



Version 3.00

Dec 28, 2001

## Transport files for exchange of field trial data

### ***Background. Purpose of using transport files***

Data from field research should be easily exchanged between different data collecting and database systems. Before transport files could be designed a well tested structure of field trial data has to be developed. Such a structure has been developed, after several years of development, at the Swedish University of Agricultural Sciences (SLU), Uppsala, Sweden.

Field trial data at SLU will be stored in relational database systems where different types of information is grouped in 16 database tables. The tables have been designed to store most types of trial data where the basic identification keys are Series (trial designs), Locations (experimental units) and Trials (identified sets of data for a selected period, in field trial normally one year).

We now suggest a format for exchanging data based on this structure. Work has also started for establishing better rules for describing collected and stored information. Information about this work may be obtained from the Unit for Field Research at SLU in Uppsala.

### ***General description of transport files***

Transport files contains text information either stored in ASCII (text) files or in some selected spreadsheet or database files. If the transport file is a text file, data is stored in the file as delimited table fields or in a fixed format (data in specified positions on the data rows). All rows in text files is terminated by a carriage return - linefeed (ASCII codes 13 + 10) sequence. If database or spreadsheet files are used, all fields or columns in that table/file should be in the text format. In spreadsheet workbook files the data to be transported should be put on the first page of the workbook.

The transport file has the following three elements: Parameter row (A), Header row (B) and Data rows (C). The contents of these rows are described below. There is only one parameter row and it must be the first row in the transport file. After that row, data are grouped depending on the type of information. Each group of data starts with a Header row followed by the corresponding data rows. Data groups may come in any order and also be repeated if new information of a type is added to the file later on.

When exporting data each system software should ensure that the data is consistent with the given parameters in the parameter row and with the specification for each data group according to the respective Header row. During import, the receiving system should warn against or prevent import of data in formats which the importing software cannot handle.

When naming transport files, use no more than 8 characters (preferably use only letters A-Z, digit 0-9 and the underscore) for the name and three characters for the filename extension. Not all systems can handle long filenames or special or national characters. For other files than text files it is recommended to use 1-2 year older versions since the receiver may not have the latest versions.

### ***Parameter row (A)***

This row holds 6-9 parameters delimited by the space and/or a given field delimiter character (text files) or entered in the first 6-9 columns fields (database/spreadsheet transport files). The parameters are:

1. Transport file version
2. Date format, required
3. Decimal character, ASCII code, required
4. Field delimiter, required for text files (set to N for other files)
5. Character set, required, 0=IBM PC (MS-DOS) , 1=ANSI (Windows)
6. Language code, like SE for Sweden, UK for England
7. Creation date, optional, optional format
8. Creation time, optional, optional format
9. Created by, optional system/user identification

1. Transport file version: An integer number between 0100 and 9999. This version is 0200.
2. Date format: The dates can be written in common date formats or as Julian numbers. If dates are transferred as Julian numbers, data exchange can be made between systems using different date formats. In this case, specify J or j as date format. In other cases the format can be given by 6-8 upper or lower case letters where y, m and d stands for year, month and day positions. Year, month and day can appear in any order and with or without a delimiter character between. Days and months should be given as numbers and either written with leading zeroes (for days and months numbered 1-9) or with delimiters between (to avoid confusion). Years can be given with two or four digits but four is recommended. If a delimiter character is used, be careful to choose one that is not a number or the same character as the decimal character (parameter 2 below) or the field delimiter (parameter 3 below). The characters - and / are commonly used. Examples of date formats: yyyy-mm-dd, mm/dd/yy, j (for Julian date numbers).
3. Decimal character: Real numbers can be exchanged either as numbers including a decimal character or with scientific notation. In the latter case, problems of using different decimal characters between different systems can be avoided. Also rounding of very small or large values may be avoided. Write E or e for real numbers in scientific notation. The numbers in the data is written in the format *nnnnnex* (or *nnnnnEx*) here *nnnnn* is the significant digits expressed as an integer number, e (or E) is the exponent identifier and x is the exponent (like 2 for  $10^2$  or -3 for  $10^{-3}$ ). The needed number of digits in the integer part of the scientific number should be enough to accommodate the precision of the exported numbers. For a data collecting system 5 digits may be enough, for single precision real numbers 7 digits may be appropriate, for double precision values, 14 digits may be needed (at least 10 is recommended). When exchanging data where the decimal character is written, the parameter should contain the corresponding ASCII value of that character, normally 46 (decimal point).
4. Field delimiter: If not a text file, set N or n for this parameter. For delimited text files, the ASCII code for the field delimiter should be given, for example 9 for TAB or 59 for semicolon (;). You can only use one delimiting character and it has to be chosen carefully, so that this character doesn't appear elsewhere in the data rows. For fixed length text files, no delimiter is used. Then set this parameter to zero (0). The position on the data rows for each piece of information is described in section C below. In this section you will also find information about the data types in each field for the different data groups.
5. Character set: This parameter defines how national text characters are coded. If they have ASCII codes according to the IBM PC (MS-DOS) standard set the parameter to 0, if they follow the ANSI (Windows) set it to 1. In the future this parameter may be extended to handle different character code page numbers (like 437 for USA, 850 for multilingual Latin I etc.). This extension is not implemented yet.
6. Language code: A short code, maximum four characters, for the country/language of the database system. Use codes like on the Internet, that is SE for Swedish, UK for English etc..
7. Creation date: Optional parameter with the date when the transport file was created.
8. Creation time: Optional parameter with the time when the transport file was created.
9. Created by: Optional parameter with information about the system or user who created the transport file.

## **Header row (B)**

Header rows precede each group of data of a certain type. The header should always have a four character data set name inside brackets [] in position 1-6 on the header row. The rest of the row may be used for optional comments. The following 16 headers/data groups has been defined:

Header    Content, specified in section C below

[FD01] Key information for trials	[FD09] Treatment values
[FD02] Trial locations	[FD10] Plot (experimental unit) values
[FD03] Trial design information	[FD11] Results of statistics analysis
[FD04] List of treatments	[FD12] Library of observed variable codes
[FD05] Detailed list of treatment	[FD13] Library of variety or pesticide codes etc
[FD06] Treatment distribution on plots/exp.units	[FD14] General text library for lists, etc.
[FD07] General trial information	[FD15] Library of data calculation formulas
[FD08] List of registered variables in trials	[FD16] Library of data extraction profiles

Only a few of the tables may actually be included in a transport file. This depend on the information to be transferred. From an analyzing laboratory typically only group [FD08] and one or both of [FD09] and [FD10] would be included. When extracting the complete data for some trials from a database, more data groups will be included, at least groups [FD01] up to [FD10].

## Data rows (C)

After each header follows the data rows. The number of fields reflect the type of data in the respective group (the 16 groups below). The purpose of and content of each group of data is described below.

The "position from-to" in the tables below apply to fixed length text files. The given data types for each field has the following meaning:

TXT Text field, maximum no. of characters, U=Upper case characters stored  
DAT Date field, length always 10 characters (as is needed for dd/mm/yyyy format and the like)  
NUM Numeric (floating/real) field., "-" allows negative values  
NMI Integer (0-32767), "-" allows negative numbers  
NML Large integers, "-" allows negative numbers.  
EXT Text field containing a reference to a multimedia (image, sound) file name  
(\*) The field is used in table indices - the combination of the fields unique for each data row

### [FD01]

Data key group establishing the relation between the Series, Location and Trial identifiers. It is recommended to include this data group in cases where the receiver of the transport file does not have this information. The trial identifier is the unique one in this data group and identifies a period of data (normally one year) for a certain trial Location belonging to a trial Series.

Field	Data type	Position from-to	Comment/use
Trial identifier (*)	TXT 12 U	1-12	The main data grouping key
Series identifier	TXT 12 U	13-24	Related to data with series identifier
Location identifier	TXT 12 U	25-36	Related to data with location identifier
Period (year, subpart)	TXT 12 U	37-48	Year as YYYY, optional subpart added
Crop	TXT 24	49-72	Crop/substrate identifier
District-1	TXT 4	73-76	Code for official agriculture regions
District-2	TXT 4	77-80	Other region/district/station code
Principal investigator	TXT 60	81-140	Responsible for carrying out this trial
Replications	NMI 2	141-142	Number of treatment replications (blocks)
Harvests	NMI 2	143-144	Number of harvests (replied time periods)
Crop rotation year	TXT 4	145-148	Rotation number, for example age of ley
General comments - 1	TXT 120	149-268	Any other comments about this trial (4 fields)
General comments - 2	TXT 120	269-388	
General comments - 3	TXT 120	389-508	
General comments - 4	TXT 120	509-628	
General comments - 5	TXT 120	629-748	

### [FD02]

Data group with information about the location of experiments.

Field	Data type	Position from-to	Comment/use
Location (*)	TXT 12 U	1-12	Location identifier
Location host/owner	TXT 60	13-72	Name & phone
Host address	TXT 60	73-132	Address to host/location
Soil classification	TXT 15	133-147	Classes/codes for soil texture/humus
Soil description/comment	TXT 60	148-207	Description, soil details
Altitude or wgs84 coord.	NUM -	208-223	Approx. level above sea or z-coordinate
Latitude (wgs84)	NUM -	224-239	Latitude or mapping x-coordinates
Longitude (wgs84)	NUM -	240-255	Longitude or mapping y-coordinates
Map file	EXT	256-335	MS-DOS image file names recommended

### [FD03]

This data group holds trial design parameters for a series of trials (title, number of treatment factors and levels etc.).

Field	Data type	Position from-to	Comment/use
Series (*)	TXT 12 U	1-12	Series identifier
Design type	TXT 6 U	13-18	Code for trial design type
Title	TXT 60	19-78	Title of trial series
Study director	TXT 60	79-138	Name & phone to study director
Treatment code pattern	NML 5	139-143	Decoding pattern for treatment codes
Treatment symbols	TXT 5	144-148	Character pattern for printing treatment
Result template	TXT 12 U	149-160	File name for result form template
Levels of factor 1	NMI 2	161-162	Treatment levels in largest (main) plots
Levels of factor 2	NMI 2	163-164	Treatment levels in next largest plots
Levels of factor 3	NMI 2	165-166	Treatment levels in third largest plots
Levels of factor 4	NMI 2	167-168	Treatment levels in fourth largest plots
Levels of factor 5	NMI 2	169-170	Treatment levels in smallest plots
Average calculation	NMI 2	171-172	Method number, 0=none
Statistics codes	TXT 8 U	173-180	For optional use in statistics programs
Statistics statements	TXT 120	181-300	For optional use by statistics programs
Comments	TXT 120	301-420	Comments regarding the trial series
Study PM file ref	TXT 80	421-500	File reference for study documentation

### [FD04]

List of treatment combinations laid out on the experimental units as well as treatment levels listed for each treatment factor

Field	Data type	Position from-to	Comment/use
Series (*)	TXT 12 U	1-12	Series identifier
Treatment number (*)	NMI 4	13-16	Treatment number from 1 and up
Treatment code	NML 8	17-24	Treatment numerical code
Treatment field code	TXT 8	25-32	Code used for treatment in the field
Treatment type, list sorter	TXT 6 U	33-38	+=tested, F=factor average, @=printed
Treatment text	TXT 100	39-138	Treatment text/description

### [FD05]

This data group may be used to give details/specifications to each treatment factor and level where the main key may be either the series, location or trial identifier.

Field	Data type	Position from-to	Comment/use
Identifier (*)	TXT 12 U	1-12	Series, Locations or Trial identifier
Treatment number (*)	NMI 4	13-16	Treatment number from 1 and up
Treatment factor (*)	NMI 4	17-20	Factor (part) of the treatment combination
Amount specification	NUM -	21-36	Value specification
Unit specification	TXT 10	36-46	Unit/other specification
Text specification	TXT 100	47-146	Text detail

### [FD06]

Data group specifying the treatment randomization/distribution on the plots/experimental units. The trial and plot number are the identifiers and the treatment numbers corresponds to those in a treatment list [FD04]. The key data group is needed if the relations between trial and series identification is not know to the receiver of data.

Field	Data type	Position from-to	Comment/use
Trial (*)	TXT 12 U	1-12	Trial identifier
Plot number (*)	NMI 4	13-16	Plot (experimental unit) from 1 and up
Treatment number	NMI 4	17-20	Treatment number (as in [FD04])
Sub-group counter	NMI 4	21-24	Grouping counter for incomplete designs

### [FD07]

Data group for storing all kinds of general information for a trial. The information is identified by the trial identifier and a variable (property) code. Treatment details can be given in data group [FD05]

Field	Data type	Position from-to	Comment/use
Trial (*)	TXT 12 U	1-12	Trial identifier
Variable code (*)	TXT 12 U	13-24	Variable/property identifier
Date specification	DAT 10	25-34	Date of observation or general treatment
Growth stage	TXT 4	35-38	Decimal growth stage specification
Amount/Level	NUM -	39-54	Observed/applied amount/level
Text specification	TXT 30	55-84	Text for observation/general treatment

### [FD08]

Data group holding information about observed treatment averages and plot variables. This data group should always be included when one more of the data groups [FD09] to [FD11] appears. For the same trial, the same variable code should normally not exist in more than one of these data groups.

Field	Data type	Position from-to	Comment/use
Trial (*)	TXT 12 U	1-12	Trial identifier
Variable code (*)	TXT 12 U	13-24	Variable code identifier
Data origin	TXT 1 U	25	+ for original values, C for calculated
Data type	TXT 1 U	26	P=plot values, X/S=treatment values/text
Observation date	DAT 10	27-36	Date when the variable was observed
Observed area	NUM	37-52	Most common the net harvested plot, m2
Growth stage	TXT 4 U	53-56	Optional growth stage specification
Special code	NMI 4 -	57-61	Special code for relative numbers etc.

### [FD09]

Data group holding all treatment average data. If the receiver does not know the relation between the trial and series identifiers the data group [FD01] is needed. The treatment list and treatment distributions are stored in the [FD04] and [FD06] data groups. Data in this group should as a rule be combined with data group [FD08].

Field	Data type	Position from-to	Comment/use
Trial (*)	TXT 12 U	1-12	Trial identifier
Variable code (*)	TXT12 U	13-24	Variable code identifier
Treatment number (*)	NMI 4	25-28	Number as in [FD04]
Observed value	NUM -	29-44	Data value (only numeric values allowed)

### [FD10]

Data group holding all plot (experimental unit) data. Data in this group should as a rule be combined with data group [FD08].

Field	Data type	Position from-to	Comment/use
Trial (*)	TXT 12 U	1-12	Trial identifier
Variable code (*)	TXT12 U	13-24	Variable code identifier
Plot/unit number (*)	NMI 4	25-28	Number as in [FD06]
Observed value	NUM -	29-44	Data value (only numeric values allowed)

### [FD11]

This data group can hold optional text data as well as parameters obtained from statistical analysis. The stored parameters would appear in the corresponding data column in a result output. However, extra rows in data group [FD04] is needed to create output lines for such column information (for database applications).

Field	Data type	Position from-to	Comment/use
Trial (*)	TXT 12 U	1-12	Trial identifier
Text variable (*)	TXT 12 U	13-24	Dependent (Y) variable/column identifier
Treatment/line no. (*)	NMI 4	25-28	Treatment no. for text data or output line
Text or Statistics par.	TXT 8	29-36	Value/text, result from statistical analysis

### [FD12]

Text library for listing of all used registration variables (for general information as well as for treatment averages and plot values)

Field	Data type	Position from-to	Comment/use
Category (*)	TXT 12 U	1-12	Catgory/code group
Variable code (*)	TXT 12 U	13-24	Code identifying this variable
Text Description	TXT 60	25-84	Description (text)

### [FD13]

Text library for standard texts, variety codes, pesticide codes etc. Text are grouped in categories and numbered within these.

Field	Data type	Position from-to	Comment/use
Category (*)	TXT 3 U	1-3	Grouping key for text groups
Numerical code (*)	NML 8	4-11	Code within category
Standard text	TXT 36	12-47	Text description

### [FD14]

A data group for general text library information like texts for check lists, analysis forms, labels, growth stages, soil descriptions etc..

Field	Data type	Position from-to	Comment/use
Text group key (*)	TXT 12 U	1-12	Text category (group)
Text row (*)	TXT 4 U	13-16	Row within category
Text type	TXT 1 U	17	Code for the type of text
Text	TXT 88	18-105	Text line

### [FD15]

Data group with sets of calculation formulas, mainly for use in database applications. With the formulas, instruction can be given on how to make additional calculations on the included data .

Field	Data type	Position from-to	Comment/use
Formula key (*)	TXT 12 U	1-12	Formula group name
Formula row (*)	NML 4	13-16	Row within formula (sets calculation order)
Operation	TXT 1 U	17	Type of calculation (like * + - /)
Result variable	TXT 16 U	18-33	Variable code for calculation result
Input variable 1	TXT 16 U	34-49	First (only) input variable
Input variable 2	TXT 16 U	50-65	Second (optional) input variable
Parameter/constant 1	TXT 8 U	66-73	Calculation parameter/constant 1
Parameter/constant 2	TXT 8 U	74-81	Calculation parameter/constant 2
Text comment	TXT 30	82-111	Text comment about the operation

### [FD16]

Data group with specifications of data output profiles for statistics files or result reports. This group is mainly used in database applications for retrieving and presenting data.

Field	Data type	Position from-to	Comment/use
Profile key (*)	TXT 12 U	1-12	Profile group name
Profile column (*)	NML 4	13-16	Profile column order (from left)
Column properties	TXT 2 U	17-18	Codes for type/class of the column data
Column code	TXT 16 U	19-34	Variable or other code for fetching data
Decimal format	TXT 2 U	35-36	Number format (for data values only)
Text column header 1	TXT 8	37-44	Text 1 for data column in result reports
Text column header 2	TXT 8	45-52	Text 2 for data column in result reports
Text column header 3	TXT 8	53-60	Text 3 for data column in result reports
Text column header 4	TXT 8	61-68	Text 4 for data column in result reports
Text column header 5	TXT 8	69-76	Text 5 for data column in result reports
Scale factor	TXT 12	77-88	Value for re-scaling numerical data

## Examples of transport files

### 1. Text file with order of analysis at a laboratory

The file includes only three data groups. Group [FD12] is a variable description list and [FD08] a list of trials and variables to be analyzed. For one variable the group [FD09] contains a list of treatment numbers that should be analyzed. For the other variables the number of samples (in treatment or plot number order from 1 and up) was given in the special code field of the [FD08] data group. After analyzing the transport file would have been expanded by all data in [FD09] and [FD10] groups and the data in [FD08] would be updated with appropriate information (mainly the analyzing date).

#### File to laboratory:

```
0100 YYYY-MM-DD 46 59 0 1998-06-11 17:35:00 Torbjörn_Leuchovius
[FD12]
UK;MC;Moisture content, drying oven, %
UK;RE;Refuse amount, %
UK;KRW;Thousand kernel weight, grams
[FD08]
EXAMPLE;MC;+;P;;;12
EXAMPLE;KRW;+;X;;;2
EXAMPLE;RE;+;X;;;4
[FD09]
EXAMPLE;KRW;1;
EXAMPLE;KRW;3;
```

#### File from laboratory:

```
0100 YYYY-MM-DD 46 59 0 UK 28-08-1998 11:12:23 AgroLab_Eslöv
[FD12]
UK;MC;Moisture content, drying oven, %
UK;RE;Refuse amount, %
UK;KRW;Thousand kernel weight, grams
[FD08]
EXAMPLE;MC;+;P;1998-08-23;;;11      (one sample was lost, #8)
EXAMPLE;KRW;+;X;1998-08-25;;;2
EXAMPLE;RE;+;X;1998-08-25;;;4
[FD09]
EXAMPLE;KRW;1;55.3
EXAMPLE;KRW;3;52.8
EXAMPLE;RE;1;1.7
EXAMPLE;RE;2;16
EXAMPLE;RE;3;2.9
EXAMPLE;RE;4;2.6
[FD10]
EXAMPLE;MC;1;17.2
EXAMPLE;MC;2;14.2
EXAMPLE;MC;3;16.3
EXAMPLE;MC;4;20.7
EXAMPLE;MC;5;18.7
EXAMPLE;MC;6;17.9
EXAMPLE;MC;7;15.8
EXAMPLE;MC;9;16.7
EXAMPLE;MC;10;17.6
EXAMPLE;MC;11;16.8
EXAMPLE;MC;12;17.3
```

## 2. Transport file with treatments list & distribution and treatment/plot data

The file is prepared in Excel with each field in one column. All columns holds text data. Not all the content in the cells is visible. The symbol ...> means that there are more fields in this data group than shown here. This example contains only one observed variable in each of the two listed trials (to fit this page).

	A	B	C	D	E	F	G	H
1	0100	J	E	N	1	19980610	12:25:00	HS-BC
2	[FD01]							
3	HC0023	S7-450-95	BC-49-1995	1995	Malkorn	Lars Daniels..	2	...>
4	HC0024	S7-450-95	BC-50-1995	1995	Malkorn	Lars Daniels..	2	...>
5	[FD03]							
6	S7-450-95	SP	Sortförsök	CG Petters..	1115	1A	standard.rbl	...>
7	[FD02]							
8	BC-49-1995	Ambjörn Mat..	Övergran, Åby	mmh ML	Måttligt mullh..	30	591500	...>
9	BC-50-1995	CG Petters..	Kävlinge, Fjä..	mr Mo LL	Mullrik moig ...	15	593025	...>
10	[FD04]							
11	S7-450-95	1	109206	1A	+@	Obeh. maresi		
12	S7-450-95	2	108487	1B	+@	Obeh. SW 8..		
13	S7-450-95	3	109419	1C	+@	Obeh. Maud..		
14	S7-450-95	4	109101	1D	+@	Obeh. Baron..		
15	S7-450-95	5	209206	2A	+@	Beh. maresi		
16	S7-450-95	6	208487	2B	+@	Beh. SW 8..		
17	S7-450-95	7	209419	2C	+@	Beh. Maud..		
18	S7-450-95	8	209101	2D	+@	Beh. Baron..		
19	S7-450-95	9	9206	A	F@	Maresi.AN		
20	S7-450-95	10	8487	B	F@	SW 8487		
21	S7-450-95	11	9419	C	F@	Maud.SW		
22	S7-450-95	12	9101	D	F@	Baronesse.NS		
23	S7-450-95	13	100000	1	F@	Obehandlat		
24	S7-450-95	14	200000	2	F@	Behandlat		
25	[FD07]							
26	HC0023	CRP	0		0E0	Höstraps		
27	HC0024	CRP	0		0E0	Höstvete		
28	[FD06]							
29	HC0023	1	7	0				
30	HC0023	2	5	0				
	.....							
44	HC0023	16	3	0				
45	HC0024	1	2	0				
46	HC0024	2	4	0				
	.....							
60	HC0024	16	3	0				
61	[FD08]							
62	HC0023	MC	+	X	729636			
65	HC0024	Y.RD	+	P	729638	24200000E-6		
66	[FD09]							
67	HC0023	MC	1	17300000E-6				
68	HC0023	MC	2	16500000E-6				
	.....							
73	HC0023	MC	7	15800000E-6				
74	HC0023	MC	8	18300000E-6				
75	[FD10]							
76	HC0024	Y.RD	1	65376000E-6				
77	HC0024	Y.RD	2	61400000E-6				
	.....							
90	HC0024	Y.RD	15	58450000E-6				
91	HC0024	Y.RD	16	59330000E-6				
92								
93								