Ames: Interval between values of the independent variable, 0 for non-uniform interval
- EBAS: no change

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

NLHEAD 1001

ONAME

- NASA Ames: Description of the independent variable
FREE TEXT
- EBAS: only one unit supported:
days from file reference point

DATE

DX(1)

XNAME(1)

NV

[VSCAL(n), n=1,NV]

[VMISS(n), n=1,NV]

[VNAME(n)] n=1,NV

NSCOML

[SCOM(k)] k=1,NSCOML

NNCOML

[NCOM(k)] k=1,NNCOML

[X(m,1) (V(m,n), n=1,NV)]

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

NLHEAD 1001

ONAME

ORG

- NASA Ames: Number of independent variables
- EBAS: no change

SNA

MNAME

IVOL NV

DATE

DX(1)

XNAME(1)

NV

[VSCAL(n), n=1,NV]

[VMISS(n), n=1,NV]

[VNAME(n)] n=1,NV

NSCOML

[SCOM(k)] k=1,NSCOML

NNCOML

[NCOM(k)] k=1,NNCOML

[X(m,1) (V(m,n), n=1,NV)]

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

```
NLHEAD    1001
ONAME
```

- NASA Ames: scale factor for each independent variable
- EBAS: no scale factors used, fixed to 1 1 1...

```
IVOL
DATE
DX(1)
XNAME(1)
NV
[ VSCAL(n), n=1,NV ]
[ VMISS(n), n=1,NV ]
[ VNAME(n) ] n=1,NV
NSCOML
[ SCOM(k) ] k=1,NSCOML
NNCOML
[ NCOM(k) ] k=1,NNCOML
[ X(m,1) ( V(m,n), n=1,NV ) ]
```


FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

```
NLHEAD    1001  
ONAME
```

- NASA Ames: Missing value for each independent variable
Must be larger than any "good" data value for variable
- EBAS: Additional requirements/recommendations (all digits 9, at least one order of magnitude larger than the largest "good" value used)

```
ΔI  
NV  
[ VSCALE(n), n=1,NV ]  
[ VMISS(n), n=1,NV ]  
[ VNAME(n) ] n=1,NV  
NSCOML  
[ SCOM(k) ] k=1,NSCOML  
NNCOML  
[ NCOM(k) ] k=1,NNCOML  
[ X(m,1) ( V(m,n), n=1,NV ) ]
```

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

```
NLHEAD    1001
ONAME
```

- NASA Ames: Name, description and unit for each independent variable

FREE TEXT

- EBAS: Metadata for each variable

STRICT SYNTAX

```
ΔI
NV
[ VSCALE(n) ] n=1,NV ]
[ VMISS(n) ] n=1,NV ]
[ VNAME(n) ] n=1,NV
NSCOML
[ SCOM(k) ] k=1,NSCOML
NNCOML
[ NCOM(k) ] k=1,NNCOML
[ X(m,1) ( V(m,n), n=1,NV ) ]
```

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

NLHEAD 1001

ONAME

ORG

SNAME

MNAME

IVOL NVOL

DATE RDATE

DX

XI

NV

[VSCAL()] n=1, NV]

[VMISS()] n=1, NV]

[VNAM()] n=1, NV]

NSCOML

[SCOM(k)] k=1, NSCOML

NNCOML

[NCOM(k)] k=1, NNCOML

[X(m, 1) (V(m, n), n=1, NV)]

- NASA Ames: Number of special comment lines
- EBAS: no special comment lines used, set to 0

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

NLHEAD 1001

ONAME

ORG

SNAME

MNAME

IVOL NVOL

DATE RDATE

DX

XI

NV

[VSCAL ()]

[VMISS ()]

[VNAME] n=1, NV

NSCOML

[SCOM(k)] k=1, NSCOML

NNCOML

[NCOM(k)] k=1, NNCOML

[X(m, 1) (V(m, n), n=1, NV)]

- NASA Ames: Special comment lines

FREE TEXT

- EBAS: no special comment lines used

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

NLHEAD 1001

ONAME

ORG

SNAME

MNAME

IVOL NVOL

DATE RDATE

DX(1)

X

N

[

[VMISS()]

[VNAME()] n=1, NV

NSCOML

[SCOM()] k=1, NSCOML

NNCOML

[NCOM(k)] k=1, NNCOML

[X(m, 1) (V(m, n), n=1, NV)]

- NASA Ames: Number of normal comment lines
- EBAS: no change

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

```
NLHEAD    1001
ONAME
ORG
SNAME
```

- NASA Ames: Normal comment lines
FREE TEXT
- EBAS:
 - All file global (file default) EBAS metadata
 - Syntax in each line: <EBAS metadata tag>: <value>
 - Last line: column header comment (FREE TEXT)

```
[          n=1,NV
NSCO
[ SC      ] k=1,NSCOML
NNCOML
[ NCOM(k) ] k=1,NNCOML
[ X(m,1)  ( V(m,n), n=1,NV ) ]
```

FFI = 1001:

One real, unbounded independent variable (NIV=1).

Primary variables are real.

No auxiliary variables.

Independent and primary variables are recorded in the same record.

```
NLHEAD    1001
ONAME
ORG
SNAME
MNAME
```

- NASA Ames: Data
each line contains one values for the independent variable and for each dependent variable, delimited by one or more spaces
- EBAS: no changes

```
[ ... n=1, NV ]
[ ... n=1, NV
NSC
[ SC ... ] k=1, NSCOML
NNCOM
[ NCOM(k) ] k=1, NNCOML
[ X(m, 1) ( V(m, n), n=1, NV ) ]
```

EBAS Metadata

- Formal metadata are very limited in NASA Ames
- Additional metadata important for EBAS are added as *Tag: Value* pairs
 - *Examples for EBAS metadata tags: Station code, Laboratory code, Instrument name, Detection limit, Measurement uncertainty, ... (> 100 elements)*
 - *EBAS defines:*
 - *a list of allowed metadata tags*
 - *restrictions for the values of those metadata elements (data type, value range, unit, controlled vocabulary)*
 - *File global (file default) metadata specified in NCOM block of NASA Ames*
 - *Metadata for each variable can be set in the VNAME lines (overrides file global setting for this variable)*

EBAS Metadata: File global

File global metadata specified in NCOM block of NASA Ames

- NCOM block contains the normal comment lines, which “*should be used to more completely describe the contents of the file*”. This is what the EBAS metadata are intended for.
- EBAS specifies in addition to NASA Ames 1001 a strict syntax for the comment block (syntax `Tag: Value`, list of metadata tags and restrictions for the metadata values)
- The file global metadata apply generally to all variables in the file, *but they may be overridden for each variable* (default, see also per variable setting on next slide). [EBAS exporter sets a value in the file global metadata if more than 50% of the variables use the same value for the respective metadata element. Otherwise the value is omitted and specified for each variable. This can also be seen as a recommendation]
- Empty metadata values (“`tag:`”) have the same meaning as omitted metadata elements. It is therefore no difference if a line contains a tag without value or the line is deleted altogether. Both can be understood as “metadata not reported” (for whatever reason: not applicable, unknown, forgotten).

EBAS Metadata: Per variable

Metadata for each variable can be set in the VNAME lines of NASA Ames:

- VNAME lines contains the *name, description and unit* for each variable. This is what the EBAS metadata are intended for.
- EBAS specifies in addition to NASA Ames 1001 a strict syntax for the VNAME lines:
<EBAS component name>, <unit>, <tag>=<value>, ...
(mind the syntax *tag=value* for VNAME instead of *tag: value* in the NCOM block)
- File global metadata (set in the NCOM block) can be overridden individually for each variables in the VNAME lines.
- Empty metadata values (“tag=”) are interpreted as “**metadata element not reported**” (for whatever reason: not applicable, unknown, forgotten). This can be used to *reset* a metadata element to empty for one variable (when it was set to any value in the file global metadata)

EBAS Metadata: Example

Here is an example snippet of a EBAS NASA Ames file header:

```
...
sulphate_total, ug S/m3
calcium, ug/m3, Measurement uncertainty=, Detection limit=0.01 ug/m3
sulphur_dioxide, ug S/m3, Matrix=air, Medium=Cellulose, Coating/Solution=KOH
...
Matrix:                aerosol
Detection limit:       0.01 ug S/m3
Medium:                Teflon
Coating/Solution:
Measurement uncertainty: 10%
```

VNAME lines:
per variable metadata

NCOM block:
file global metadata



EBAS Metadata: Example

Here is an example snippet of a EBAS NASA Ames file header:

```
...
sulphate_total, ug S/m3
calcium, ug/m3, Measurement uncertainty=, Detection limit=0.01 ug/m3
sulphur_dioxide, ug S/m3, Matrix=air, Medium=Cellulose, Coating/Solution=KOH
...
Matrix:                aerosol
Detection limit:       0.01 ug S/m3
Medium:                Teflon
Coating/Solution:
Measurement uncertainty: 10%
```

Coating/Solution is empty, could also be left away (delete the whole line)

EBAS Metadata: Example

Here is an example snippet of a EBAS NASA Ames file header:

```
...
sulphate_total, ug S/m3
calcium, ug/m3, Measurement uncertainty=, Detection limit=0.01 ug/m3
sulphur_dioxide, ug S/m3, Matrix=air, Medium=Cellulose, Coating/Solution=KOH
...
Matrix:                aerosol
Detection limit:       0.01 ug S/m3
Medium:                Teflon
Measurement uncertainty: 10%
```

Coating/Solution is empty, could also be left away (delete the whole line)

EBAS Metadata: Example

Here is an example snippet of a EBAS NASA Ames file header:

Variable sulphate_total: Component name and Unit set (minimum), all other metadata are inherited from the file global metadata (i.e. Matrix, Detection limit, Medium and Measurement uncertainty)

```
...
sulphate_total, ug S/m3
calcium, ug/m3, Measurement uncertainty=, Detection limit=0.01 ug/m3
sulphur_dioxide, ug S/m3, Matrix=air, Medium=Cellulose, Coating/Solution=KOH
...
Matrix:                aerosol
Detection limit:       0.01 ug S/m3
Medium:                Teflon
Coating/Solution:
Measurement uncertainty: 10%
```

EBAS Metadata: Example

Here is an example snippet of a EBAS NASA Ames file header:

Variable calcium: Component name, Unit, and Detection limit set, Measurement uncertainty is reset to **not defined**, all other metadata are inherited from the file global metadata (i.e. Medium)

```
...
sulphate_total, ug S/m3
calcium, ug/m3, Measurement uncertainty=, Detection limit=0.01 ug/m3
sulphur_dioxide, ug S/m3, Matrix=air, Medium=Cellulose, Coating/Solution=KOH
...
Matrix:                aerosol
Detection limit:       0.01 ug S/m3
Medium:                Teflon
Coating/Solution:
Measurement uncertainty: 10%
```

EBAS Metadata: Example

Here is an example snippet of a EBAS NASA Ames file header:

```
...
sulphate_total, ug S/m3
calcium, ug/m3, Measurement uncertainty=, Detection limit=0.01 ug/m3
sulphur_dioxide, ug S/m3, Matrix=air, Medium=Cellulose, Coating/Solution=KOH
...
Matrix: aerosol
Detection limit: 0.01 ug S/m3
Medium: Teflon
Coating/Solution:
Measurement uncertainty: 10%
```

Note: Detection limit must be set on this variable, because otherwise we end with a variable calcium, with unit ug/m3 and a detection limit in a different unit (ug S/m3), which is **not allowed**

EBAS Metadata: Example

Here is an example snippet of a EBAS NASA Ames file header:

Variable sulphur_dioxide: Component name, Unit, Matrix, Medium and Coating/Solution set, all other metadata are inherited from the file global metadata (i.e. Detection limit and Measurement uncertainty)

```
...
sulphate_total, ug S/m3
calcium, ug/m3, Measurement uncertainty=, Detection limit=0.01 ug/m3
sulphur_dioxide, ug S/m3, Matrix=air, Medium=Cellulose, Coating/Solution=KOH
...
Matrix: aerosol
Detection limit: 0.01 ug S/m3
Medium: Teflon
Coating/Solution:
Measurement uncertainty: 10%
```

Missing Values

NASA Ames 1001:

- VMISS is a value indicating missing or erroneous data
- Must be larger than any "good" value of the respective variable recorded in the file.
- The value of VMISS defined in the file header is the same value that appears in the data records for missing/bad values of the respective variable

EBAS adds some requirements and recommendations:

- The number format of the missing value code specifies the number format of the corresponding data column in the data section (use of decimal or scientific notation, number of digits left and right of the decimal point).
- The missing value *should* be at least one order of magnitude higher than the highest value for this variable in the data section
- All digits in the missing value *must* be nines.



Flag variables

- Each measurement value can be accompanied by one or more flags
(Full list of flags: <http://www.nilu.no/projects/ccc/flags/>)
- Thus, in addition to data variables (measurements), flag variables are needed
- A flag variable may apply to one or more data variables
 - In order to use one flag variable for more than one data variable, the whole sequence of flags in the file must be exactly the same for all variables sharing the flag variable
 - A flag variable always **follows** the data variable(s) it applies to. When a flag variable applies to more than one data variable, the data variables must be in sequence, directly followed by the applicable flag variable. (With other words: A flag variable applies to all data variables between the previous flag variable and the current one.)

Flag variables: VNAME

- Variable names for flag variables *always* start with the keyword `numflag`
- Recommendation:
 - If the flag variable applies to one single data variable, the VNAME of the data variable is used, prepended with the keyword `numflag` and the unit is changed to `no unit`.

This way, it is additionally *explicitly* stated which variable the flag belongs to.

Example:

```
calcium, ug/m3  
numflag calcium, no unit  
chloride, ug/m3  
numflag chloride, no unit
```

- If the flag variable applies to more than one data variables, the VNAME should be only `numflag, no unit`

Example:

```
calcium, ug/m3  
chloride, ug/m3  
numflag, no unit
```

Flag variables: values (flag encoding)

Requirements:

- As defined before, each measurement can be accompanied by one or more flags
- Each flag is represented by a three digit number
- NASA Ames 1001: all variables must be of type real

Encoding the EBAS flags into one real value:

- the flag value is always a float number with value 0 left of the comma
- Right of the comma, groups of three digits are used to express the single EBAS flags

Examples:

0.999

One flag (999), meaning "*Missing measurement, unspecified reason*". Please keep in mind, that this is not a missing value for the flag variable (it must be applied to a missing value of the corresponding data variable though)

0.999000

Same as above, except that there would be space for a second flag (which is not set)

0.676647392

Flag value encoding three different EBAS flags. 676: "station inside cloud (visibility < 1000 m)", 647: "Fire/wood burning nearby" and 392: "Data completeness less than 75%"

No flag set. Normal, unflagged, valid measurement.

Flag variables: Missing values

- ***Flag values can never be missing!***

Measurement can be missing, but the corresponding flag value will always be valid (e.g. 0.999). “No flag” is encoded as 0.000, also not missing.

- Nevertheless, NASA Ames *requires* a missing value for each variable. As far as NASA Ames 1001 is concerned, an EBAS flag variable is just a regular dependent variable, thus we need to define a missing value in line 12 (VMISS)
- EBAS: Missing value *should* be at least one order of magnitude higher than the highest possible value and all digits *must* be nines
 - The highest possible value for a flag variable is 0.9999990980... (EBAS flags 999, 990, 980 and so on). The number of digits needed varies from file to file and from variable to variable (three times the maximum number of flags that will be used at the same time).
 - Thus the missing value for a flag column should be 9.999999999...

Thanks!

