

SDSU Department of Geography
Interdisciplinary Knowledge Synthesis (IKS) Group

SEKS-GUI v.0.6 User Manual



May 2006

CONTENTS

Preface.....	1
PART I	
INSTALLATION OF THE SEKS-GUI PACKAGE.....	2
1. System Requirements.....	2
2. Installation Notes.....	2
3. Testing BMElib.....	4
PART II	
INTRODUCTION TO SEKS-GUI.....	5
<i>Section A</i>	
The SEKS-GUI Structure.....	5
BME or GBME?.....	6
<i>Section B</i>	
What information is needed in order to run the SEKS-GUI.....	7
1. Hard data file.....	7
2. Soft data file.....	8
3. Output format file.....	12
<i>Section C</i>	
Starting a SEKS-GUI session.....	14
PART III	
SEKS-GUI SCREENS.....	15
<i>Section A</i>	
Common screens: Input for all types of investigations.....	15
1. Splash Screen.....	15
2. "Choose a Task".....	16
3. "Hard Data Wizard" – Part I.....	18
4. "Hard Data Wizard" – Part II.....	20
5. "Soft Data Wizard" – Part I.....	21
6. "Soft Data Wizard" – Part I (continued).....	22
7. "Soft Data Wizard" – Part II.....	23
8. "Output Configuration".....	24
9. "Exploratory Analysis" – Part I.....	26
<i>Section B</i>	
BME Analysis Screens in SEKS-GUI.....	27
1. BME "Exploratory Analysis" – Part II.....	27
Introduction.....	27

Instructions.....	27
2. BME “Exploratory Analysis” – Part III.....	33
Introduction.....	33
Instructions.....	34
3. BME “Covariance Analysis” – Part I.....	37
Introduction.....	37
Instructions.....	38
4. BME “Covariance Analysis” – Part II.....	42
5. “BME Estimations Wizard”.....	47
<i>Section C</i>	
GBME Analysis Screens in SEKS-GUI.....	52
1. GBME “Exploratory Analysis” – Parts II and III.....	52
Introduction.....	52
Instructions.....	52
2. “GBME Estimations Wizard”.....	55
Introduction.....	55
Instructions.....	56
<i>Section D</i>	
Visualizations with SEKS-GUI.....	61
1. “Visualization Wizard”.....	61
PART IV	
SEKS-GUI EXAMPLES INCLUDED IN PACKAGE.....	70
1. BME and GBME S/T study of Total Ozone concentrations over the United States.....	70
BME Analysis.....	70
GBME Analysis.....	71
2. BME and GBME Spatial study of Arsenic in Bangladesh drinking water.....	72
BME Analysis.....	73
GBME Analysis.....	74
PART V	
SEKS-GUI UTILITIES INCLUDED IN PACKAGE.....	75
1. Data exchange with shapefiles: Converting shapefiles into text.....	75
2. Data exchange with shapefiles: Converting text into shapefiles.....	78
3. A start-up set of files for the creation of masking files.....	81
PART VI	
BIBLIOGRAPHY.....	83

PREFACE

Welcome to the Spatiotemporal Epistemic Knowledge Synthesis Graphical User Interface (GUI), or **SEKS-GUI**. This package is currently the combination of the scientific software libraries BMElib (Bayesian Maximum Entropy library) and GBMElib (Generalized BME library) and the GUI related files. Forthcoming versions of the SEKS-GUI will include software libraries based on stochastic logic theory and techniques (e.g., material conditional and biconditional).

BMElib and **GBMElib** are stand-alone software libraries for space-time modeling, prediction (estimation) and mapping, which introduce many new features and implement innovative approaches of Temporal Geographical Information Science (TGIS; see, Christakos *et al.*, 2002). These libraries differ in that they have different requirements and employ different concepts and techniques of space-time modeling and mapping. In particular, BMElib processes detrended, normally distributed data sets; it provides explicit information on estimates of the underlying data trends; the user models data correlations by fitting permissible (ordinary) covariance models to the raw data. The GBMElib is especially applicable in the study of heterogeneous space-time dependences and complex patterns (e.g., geographical non-homogeneity and temporal non-stationarity); it allows a less involved exploratory data analysis of the raw data than BME; it involves the automatic calculation of heterogeneity orders and (generalized) covariances within local space-time neighbors.

Both the BMElib and GBMElib software libraries currently run on Matlab. Therefore, for the use of each one of them as a stand-alone package it is required that the user has a working knowledge of the Matlab command line. SEKS-GUI is built with the user in mind. It provides a friendly interface for these libraries, eliminates the need to know programming in order to use them, and provides a unified framework of all the individual steps of space-time modeling, prediction and mapping. This is the manual guide for the latest version of the GUI (SEKS-GUI Version 0.6).

For an initial familiarization or a “fast start” with SEKS-GUI, two examples are presented at the end of the manual and are included in the SEKS-GUI package. We hope that the users will find this manual guide helpful when using the SEKS-GUI package in their various projects. Our goal is to deliver updated GUI versions that will allow more flexibility and feature-rich access to the most recent theoretical approaches in interdisciplinary knowledge synthesis, in general, and TGIS, in particular. We will greatly appreciate your feedback on the ideas, features, functionality and aesthetics of the SEKS-GUI.

IKS Group, SDSU-Dept of Geography
Director: Dr. George Christakos (gchrista@mail.sdsu.edu)
Contact (SEKS-GUI): Dr. Hwa-Lung Yu (hlyu@mail.sdsu.edu)

INSTALLATION OF SEKS-GUI PACKAGE (BMELib SCIENTIFIC SOFTWARE & SEKS-GUI)

1. System Requirements

1. Personal computer running Windows OS, Mac OS X, or Intel Linux with Matlab pre-installed. The main requirement is that the Matlab version used is v. 6.5 (Release 13 of June 2002) or more recent. BMELIB version 2.0b will run on versions as early as Matlab 6.1, but more recent ones are necessary for the SEKS-GUI. Note that SEKS-GUI will not run with versions earlier than Matlab 6.5.

⇒ If you are unsure about the Matlab version you are using, you can check this software version by launching Matlab (see subsection 2.3 below). Then type on the command line:

```
>> version
```

and press the return key. The number you see should be 6.5 or higher.

⇒ Additional information: Key functions in BMELib and GBMELib have been compiled with a Matlab-approved Fortran compiler on Windows OS, with GCC v.3.4.3 on Mac OS X.4.6, and with GCC v.3.3.4 on Intel Linux Suse 9.2.

2. Installation notes

For installation of the SEKS-GUI package you do not need administrative rights on your computer. Simply follow the steps below to move the package files into a desired position in your documents.

1. Select a folder where the material will reside.

Example: You can create a folder “SEKSGUIfolder” to store the package. This can be: “C:\[Your_WindowsOS_path_to]\My Documents\SEKSGUIfolder” in Windows, or: “[Your_MacOS_path_to_your]\Documents\SEKSGUIfolder” under Mac OS.

In this manual we assume that the package will reside in “SEKSGUIfolder”.

2. Move “SEKS-GUI0.6package.zip” into the “SEKSGUIfolder”.

3. Unzip the file. The contents are the following items:

- i) “SEKS-GUI0.6”: Folder that contains all the graphical user interface (GUI) files.
- ii) “SEKS-GUI0.6manual.pdf”: The current manual.
- iii) “startup.m”: A file with information so that the Matlab application can locate files.
- iv) “guiExamples”: Folder containing GUI examples.

Move all the above items directly into the SEKSGUIfolder, in case the uncompressing application you use has put them under a new additional folder by default. Upon successful unzipping of the above items and having them under the “SEKSGUIfolder” you may delete, if you want, any additional folder created and the “SEKS-GUI0.6package.zip” file.

4. Start Matlab and navigate its current working folder to the “SEKSGUIfolder” using the bar on the top of the Matlab Command window: Use the buttons on the top right-hand side to locate the desired folder within your filesystem (Fig. I.1).

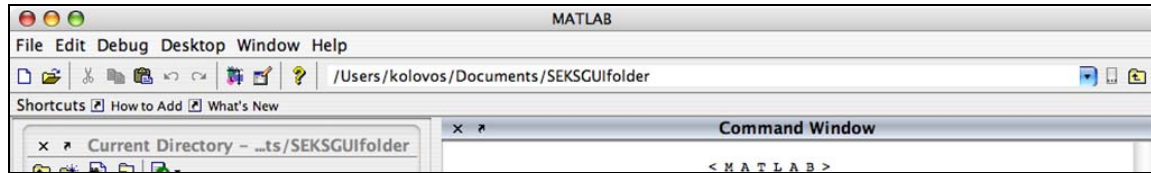


Fig. I.1

Alternatively, you can perform the navigation using the Current Directory window (Fig. I.2) to the left of the Matlab Command window (the main window when you launch Matlab).

If the Current Directory window is not showing, you can find it and select it under the “Desktop” menu of the Matlab main window, and it should then appear on the side of the main window.

⇒ If this process has been performed once on a computer, you may find that next time you launch Matlab on the same computer you can go directly to the SEKSGUIfolder (or any other previously visited folder) by clicking on the downward arrow on the side of the Current Directory bar (Fig. I.1) at the top of the Matlab Command window. This action will display a list of the most recently visited folders, so you could select the SEKSGUIfolder without having to navigate your way there again.

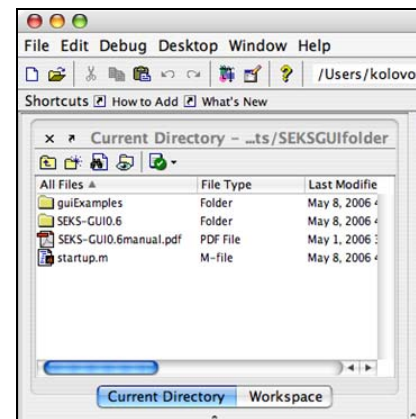


Fig. I.2

5. Once Matlab indicates above the Command window that you are in the “SEKSGUIfolder”, type “startup” at the Matlab command line and then press the Enter or the Return key.

```
>> startup
```

This will render all of the material in the SEKS-GUI folders accessible in your session by adding them into the Matlab search path. A confirmation message should display in the Matlab Command window (Fig. I.3).

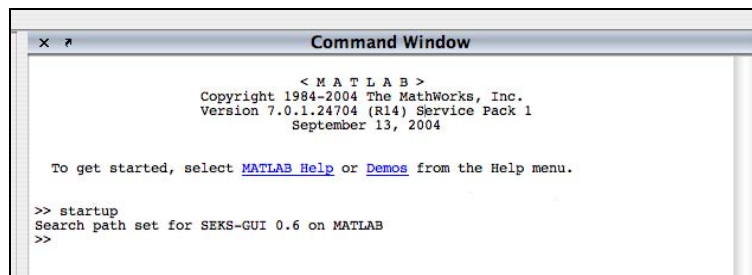


Fig. I.3

You are now ready to use BMElib and GBMElib on Matlab.

⇒ Steps 4 and 5 above must be performed every time you start Matlab, if you want to use SEKS-GUI (and BMElib or GBMElib) in a session.

★ Step 5 above may be skipped if you start Matlab from within the “SEKSGUIfolder”. This option is available for Windows OS users if you define a Matlab shortcut that starts from within this folder when invoked. This option is also available for Mac OS and Linux users, if you start Matlab from the command line of a terminal while in the “SEKSGUIfolder”. This information is slightly more advanced; for all good purposes you can follow step 5 to start SEKS-GUI.

3. Testing BMElib

This is not a necessary step in the SEKS-GUI installation. It is a recommended action, though, the first time you ever run BMElib on a computer to ensure that important BMElib functions work properly on the particular hardware.

1. Start Matlab. The Matlab Command Window should appear. Make sure that Matlab already has the SEKS-GUI package in its path (steps 4 and 5 in previous paragraph 2).

2. Type “MVNLIBtest” in the command line of the Matlab Command Window and then press the Enter or the Return key.

```
>> MVNLIBtest
```

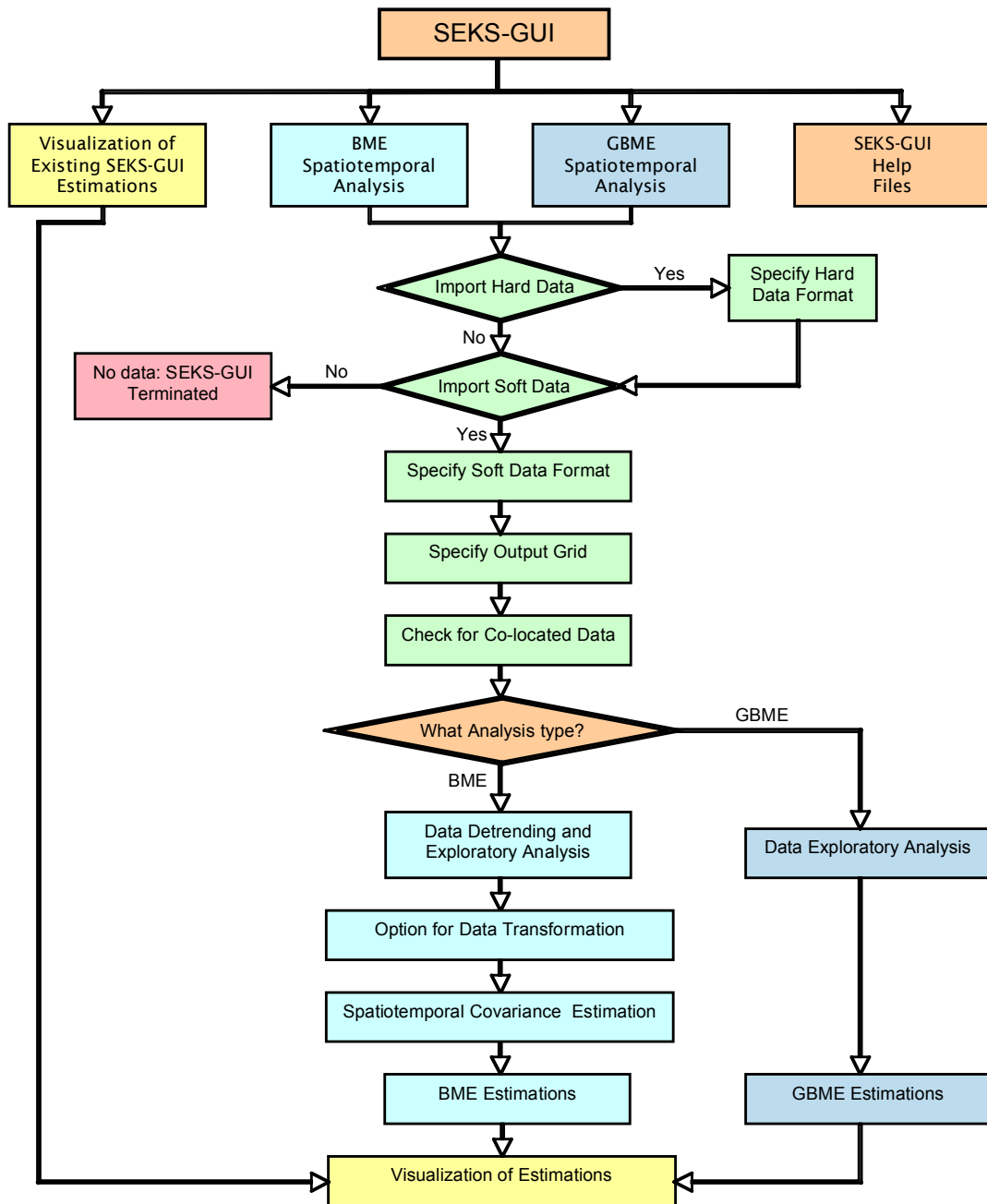
Wait for 1-2 seconds for some testing calculations to appear on the Matlab Command Window. You should see the message “test complete” at the bottom of the window soon afterwards. This signifies that BMElib is fully functional on your computer.

⇒ If any errors should appear, this is evidence that some key BMElib functions are not properly compiled for your computer architecture/software. In such a case please report the error by contacting/sending an email to akolovos@mail.sdsu.edu. Carefully describe the matter and your operating system details, so that the problem can be examined.

INTRODUCTION TO SEKS-GUI

Section A – The SEKS-GUI Structure

SEKS-GUI makes use of BME and GBME spatiotemporal analysis. The flowchart is an overview of the procedures that take place in SEKS-GUI for both types of analyses.



BME or GBME?

BME and GBME share the same methodological framework of Knowledge Synthesis. However, they exhibit significant differences in the underlying theory and practical implementation, which should be taken into account by the user when choosing between them.

The two techniques may use distinct modeling assumptions and emphasize different general knowledge bases, as exhibited in the Table II.A.1 below.

	BME	GBME
Heterogeneity	Decomposition of original attribute into a mean function and a residual attribute (separate, explicit steps)	Calculation of space-time increments in original attribute distribution (steps incorporated in estimation process)
Space-Time dependence structure	Not available	Calculation of heterogeneity orders ν/μ
Type of correlation analysis	Performs on a global scale approach; scale depends on the user choice of the output grid	Local space/time scale
Covariance functions	Ordinary covariance functions (must be positive-definite)	Generalized covariance functions (need only to be conditionally positive-definite)
Data requirements	Minimum data required throughout estimation grid (no explicit number). May generate estimates even if data not present locally	Minimum data required in each estimation node neighborhood. Best performance with as many possible, uniformly distributed data

Table II.A.1

Therefore, under the conditions applicable to a particular investigation, the user may consider the implementation of either of these methods or the one most suitable. Alternatively, a combination of BME and GBME can be feasible by means of applying the BME technique in some sub-regions of the study area and the GBME technique in some others. Finally, the two different techniques may be both applicable, and the user may wish to derive the outcome from each of them separately for a comparative study.

Section B – What information is needed in order to run the SEKS-GUI

- ⇒ The following information describes what you will need to have at hand before you begin working with the SEKS-GUI. For your convenience, please read the current section before running the software, and prepare your material as described in the following guidelines.
- ⇒ If you only wish to test-run the program, there exist built-in, full-scale examples in the SEKS-GUI package, so you do not need to have any additional files or information at hand. The locations of the example files you will be asked for during execution of the SEKS-GUI are provided in PART IV towards the end of this manual.

The SEKS-GUI enhances and facilitates your space/time analyses by allowing you to use the powerful BMElib and GBMElib software libraries in a user-friendly manner. Your investigation input is necessary to run the GUI. You can make optimal use of the GUI by having the requested input information available at hand. Depending on your available data and study types, 3 files containing information will be asked at different times:

1. A hard data file: Based on the Knowledge Synthesis framework, hard data are exact measurements (i.e., data without significant uncertainty in their values for the particular case study). This file should be containing coordinates and data value provided in one of the following forms: ASCII (text) file, Microsoft Excel (.xls), or GeoEAS – formatted text file. On the screen where hard data are provided by the user (see following sections III.A.3 and III.A.4) it will be required that the user specifies the type of data file as one of the above. The following notes and file formatting rules must be observed:

- I. If you choose to use a text file, values should be space- or tab-separated, occupy continuous lines, and there should be 1 datum information per line. Each datum line should display up to 3 spatiotemporal coordinates (up to 3 spatial coordinates in the purely spatial case, or up to 2 spatial coordinates plus 1 temporal in the spatiotemporal case) followed by the datum value. The display in each line must be consistent with those of the other lines in the same file (e.g., in each line the first number must stand for the same entity). Finally, you will need to know what each of the columns in your text file stands for (e.g., x-coordinates are in the 1st column).
- II. In case you use an Excel spreadsheet, please ensure your information is in the first spreadsheet of the file (if there are more than one spreadsheets in the file). Values should be stored in neighboring cells, and each row should contain information for exactly 1 datum. The formatting rules are the same as in the previous case (I).
- III. If you will be using the GeoEAS format, then all you need to know is where your input file resides when asked, and how the data are positioned within the file. That is, you will need to know the columns where the spatiotemporal coordinates and the variable values are stored.

2. A soft data file: Based on the Knowledge Synthesis framework, soft data are measurements associated with some known degree of uncertainty. This file should be provided in the form of an ASCII (text) file or in a Microsoft Excel file (.xls).

- I. In a text file, values should be space- or tab-separated, occupy continuous lines, and there should be 1 datum information per line. Each datum line should be formatted by first displaying up to 3 spatiotemporal coordinates (up to 3 spatial coordinates in the purely spatial case, or up to 2 spatial coordinates plus 1 temporal in the spatiotemporal case), and the datum value at the end. The contents in each line must be consistent with those of the other lines in the same file (e.g., in each line the first number must stand for the same attribute, etc.) Finally, you will need to know what each of the columns in your text file stands for (e.g., x-coordinates are in the 1st column, etc.)
- II. In case you use an Excel spreadsheet, please ensure your information is in the first spreadsheet of the file (if there are more than one spreadsheets in the file). Values should be stored in neighboring cells, and each row should contain information for exactly 1 datum. The formatting rules are the same as in the previous case (I).

When it comes to assimilating soft information, SEKS-GUI is a powerful tool that can accept and process a variety of types/categories of soft data.

- ⇒ Only one type of soft data may be used in the same investigation, i.e., different types cannot be mixed in the same soft data file!
- ⇒ If your study involves more than one soft data types you can convert them to the same single soft data type and store them in one soft data file for use in the investigation. In particular, types shown in the following in (a) and (b1) can be converted, e.g., into the soft data types exhibited in Examples 5 and 7 shown later in (b2) – interval soft data are essentially uniform distributions. Given any of the fully described distributions presented in (a) or (b1), the user can estimate in advance the distribution probability density $p_{1A}, p_{2A}, \dots, p_{NA}$ at a set of N values $l_{1A}, l_{2A}, \dots, l_{NA}$ of the quantity of interest. In this way, following the guidelines of either of the Examples 5 or 7 below, any soft data collection of mixed types can be stored in the same file for use in an investigation.

Each of the following examples describes one of the different soft data types that can be used with SEKS-GUI. These include:

(a) **Interval soft data:** Information is provided in terms of the lower and upper bounds of each datum, whose values may range uniformly anywhere in the given interval. Each line in the interval soft data file must contain the datum coordinates, followed by the lower interval bound, and then followed by the upper interval bound.

Example: “ $x_A \ y_A \ z_A \ l_A \ u_A$ ” can be the line describing one datum A , which is located on a plane at (x_A, y_A) at the temporal instance z_A [the same line also describes a soft datum at (x_A, y_A, z_A) in a purely spatial 3-D case], and its value may be measured with equal probability to be anywhere within the interval $[l_A, u_A]$.

(b) **Probabilistic soft data:** Each datum is provided in the form of a probability density function (PDF) describing its probability distribution across all the valid values of the quantity of interest. The PDF may be either fully known by virtue of its characteristics, or it may be described by the user in terms of individual probabilities assigned to a pre-defined number of bins across the PDF span – there are several different soft data types that fall in this category. Each line in the interval soft data file must contain the datum coordinates, followed by the appropriate datum PDF description. The examples in the following pages cover the spectrum of the acceptable probabilistic soft data forms in SEKS-GUI.

b1. *Probabilistic soft data with fully described PDF characteristics*: SEKS-GUI accepts soft data in the form of gaussian, uniform or triangular distributions.

Example 1: Case where PDF is a gaussian distribution with known mean and variance.

Let us assume one such datum A that is a gaussian distribution $N(m_A, v_A)$ with mean m_A and variance v_A , and is located at coordinates (x_A, y_A) in space and at z_A in time, as portrayed in Fig. II.A.1. The description of this datum in the soft data file should look like this (the quotes not to be included in the file; Excel files should feature these values in consecutive cells in the same row):

`"xA yA zA mA vA"`

Note that the same line also describes a soft datum at (x_A, y_A, z_A) in a 3-D purely spatial case.

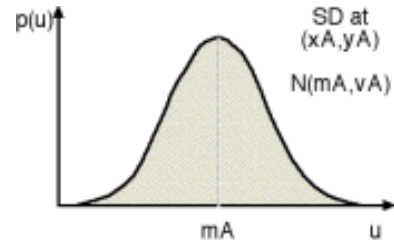


Fig. II.A.1

Example 2: Case where PDF is a uniform distribution with known mean and variance.

Each datum is described in 1 line that should look like this (the quotes not to be included in the file; Excel files should feature these values in consecutive cells in the same row):

`"xA yA mA vA"`

For the current soft data type, the above line defines one probabilistic datum A portrayed in Fig. II.A.2, which is located at coordinates (x_A, y_A) in a purely spatial 2-D case. The description line also contains the information that the datum uniform distribution $U(m_A, v_A)$ has mean m_A and variance v_A .

Note that the same line above also describes a soft datum at x_A , at the temporal instance y_A in a 1-D spatial and temporal case.

Also, note that uniform distribution soft data are very similar to soft interval data of the category (a). Indeed, soft intervals as defined are uniform distributions for which their upper and lower limits are known, rather than their means and variances.

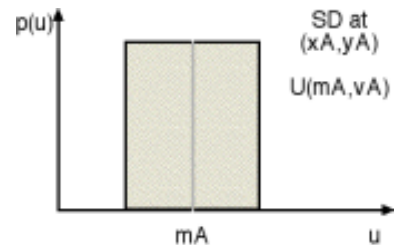


Fig. II.A.2

Example 3: Case where PDF is a triangular distribution with known mean and limits.

Each datum is then described in 1 line that should look like this (the quotes not to be included in the file; Excel files should feature these values in consecutive cells in the same row):

`"xA yA zA u1A mA u2A"`

For the current soft data type, the above line describes one probabilistic datum A portrayed in Fig. II.A.3, which is located at coordinates (x_A, y_A) in space and at z_A in time. Its PDF is a triangular distribution $T(u_{1A}, m_A, u_{2A})$ that spans within the interval $[u_{1A}, u_{2A}]$ and has a mean m_A .

Note that the same line also describes a soft datum at (x_A, y_A, z_A) in a 3-D purely spatial case.

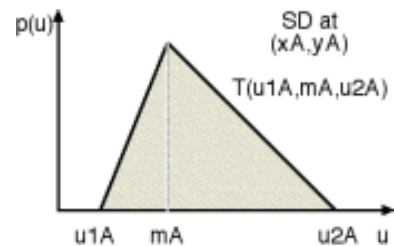


Fig. II.A.3

b2. *Probabilistic soft data with user-described PDF* in terms of probabilities at pre-defined bins. This is a useful alternative in the events the soft datum PDF is some other than the above supported shapes, when the user has information in bins or range brackets, or otherwise when the available uncertain information exhibits a more complex behaviour or PDF form.

Example 4: Case where PDF bins do not necessarily have the same size, and the PDF value is constant within a given bin. A simple distribution of this type is portrayed in Fig. II.A.4. Each datum is described in 1 line, and the corresponding entry in the file for the Fig. II.A.4 soft datum should look like this (the quotes not to be included in the file; Excel files should feature these values in consecutive cells in the same row):

`"xA yA zA 3 l1A l2A l3A l4A p1A p2A p3A"`

In the above line, a probabilistic datum A is located at (x_A, y_A) on a plane at the temporal instance z_A in the spatiotemporal case. Its PDF has 3 bins, whose 4 limits are l_{1A} l_{2A} l_{3A} and l_{4A} . Within each of the bins the datum PDF has the constant values p_{1A} p_{2A} and p_{3A} .

When using data of this type, the user may include in the same file data that have variable number of bins, and consequently different number of parameters per line entry. Note that the same line above also describes a soft datum of the same type at (x_A, y_A, z_A) in a purely spatial 3-D case.

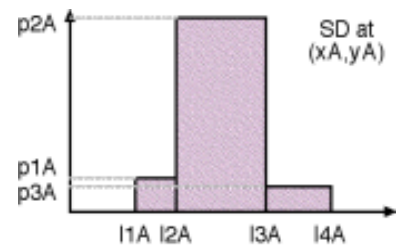


Fig. II.A.4

Example 5: Case where PDF bins do not necessarily have the same size and the PDF value changes linearly within a bin. Each datum is described in 1 line. Data would typically look like the simple distribution in Fig. II.A.5.

In this figure, we assume a probabilistic datum A located at (x_A, y_A) on a plane at the temporal instance z_A . Its PDF has 3 bins, whose 4 limits are l_{1A} l_{2A} l_{3A} and l_{4A} . Within each of the bins the datum PDF changes linearly to the values p_{1A} p_{2A} p_{3A} and p_{4A} at each of the bin limits, respectively.

The line corresponding to that datum in the soft input file should look like this (the quotes not to be included in the file; Excel files should feature these values in consecutive cells in the same row):

`"xA yA zA 3 l1A l2A l3A l4A p1A p2A p3A p4A"`

When using data of this type, the user may include in the same file data that have variable number of bins, and consequently different number of parameters per line entry. Note that the above line also describes a soft datum of the same type at (x_A, y_A, z_A) in a purely spatial 3-D case.

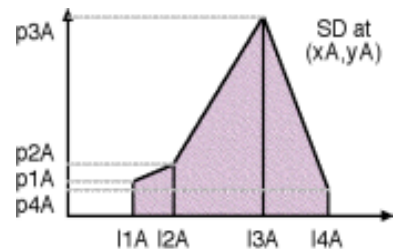


Fig. II.A.5

Example 6: Case where PDF bins have the same size and the PDF value is constant within a given bin. Each datum is described in 1 line. Data would typically look like the simple distribution in Fig. II.A.6.

In this figure, we assume a probabilistic datum A located at (x_A, y_A) on a plane at the temporal instance z_A . Its PDF has 3 bins, whose lower limit is l_{1A} , its upper limit is l_{2A} , and the distance between any two consecutive bins is ds_A . Within each of the 3 bins the datum PDF has the constant values p_{1A} , p_{2A} and p_{3A} , respectively. The line corresponding to that datum in the soft input file should like this (the quotes not to be included in the file; Excel files should feature these values in consecutive cells in the same row):

`"xA yA zA 3 l1A dsA l2A p1A p2A p3A"`

When using data of this type, the user may include in the same file data that have variable number of bins, and consequently different number of parameters per line entry. Note that the above line also describes a soft datum of the same type at (x_A, y_A, z_A) in a purely spatial 3-D case.

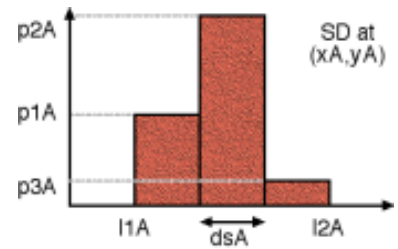


Fig. II.A.6

Example 7: Case where PDF bins PDF bins have the same size and the PDF value changes linearly within a bin. Each datum is described in 1 line. Data would typically look like the simple distribution in Fig. II.A.7.

In this figure, we assume a probabilistic datum A located at (x_A, y_A) on a plane at the temporal instance z_A . Its PDF has 3 bins, whose lower limit is l_{1A} , its upper limit is l_{2A} , and the distance between consecutive bins is ds_A . Starting from the lower limit, the datum PDF changes linearly from the value p_{1A} to the values p_{2A} , p_{3A} and p_{4A} at each of the following bin limits. The line corresponding to that datum in the soft input file should like this (the quotes not to be included in the file; Excel files should feature these values in consecutive cells in the same row):

`"xA yA zA 3 l1A dsA l2A p1A p2A p3A p4A"`

When using data of this type, the user may include in the same file data that have variable number of bins, and consequently different number of parameters per line entry. Note that the above line also describes a soft datum of the same type at (x_A, y_A, z_A) in a purely spatial 3-D case.

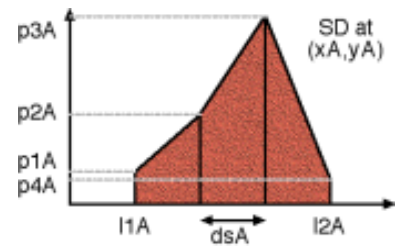


Fig. II.A.7

- ⇒ Note that in the above examples 4-7 the values used to describe the probabilistic nature of soft data are the probability density function values at the corresponding bins – **not** the actual probabilities for the bins. The probabilities are geometrically represented by the area under the bins, instead.
- ⇒ The total area under a PDF must be equal to the total probability for the occurrence of each datum, i.e., equal to 1. If this is not the case with any of the user-provided data, then the user is warned about any discrepancies. SEKS-GUI renormalizes the data in question to comply with the rule, and continues the investigation. However, it lies in the user's discretion to provide meaningful data, in the sense that SEKS-GUI cannot tell whether any such discrepancies are mistakes in the data entries.

3. An output format file that contains information about the estimation grid: Once the data are provided, the user must specify the nodes in space/time (as applicable) where SEKS-GUI will generate estimates. The estimation area definition is up to the user's needs and discretion. The distance between the nodes depends on the scale on which the user wants the results. The same is true regarding the number of nodes considered in each direction. Typically, one would want the output area to extend close to the spatiotemporal extent of the available data.

- ⇒ Too many nodes inbetween data locations may not necessarily reveal new information regarding the behavior of the field under study, and may consume computational resources unnecessarily.
- ⇒ Too many nodes extending further away from the measurements population may not even yield any results on them at all, if these locations are too distant from the data to be correlated with them.
- ⇒ This version of SEKS-GUI accepts orthogonal, regular grids as output grids. There may well be cases where your area of interest does not cover an orthogonal area similar to the output grid type, as, e.g., when data are only provided on a stretch whose spatial coordinates cross the designated output grid diagonally. It is likely then, based on the previous remark, that some nodes on the grid located further outside the data populated stretch may not be assigned an estimate if they cannot be correlated to the input information. This is an expected event and it should not be alarming: You can still make good use of the grid, and obtain estimates on as many nodes as this field's correlation will allow on the grid. Nodes without estimates can be masked out of the map at a later stage using image manipulation software.

This version of SEKS-GUI supports information for a custom orthogonal grid, whose characteristics should be provided by the user in an ASCII text file or a Microsoft Excel (.xls) spreadsheet.

- I. If you choose to use a text file, values in the file should be space- or tab-separated, occupy continuous lines, and there should be exactly 1 line dedicated to the information on each of the dimensions used.
- II. In case you use an Excel spreadsheet, please ensure your information is in the first spreadsheet of the file (if there are more than one spreadsheets in the file). Values should be stored in neighboring cells, and there should be exactly 1 row dedicated to the information on each of the dimensions used.

The following page has details on defining the grid information within your input file.

The grid information in a file can be only one of the following types:

- (a) Grid limits and spacing between nodes in each dimension.
- (b) Grid limits and number of nodes in each dimension.
- (c) Lower grid limit (origin), number of nodes and node spacing in each dimension.

Illustrations of the above types in the case of an imaginary spatiotemporal grid are shown in Fig. II.A.8. This grid features $n_x=4$ nodes in the x-axis, $n_y=3$ nodes in the y-axis and $n_t=2$ nodes in the temporal axis (or the vertical axis, if this is a purely spatial 3-D case), as well as the characteristics shown in the figure, i.e., the grid ranges from x_{min} to x_{max} in the x-direction, from y_{min} to y_{max} in the y-direction, and from t_{min} to t_{max} in the temporal continuum (or in the vertical direction in a purely spatial 3-D case), and yielding estimates at nodes spaced dx , dy , and dz , in each direction, respectively.

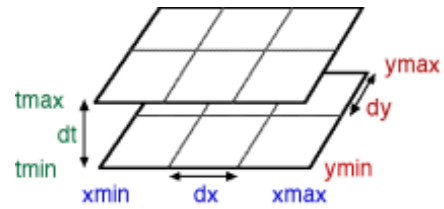


Fig. II.A.8

Example 1: In case it is convenient to provide output grid information of a type (a), the data file should look like this (the quotes not to be included in the file):

```
"xmin dx xmax  
ymin dy ymax  
tmin dt tmax"
```

Example 2: In case it is convenient to provide output grid information of a type (b), the data file should look like this (the quotes not to be included in the file):

```
"xmin xmax nx  
ymin ymax ny  
tmin tmax nt"
```

Example 3: In case it is convenient to provide output grid information of a type (c), the data file should look like this (the quotes not to be included in the file):

```
"xmin nx dx  
ymin ny dy  
tmin nt dt"
```

If using an Excel file for the input, for each of the above lines the file should feature the line values in consecutive cells in the same row.

Section C – Starting a SEKS-GUI session

1. Start Matlab – the Matlab Command Window should appear. Make sure Matlab already has the BMElib and SEKS-GUI packages in its path (perform steps 4 and 5 in PART I, Paragraph 2).

2. Once you know Matlab has the BME packages in its path you can navigate Matlab elsewhere and still run the BME software from any other folder. Let us assume that you choose to remain within the “SEKSGUIfolder”.

3. You can start a SEKS-GUI session by typing “seksgui” in the command line of the Matlab Command Window as below (Fig. II.B.1), and then press the Enter or the Return key.

```
>> seksgui
```

The SEKS-GUI splash screen will automatically appear, thus starting a new SEKS-GUI session (see paragraph III.A.1 in the following).

⇒ All steps 1-3 in this paragraph must be performed every time you start Matlab, if you want to use SEKS-GUI in a session.

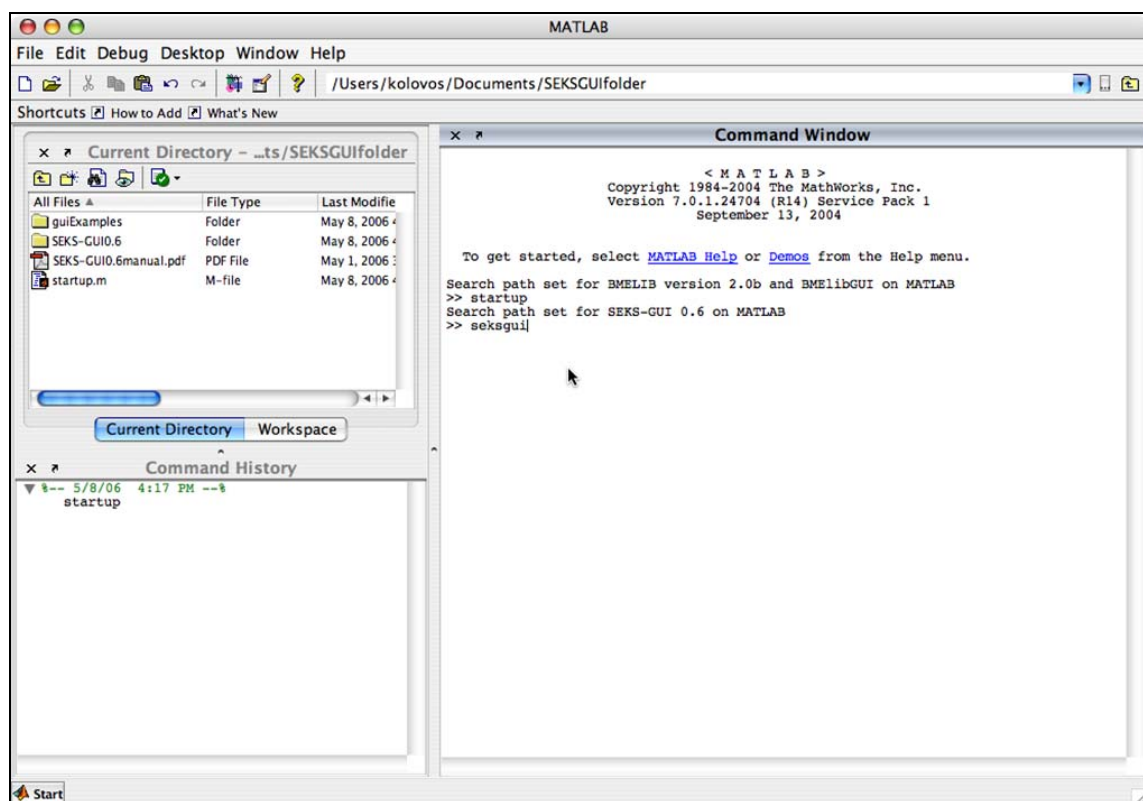


Fig. II.C.1

4. In order to exit SEKS-GUI at any instance you can push the “Main Menu” button in your screen. You will need to confirm the action, as this will result in erasing all your data stored in the memory of that session (data saved in files will not be affected), and then push the “Exit” button in the Main Menu screen (see paragraph III.A.2 in the following).

SEKS-GUI SCREENS

Section A – Common Screens: Input for all Types of Investigations

1. Splash Screen

Some general information is displayed (Fig. III.A.1). The screen shows for 4 seconds, and then it automatically closes. At this time, the program proceeds to Screen 2 so that the user may choose a task.

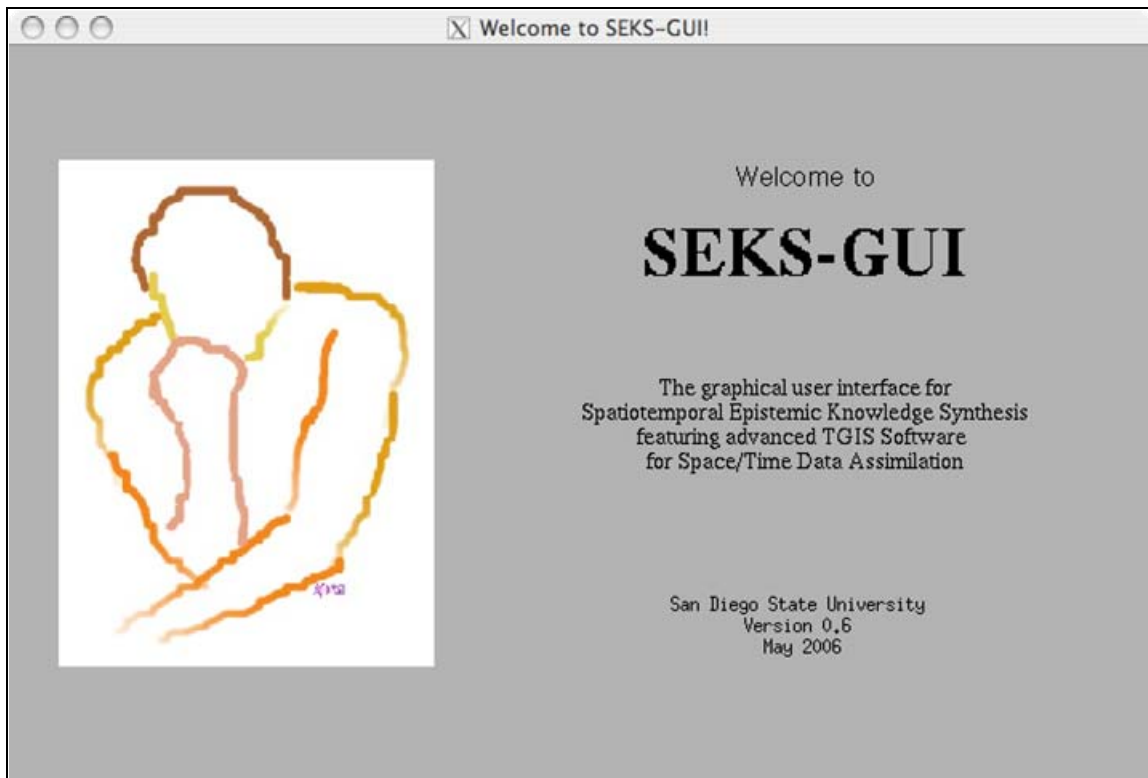


Fig. III.A.1

2. Screen 2: “Choose a Task”

You are presented with a list of available tasks. Once you make a choice by clicking on a line in the list, the line will be highlighted. Then push the “Start” button to begin (Fig. III.A.2).

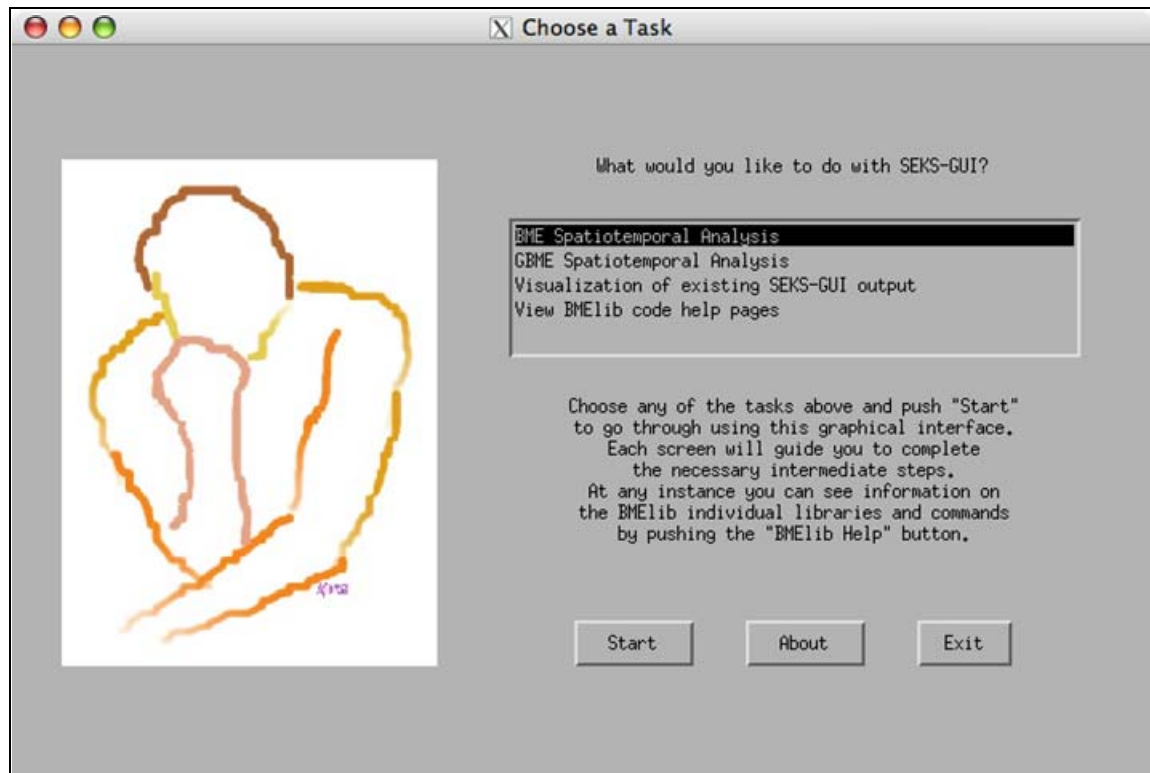


Fig. III.A.2

“BME Spatiotemporal Analysis” uses the BME methodology. You can read information on the BME analysis and instructions on the BME-related screens further later in Section B of PART III.

The “GBME Spatiotemporal Analysis” option will allow you to take advantage of the powerful v/μ -Spatiotemporal Random Field theory to perform your analysis, as an alternative way to using BME. More information on GBME and detailed instructions for the dedicated GBME screens are presented in Section C of PART III.

The “Visualization of existing SEKS-GUI output” option will take you to “Visualization Wizard”, the last of the SEKS-GUI screens. If you have estimation results from previous analyses using SEKS-GUI you can follow this path to reproduce maps of your results without having to go through all the SEKS-GUI screens inbetween. You can find a description of this screen’s features in Section D of PART III.

The “View BMElib code help pages” task provides access to the standard BMElib help pages. When the option is highlighted and “Start” is pushed, an external window is launched (Fig. III.A.3). This option displays the help available for the lists of the BMElib function libraries (upper left hand window in Fig. III.A.3) and individual functions (upper right hand window in Fig. III.A.3). That help shows when a particular name is typed in the Matlab Command Window using the Matlab “help” command. All of these help pages are now easily accessible through the GUI. The result for each request (which is the Matlab help output) is displayed in the lower window in Fig. III.A.3. When done, push “Close Help” to close the external BMElib code help window.

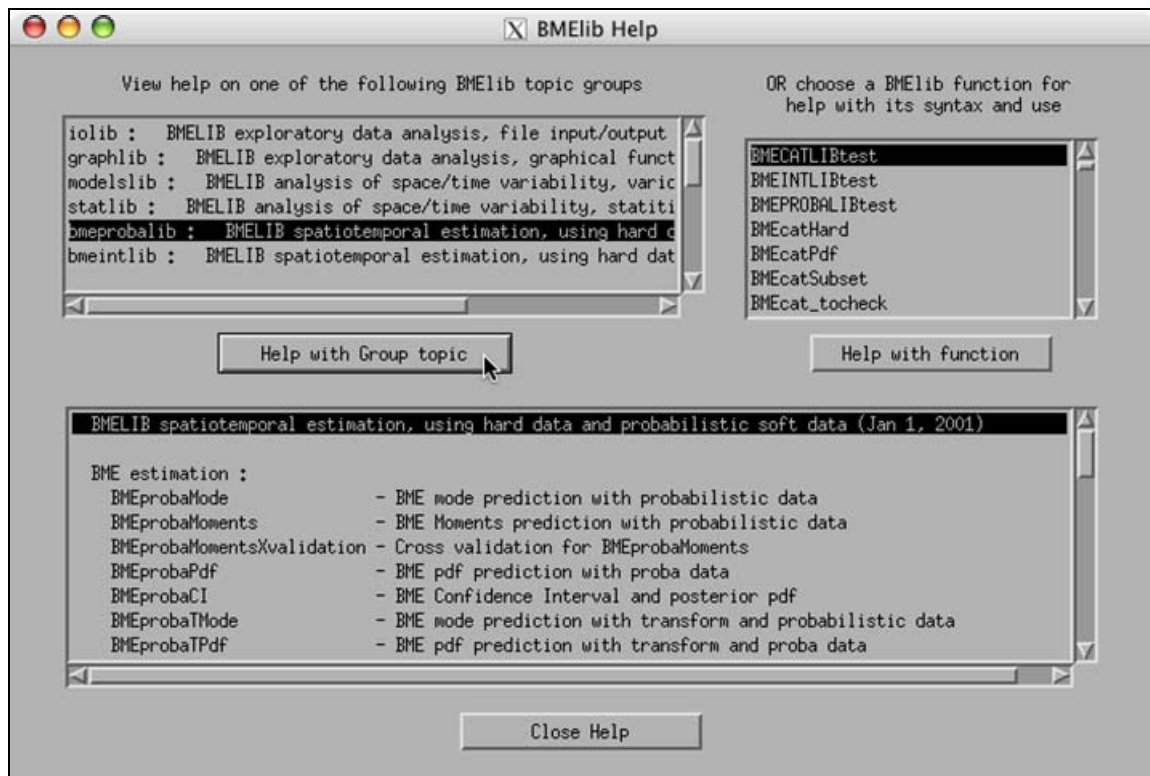


Fig. III.A.3

Back in the “Choose a Task” screen, the “About” button provides brief information on the SEKS-GUI.

The “Exit” button is used to terminate the running SEKS-GUI upon confirmation from the user (Fig. III.A.4).

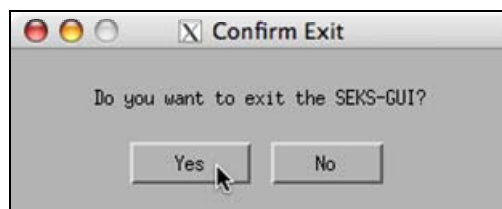


Fig. III.A.4

3. "Hard Data Wizard" – Part I

In this screen hard data information is entered into the system. According to the theoretical knowledge synthesis framework, this type of information should consist of individual values or measurements that are considered to be accurate for the scope of your study. You are now asked to enter this information. If no such data are available push the "Next" button to skip the Wizard's Part II and to be taken to the "Soft Data Wizard" screens (see details in paragraph 5).

1. The hard data (HD) can be in an ASCII (text) file, in an Excel (.xls) file, or in a GeoEAS preformatted file, as explained in this manual's PART II, Section B.1. If you opt to go for the Excel format, your data need to be saved in the first (if there are multiple in the file) or single spreadsheet of the file. Also, GeoEAS is one of the standard formats available for data files – only ask for this option if your HD are so formatted. If you will be using hard data, you have to make a choice from the three available by pushing the appropriate one of the three buttons ("ASCII text", "Excel format" and "GeoEAS format"). Only one of them can be selected at a time (Fig. III.A.5).



Fig. III.A.5

2. Push the button “Browse for Hard Data file”. Be prepared to navigate to where your hard data file is stored and select the desired file. After a successful choice the data filename should appear in the message area next to the button.

⇒ Please make sure you have a valid HD file at hand, as instructed in Section II.B earlier. SEKS-GUI is able to perform a rudimentary check on the file content and warn the user of possible inconsistencies before taking the next step (Fig. III.A.6). However, it is up to the user to provide a suitable input file. Also, be cautious to provide accurate information, as SEKS-GUI can not guess from raw input what plain numbers may stand for.

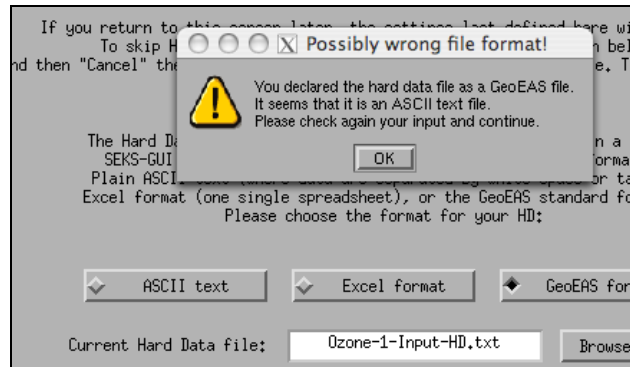


Fig. III.A.6

- ⇒ If the “Cancel” button is pushed during navigation then no HD file will be selected, and a related message will appear.
- ⇒ If a HD file is already chosen and this button is pushed again and, further, the action gets cancelled using “Cancel”, then the file formerly selected is cleared from memory and needs to be entered again using the “Browse...” button.
- ⇒ If you return to this screen from a following one, the HD filename info is kept in memory.

3. Is this a spatial-only investigation? If yes, push the “Space-Only domain” button. If the button is pushed again it will be deselected. The default is a deselected button and implies a spatiotemporal analysis.

⇒ A mistaken choice may cause inconsistencies with the data and errors at a later point.

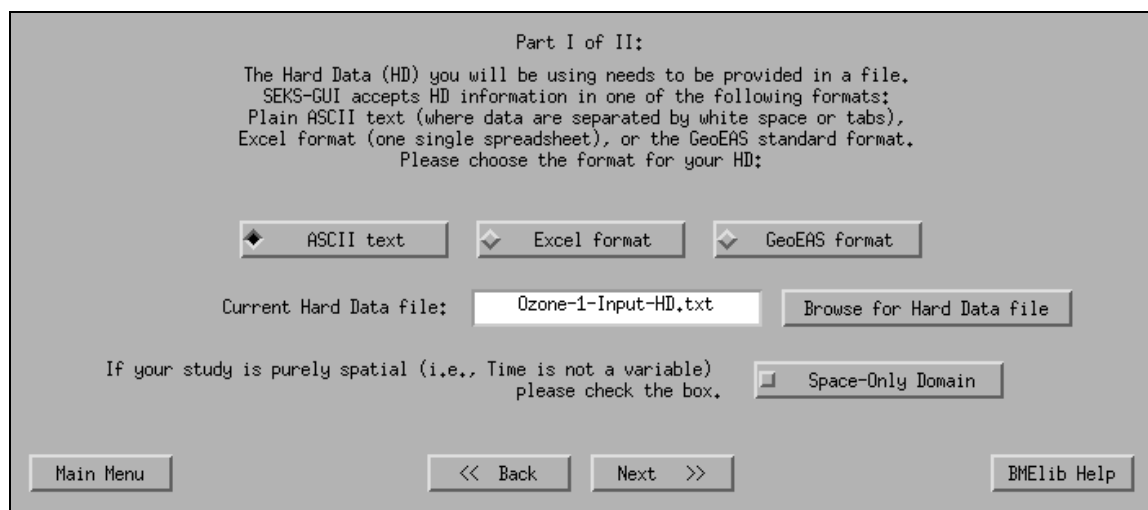


Fig. III.A.7

4. When done, push the “Next” button to proceed (Fig. III.A.7).

4. "Hard Data Wizard" – Part II

1. In this screen details about the HD file are requested to continue. This screen appears only if a HD filename has been specified in the previous one, i.e., only when hard data are used. Following the choice of the HD file, you now have to provide the column numbers where the requested attributes can be found in the file, as Fig. III.A.8 shows. Provide a number in a box only if the corresponding coordinates are used.

- ⇒ A maximum of 3 dimensions (time, if considered, must be included as the last of the reported dimensions) is currently supported by SEKS-GUI.
- ⇒ Please make sure you have a valid HD file at hand, as instructed in PART II, Section B.1 earlier. Also be cautious to provide accurate information, as SEKS-GUI can not guess from raw input what plain numbers in a file may stand for.
- ⇒ If you work on a spatiotemporal investigation the information on the time file column must be entered as the last of the coordinates (i.e., in the y-axis box in the spatial 1-D case, or in the z-axis box in the spatial 2-D case).

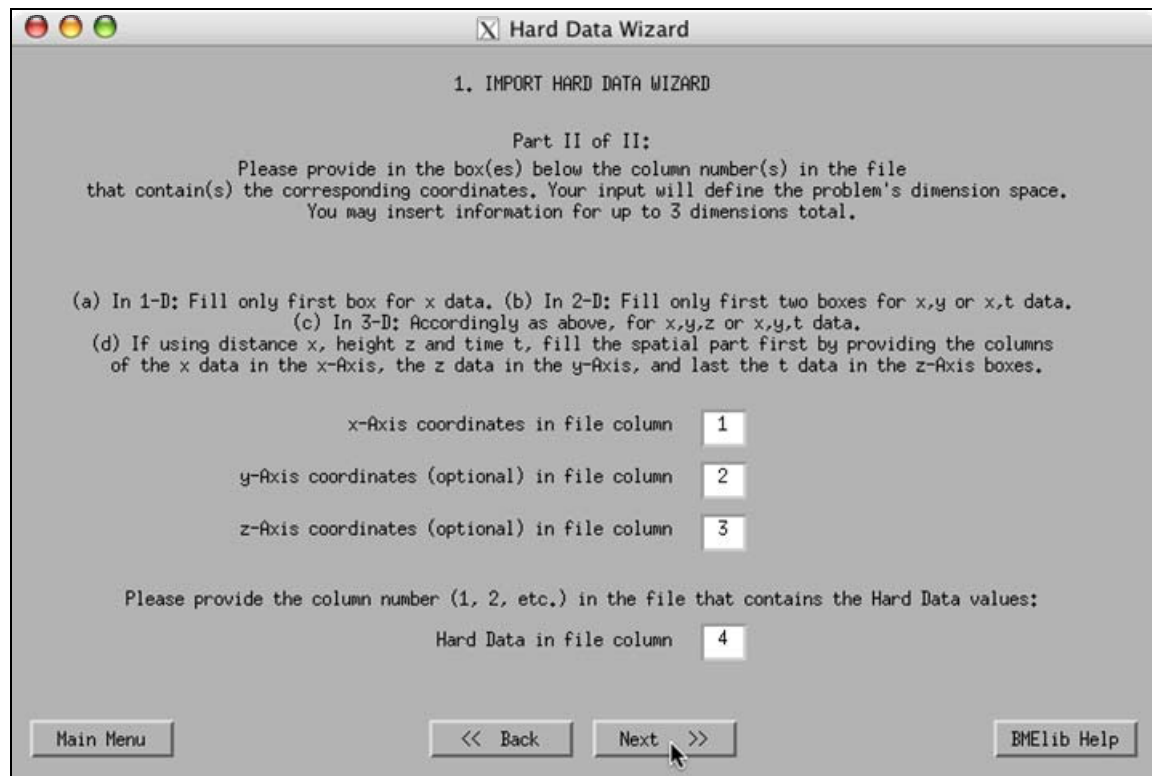


Fig. III.A.8

2. When done, push the "Next" button to proceed.

- ⇒ If the "Back" button is pushed in the following screen, then SEKS-GUI will not return to this one (part II of the hard data wizard), but rather to the part I of the hard data wizard. In such a case the last declared HD file name is kept in memory.

5. "Soft Data Wizard" – Part I

In this screen, soft data information is entered into the system. According to the knowledge synthesis framework, this is information that entails some degree of uncertainty within the scope of your study. Such data and their associated uncertainty can be entered in one of the SEKS-GUI acceptable formats (see PART II, Section B.2). If no such data will be used, push the "Next" button to skip Part II and to be taken to output grid definition screen (see paragraph 8 ahead).

⇒ If no input is present of either hard or soft data and the user attempts to continue, an error screen will appear and the user will be prompted to start anew and choose a task in the screen III.A.2.

1. As discussed in PART II, Section B.2, the soft data (SD) handled by SEKS-GUI can be intervals and probabilistic data. Use to drop-down menu shown in Fig. III.A.9 in order to choose between the variety of options offered by the SEKS-GUI.

Probabilistic SD can be probability density functions (PDFs) where the user provides the characteristics for selected PDF types (gaussian, uniform or triangular); alternatively, the probability densities can be provided as a series of constant values (histogram form) within equal or possibly variably sized bins, or as a series of values at the limits of such bins (linear form) changing linearly between consecutive limits.

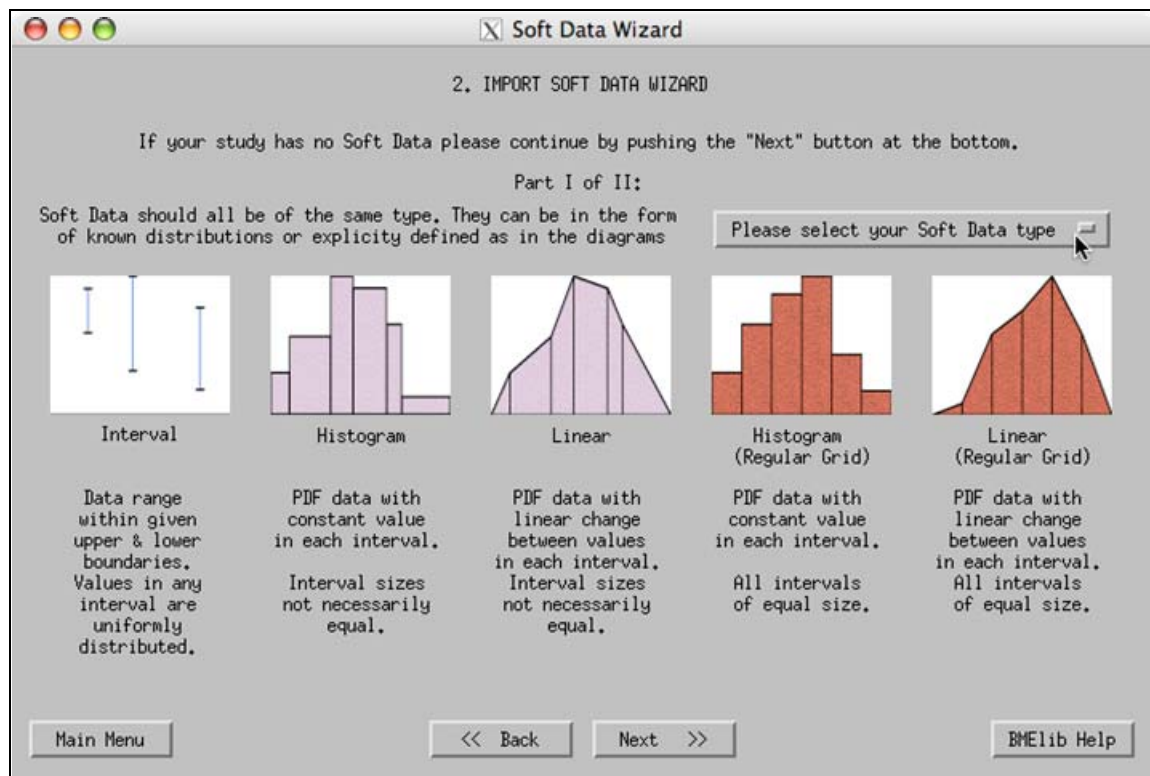


Fig. III.A.9

2. When done with your SD type choice, push the "Next" button to proceed.

6. “Soft Data Wizard” – Part I (continued)

This screen will appear if only soft data and no hard data are used.

1. First, the user needs to enter in the upper box the number of dimensions to be used in the current study (1, 2, or 3 – time, if present, must be included in this number).

2. Designate the domain of your investigation. Is it composite space-time? Or, is it a spatial-only task? If the latter is true, push the “Space-Only domain” button. If the button is pushed again it will be deselected. The default is a deselected button and implies a spatiotemporal analysis.

⇒ A mistaken choice may cause inconsistencies with the data and errors at a later point.



Fig. III.A.10

2. When done, push the “Next” button to proceed (Fig. III.A.10).

7. "Soft Data Wizard" – Part II

1. Based on the considerations of PART II, Section B.2 (regarding eligible formats of the soft data files for input in the SEKS-GUI) push the "Browse for Soft Data file" button to enter your SD information. Be prepared to navigate to where your soft data file is stored and select the desired file. After a successful choice the output info filename should appear in the message area next to the button (Fig. III.A.11).

- ⇒ Please make sure you have a valid SD file at hand, as instructed in PART II, Section B.2 earlier. SEKS-GUI is able to perform a rudimentary check on the file content and warn the user of possible inconsistencies before taking the next step. However, it is up to the user to provide a suitable input file. Also be cautious to provide accurate information, as SEKS-GUI can not guess from raw input what plain numbers may stand for.
- ⇒ If the "Cancel" button is pushed during navigation then no SD file will be selected, and a related message will appear.
- ⇒ If a SD file is already chosen and the "Browse..." button is pushed again and, further, the action gets cancelled using "Cancel", then the file formerly selected is cleared from memory and needs to be entered again using the "Browse..." button.
- ⇒ If you return to this screen from a following one, the SD filename info is kept in memory.

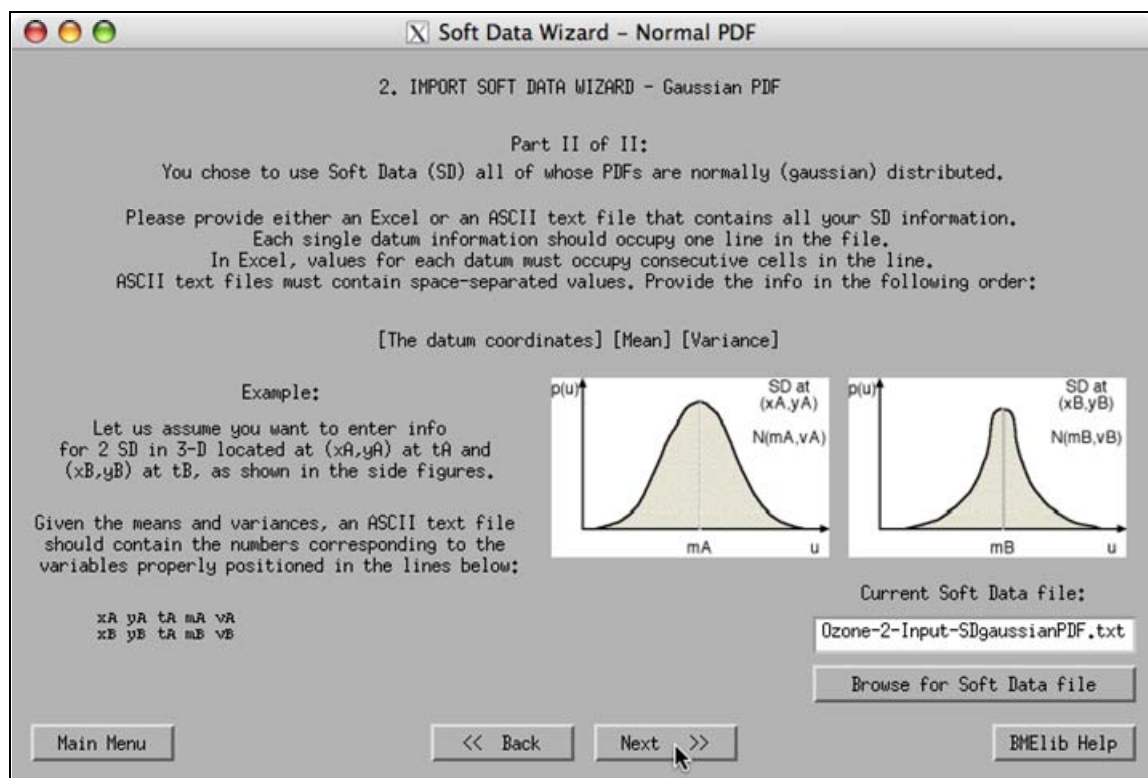


Fig. III.A.11

2. When done, push the "Next" button to proceed.

- ⇒ If the "Back" button is pushed in the following "Output Configuration" screen, then SEKS-GUI will not return to this one, but rather the starting screen of the soft data wizard presented in paragraph III.A.5.

8. "Output Configuration"

In this screen SEKS-GUI prompts the user to define the spatiotemporal (or spatial, in the spatial-only case) locations where estimates will be obtained.

1. Use the drop-down menu shown in Fig. III.A.12 to choose one of the available options:
 - By selecting option A, you will need to provide the limits and node distancing for each of the dimensions used.
 - By selecting option B, you will be asked for the limits and the numbers of nodes for each of the dimensions used.
 - By selecting option C, it will be necessary to define the grid origin (lower limit), the number of nodes and the distance between nodes for each of the dimensions used.
- ⇒ Selection or re-selection of an input format from the drop-down menu will clear any output info file information previously provided by the user to prevent potential errors. For that reason the user needs to always perform the current step 1 before proceeding to the following step 2.

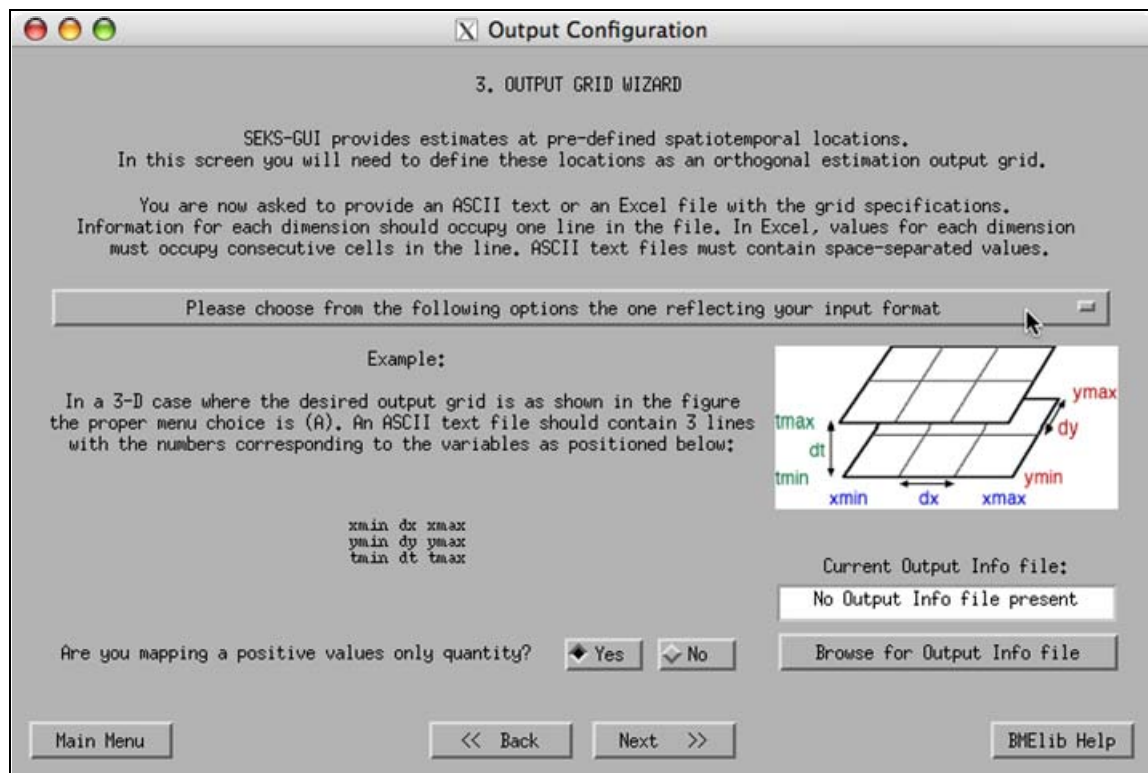


Fig. III.A.12

2. According to the choice made in the previous step, the user must now provide an ASCII or Excel text file that contains the necessary information in the appropriate format, as discussed in PART II, Section B.3. When the file is at hand, push the "Browse for Output Info file" button. Be prepared to navigate to where your output configuration data file is stored and select the desired file. After a successful choice the output info filename should appear in the message area next to the "Browse..." button (see Fig. III.A.13).

- ⇒ As previously stated, it is up to the user to provide a suitable input file. Also be cautious to provide accurate information, as SEKS-GUI cannot guess from raw input what plain numbers may stand for.
- ⇒ Selection or re-selection of an input format from the drop-down menu (see previous step 1) will clear any output info file information previously provided by the user to prevent potential errors. For that reason the user needs to always perform step 1 before proceeding to the current step.
- ⇒ If the “Cancel” button is pushed during navigation then no output info file will be selected, and a related message will appear.
- ⇒ If an output info file is already chosen and this button is pushed again and, further, the action gets cancelled using “Cancel”, then the file formerly selected is cleared from memory and needs to be entered again using the “Browse...” button.
- ⇒ If you return to this screen from a following one, the output info filename info is kept in memory.

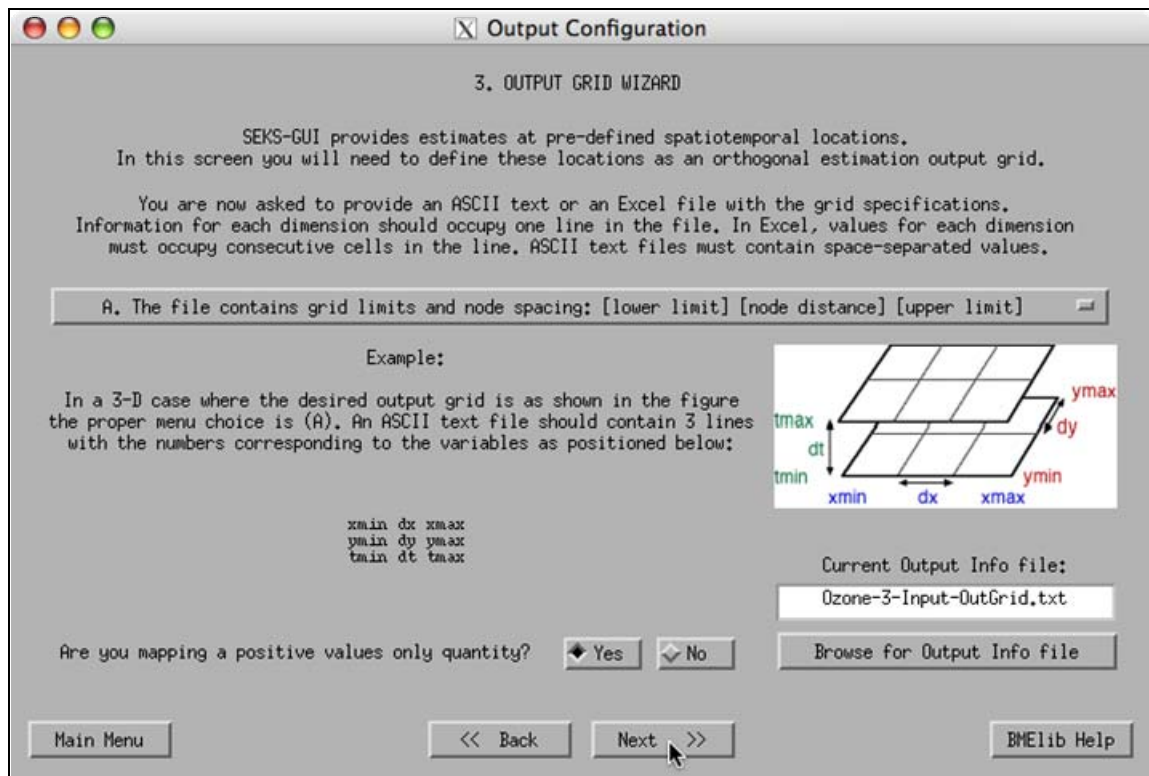


Fig. III.A.13

3. Notice that towards the bottom of the screen the user is asked to state whether the field in the study can take negative values or not. SEKS-GUI automatically scans the input information to provide an initial answer, based on whether the user data contain negative numbers or not. However, if the field is allowed to span into negative values the user should designate that by explicitly clicking on the “No” button. SEKS-GUI will use this response to prevent estimation distributions of positive-only quantities to extend to negative numbers.

4. When done, push the “Next” button to proceed.

9. “Exploratory Analysis” – Part I

In this screen you are not required to take any action. The data are evaluated by the program and a check for duplicates is performed, as duplicate coordinates result in covariance matrix singularities. The same adverse effect in the estimations has been found to occur also in the case when data points are very close spatial neighbors. The current version of SEKS-GUI automatically detects close proximity and handles it as co-location. The degree of proximity is defined individually for each separate investigation; it is independent of the spatial measurement units, and is rather assessed on the basis of the output grid dimensions, as set in the previous screen 8.

In this version of SEKS-GUI, co-located hard data in space/time are averaged, whereas soft data co-located with other hard/soft data are dealt with by slight random spatial displacements (i.e., soft duplicates are not removed from the set). Once the check has been performed, the results are displayed in the corresponding boxes on the screen. At this point you can push the “Next” button to proceed (Fig. III.A.14).

⇒ Please wait until the calculations are finished. Do not proceed to the next screen while there are still “Calculating...” messages in this screen’s boxes.

4. DATA EXPLORATORY ANALYSIS

In this section we will be reviewing the information previously supplied and will be making arrangements for the mapping stage to follow later.

Using BMELib, SEKS-GUI currently supports operating on a spatiotemporal random field of normally distributed variables of a zero (or constant) mean trend. For compliance, in the BMELib analysis the data you provided will be detrended and open for transformation. In the BMEEnumu analysis none of the above steps will be necessary.

Part I of III: Check for Duplicates

Data that are very close or whose coordinates sets coincide can affect the covariance analysis. SEKS-GUI uses a BMELib-based function that averages the values of such Hard Data. Soft Data that are co-located with other Hard or Soft Data are dealt with by means of slight displacements. You are not required to take any action during this step. The outcome is provided below.

		Hard Data	Soft Data		Total
Before Duplicates Check:	Number of points:	377		3075	3452
	Duplicates found:	0	w/ HD 0	w/ SD 0	0
After Duplicates Check:	Number of points:	377		3075	3452

Main Menu << Back Next >> BMELib Help

Fig. III.A.14

At this point in the User Manual, if you are continuing with BME investigations proceed to the following page to Section III.B. If you are using GBME analysis skip the following section instructions and proceed to Section III.C.

Section B – BME Analysis Screens in SEKS-GUI

1. BME “Exploratory Analysis” – Part II

A. Introduction

BMElib operates correctly on normally distributed residual values of a quantity, i.e., on detrended information that follows a gaussian distribution. The following 2 screens fulfill the objective of bringing the user raw information into the above suitable processing form.

The first action that needs to be taken with the user information is to remove any trends in the data set and obtain the data residual values. In this screen the user obtains a mean trend from the data distribution. The detrending takes place by applying a gaussian kernel across the dataset.

In this version of SEKS-GUI, the kernel searches for neighboring data within user-defined ranges in space/time, and extracts the trend by applying a smoothing moving window. However, if any unusually high (or low) values exist in the data set, then the moving window may be biased by these values. In this manner, the smoothing effect may produce artifacts caused by extreme values in the data, thus drastically affecting the trend estimate. SEKS-GUI addresses this issue by identifying and isolating from the detrending process potential outliers in the user data. In particular, extreme outliers are excluded from the data distribution trend estimation using criteria based on the box plot graphical techniques. The data used for the estimations are not altered by this process.

Estimation of the mean trend relies on the use of point values, which are normally represented by hard data. The user data set, though, may contain a limited amount of them or no hard data at all. SEKS-GUI resolves this shortcoming by employing the soft data in the trend estimation, as well. For this procedure, the means of the soft distributions (or the middle points of the soft intervals) are considered as a soft data approximation, which is a compromise to potentially insufficient information conditions. This approximation also proves to be useful in the coming stage of covariance modeling, as will be shown in paragraph III.B.3, and is independent from the BME estimation stage where soft information is considered in full as provided by the user.

Based on the output grid size, the kernel smoothing radius is set to 1/10 of the shortest extent among all grid directions. The maximum search radius for data to contribute to the trend in space/time is adjustable by the user. The default setting (starting values) that appears on the screen in the beginning is half the size of the longest extents among all grid directions in space and time, respectively.

⇒ The user can not proceed to the following screens unless there are trend data available to the SEKS-GUI.

B. Instructions

1. On the left-hand side of the screen statistics on the data set are displayed (Fig. III.B.1). Depending on the presence of hard and/or soft data, these statistics are based on the hard data and the soft data approximations, and only refer to the whole data set (i.e., the user cannot ask for the statistics of subsets at individual temporal instances in the spatiotemporal case). The user can choose from the drop-down menu whether the non-detrended (“Non-D”)

or the detrended (“Detr”) data set statistics will be on display. The latter are only available only after the mean trend has been estimated, or after the user has loaded a previously saved mean trend file for the particular data set (see paragraph 3 below).

- ⇒ If the user asks for statistics that are unavailable, then the data statistics boxes display the message “N/A”.
- ⇒ The data statistics boxes cannot be edited by the user. They are only available for display of values.

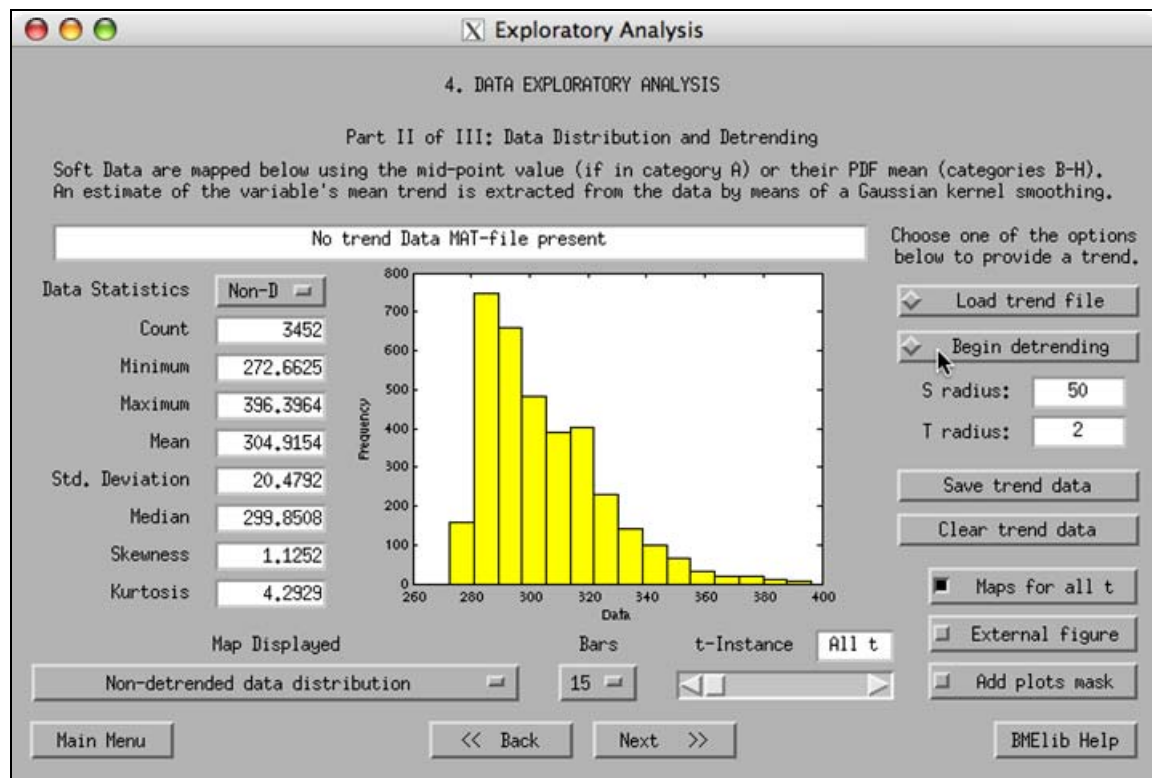


Fig. III.B.1

2. If you are visiting this screen for the first time in a study, set the desired maximum search radius for space and time (temporal maximum radius not available in purely spatial cases) in the “S-radius” and “T-radius” boxes, respectively, on the right-hand side of the screen. These are the distances around each location within which the kernelsmoothing algorithm will search for spatiotemporal neighbors to obtain a local average. Then push the “Begin detrending” button to extract the mean trend from the experimental data (Fig. III.B.1). Detrending may take a while to complete depending on the amount of data to process.

- ⇒ Initial space and time radii (as appropriate) are automatically set by SEKS-GUI upon visiting this screen, on the basis of the output grid dimensions, as set in the “Output Configuration” screen of paragraph III.A.8. The user may or may not need to modify these values to achieve the desired degree of smoothness in the estimated mean trend.
- ⇒ The larger the radius value is, the smoother the trend will appear. The outcome of this stage depends purely on the user judgment and intuition, as the actual trend is an unknown.
- ⇒ Please be patient and wait until the calculations come to an end. Matlab and the SEKS-GUI can not respond during the calculation time to any other commands. Please refrain from pushing buttons at this time, as these actions are queued and may result in

unwanted events after the calculations are done. Calculations may take a lot of time, depending on the volume of data in the data set and the hardware you are using.

- ⇒ At the end of the calculations you will be prompted to save the outcome data in a file. If you consider the results satisfactory it is strongly suggested to save them in a suitably named file, as you may later wish to return or re-run this study (see further below on how to evoke previously saved data). The data are saved in a Matlab format which has the ending “.mat” at the folder you specify, and can not be viewed independently unless they are loaded within the Matlab environment. If you decline to save the trend data but instead prefer to explore the maps output first, you can save them at a later point, but only while still viewing this screen, by pushing the “Save trend data” button.

3. If you are returning to this screen or re-running the same data set study, you may wish to load a previously saved version of the mean trend (having followed the previous step 1 at an earlier time). In that case, when the screen appears opt to push the “Load trend file” button. You will be asked to navigate in your filesystem to find the Matlab-formatted file ending in “.mat” that contains the trend data. Once selected, the file name will appear in the message box (Fig. III.B.2) and the detrended data set statistics will appear in the boxes on the left-hand side.

- ⇒ Please make sure you provide proper input, because SEKS-GUI cannot guess on the file contents.

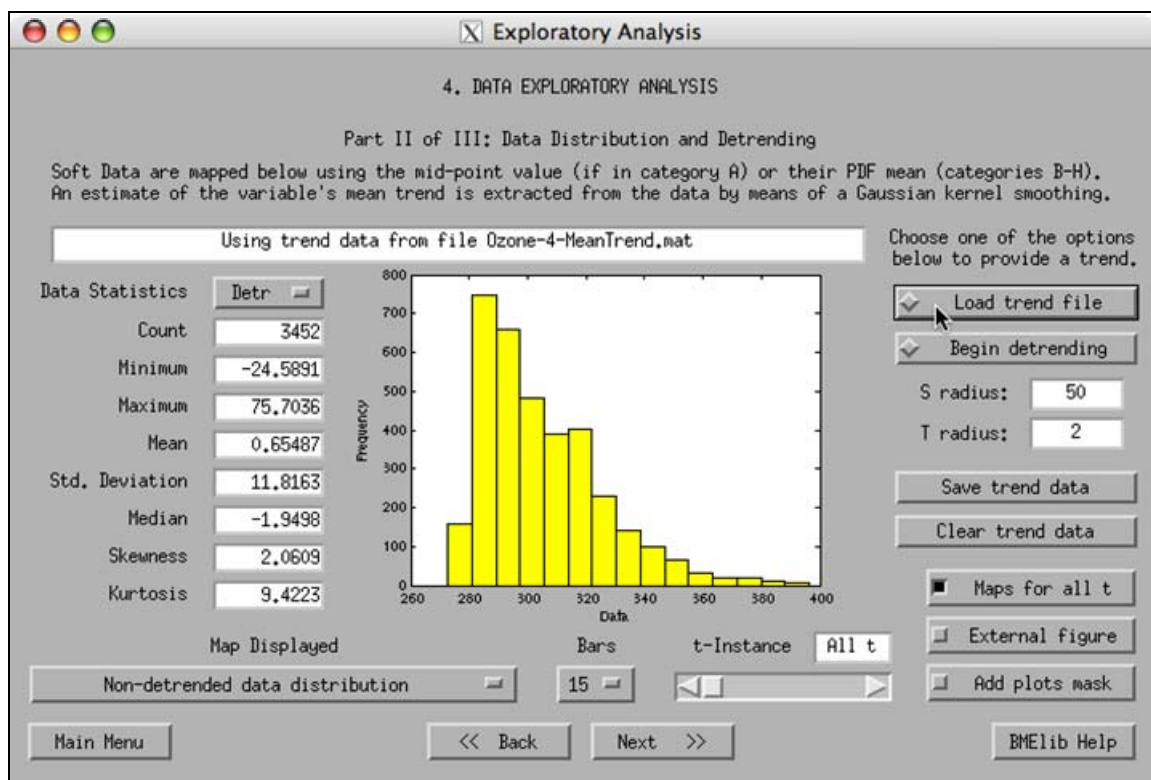


Fig. III.B.2

4. The message box on the upper part of the screen communicates useful messages to the user and cannot be edited.

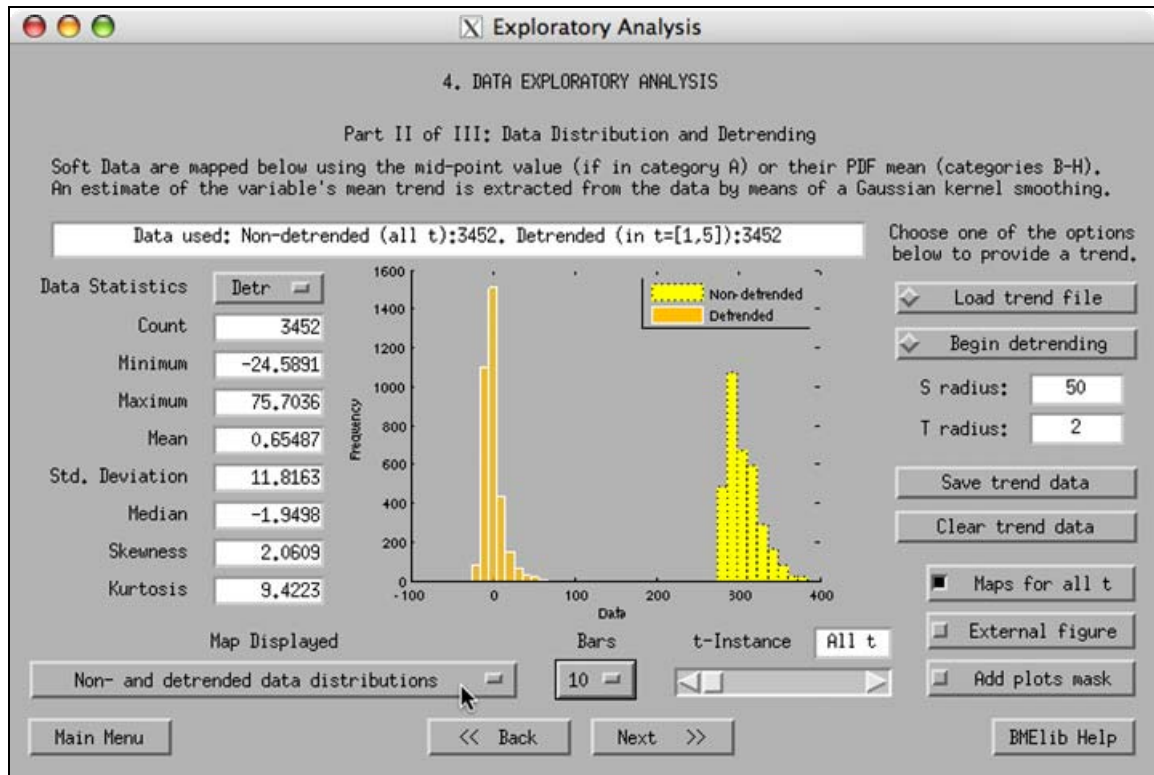


Fig. III.B.3

5. The user can plot a variety maps from the drop-down menu (under "Map Displayed"). At any time the user can choose to view maps of:

- All data locations
- Hard data locations
- Soft data locations
- Markerplots of all data (hard data and soft data approximations)
- Colorplots of all data (hard data and soft data approximations)
- Non-detrended data distribution

Once trend data are available, the following maps can also be created:

- Detrended data distribution
- Non- and detrended data distributions (Fig. III.B.3)
- Mean trend of variable in space, or at a chosen t-instance in space-time (Fig. III.B.4)

If data necessary for a particular request are not available, a message will appear in the message box on the screen.

- ★ Use the “Bars” drop-menu to display the histogram plots with more or fewer bars as desired.
 - ★ Use the “t-Instance” slider (or write a suitable number in the “t-Instance” box) to view maps at a particular instance in time, as in Fig. III.B.4 (not available in purely spatial cases).
 - ★ Activate or de-activate as desired the “Maps for all t” button to see time-aggregated maps of data locations and distributions or maps at user-selected instances, respectively (not available in purely spatial cases, or when requesting maps of individual instances of the mean trend).
 - ★ Use the “External figure” button to display the plot requests that will follow in an external, independent window when activated (Fig. III.B.5), or to return to the in-screen display when de-activated.
- ⇒ When the “External figure” button is activated, it enables complete control of the plot by making use of Matlab tools (e.g., axes rotation, renaming, etc.). Also, this feature allows the user to print the particular figure using the independent window menu. For more information on handling plots in separate figure windows please consult Matlab Help.
- ⇒ When plotting maps with legends, Matlab versions pre-dating the 7.1, R14, Service Pack 3 version may produce legend-related error messages in the main Matlab Command Window. These do not affect the SEKS-GUI execution, and can be ignored.

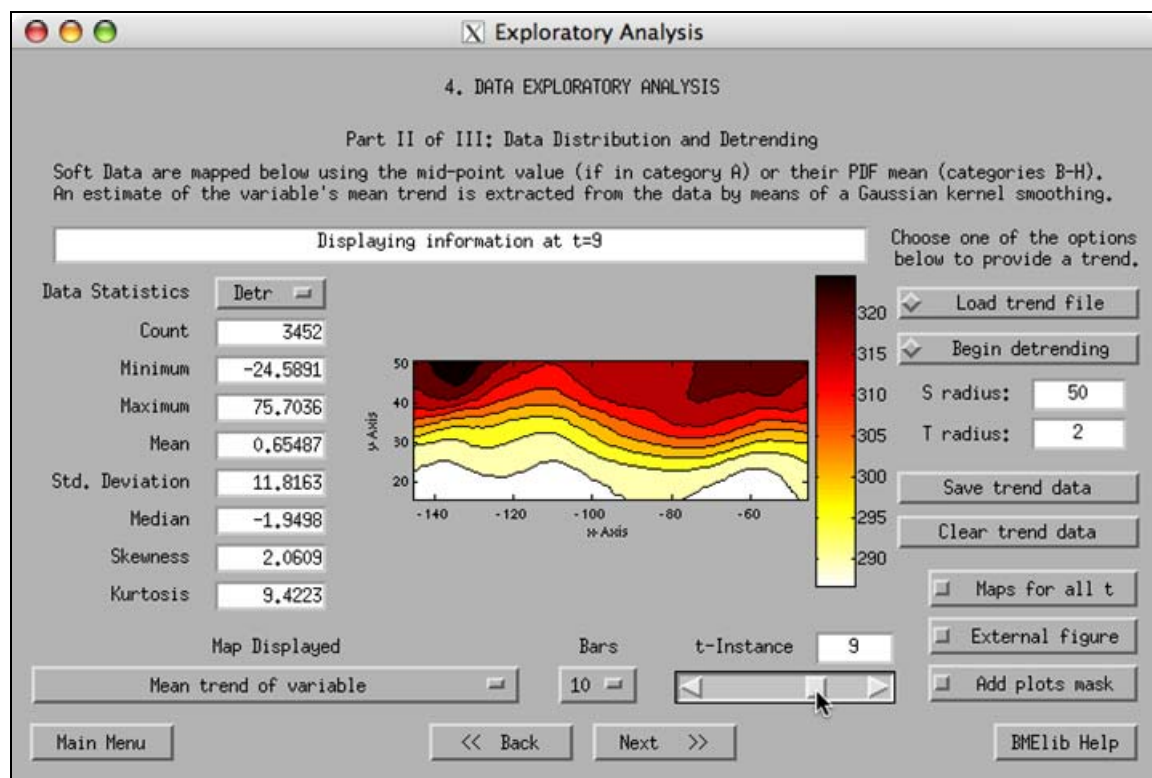


Fig. III.B.4

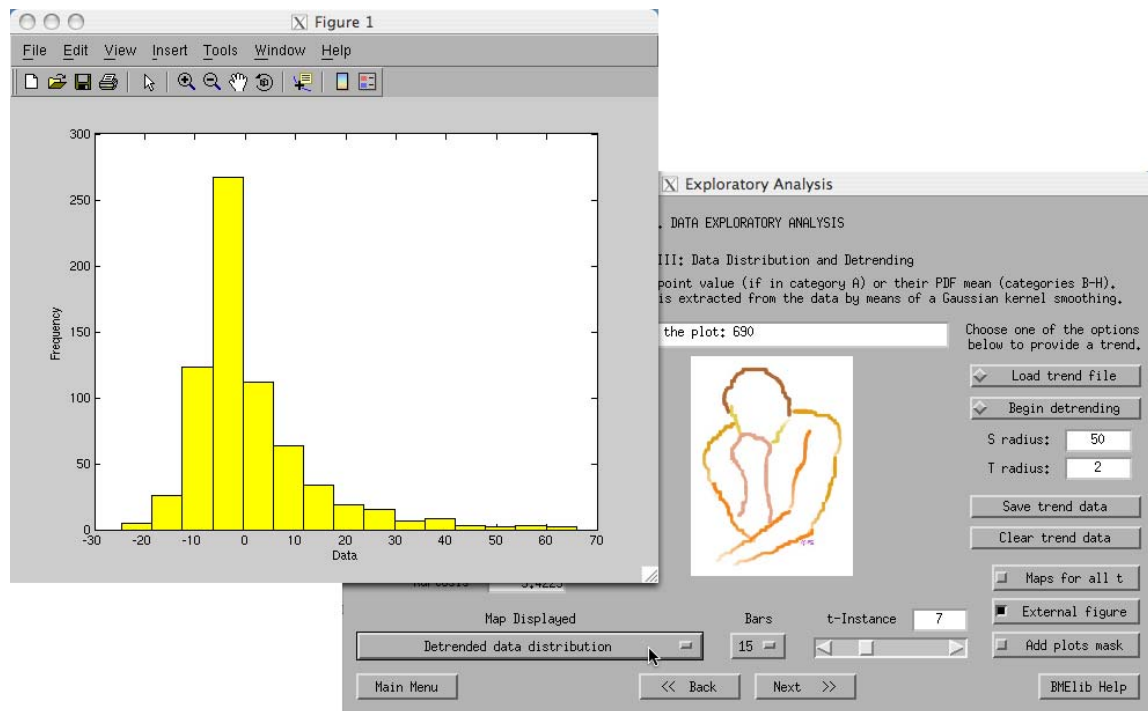


Fig. III.B.5

6. A more advanced feature for map presentations is plot masking, and requires some knowledge of Matlab programming. This is a useful feature if you would like to show the results that appear in part of your output area (e.g., by masking out the portion of a map outside a country's borders). It is more advanced because it requires that the user provides a Matlab ".m" code file with the masking information. If you can program in Matlab, you can create a map that produces the desired mask over the output grid area using suitable coordinates (see also paragraph V.3 in this manual for a basic utility provided with SEKS-GUI to assist you in the mask creation). Upon activation of the "Add mask to plots" button the user is prompted to locate such a masking code Matlab file within the user's filesystem. If you push this button accidentally you can cancel the file search. If the button is activated you can push it again to de-activate it.

7. When done, push the "Next" button to proceed.

2. BME “Exploratory Analysis” – Part III

A. Introduction

BMElib operates correctly on normally distributed residual values of a quantity, i.e., on detrended information that follows a gaussian distribution. This screen continues the task of bringing the user data into the required form, after detrending of the data in the previous Screen. In particular, the user reviews the Cumulative Density Function (CDF) of the detrended data, and is now expected to decide on whether the data set used needs to undergo a transformation. This data transformation aims to reshape the detrended data set from the original-space into a space where their distribution will resemble a normal one. The transformations are based on the detrended hard data and soft data approximations set. If there are soft data in the study, the actual soft information used in the estimations is consequently translated into the transformation space based on the transformation choice.

The measure of deviation from the normal distribution is the comparison of the data CDF with the normal distribution CDF that has the same mean and variance as the data in the corresponding space. For example, let us consider the case of a logarithmic transformation of the detrended user values. The user data statistics in the log-space can be seen on the left-hand side of the screen. The detrended, log-transformed user data distribution is checked for normality by comparing its CDF to the normal distribution CDF which is defined by the mean and variance of the detrended, log-transformed data.

SEKS-GUI allows the user to choose among the following options:

- a. No transformation: The detrended data set is left unaltered and the user proceeds to the estimations with data in the original-space.
- b. N-scores (also known as Normal Scores or Gaussian Anamorphosis; Olea, 1999) transformation: The detrended data set is suitably transformed to follow a gaussian distribution of mean of 0 and variance equal to 1. The resulting data values that will be used in the estimations are in the N-score-space.

Back-transformation is allowed by means of the transformation N-score matrix that establishes corresponding values between the original and the N-score spaces. It is possible that some extreme values in the estimates may not back-translate properly using the N-score matrix. SEKS-GUI addresses this issue by using the following commonly used technique: Upper and lower value bounds are set in the transformation. These bounds are set individually for each investigation, depend on the data span in the particular study, and provide a cushion on which possibly extreme estimates can back-transform into the original-space. All of the above actions are performed automatically in the SEKS-GUI and are seamless to the user.

- c. Box-Cox transformation: The detrended data set is tested with a series of power transformations based on a parameter λ typically ranging within $[-2,2]$. The transformation eventually uses the value of λ that brings the data distribution closest to a normal one. If $\lambda=0$, then the Box-Cox transformation is defined as the logarithmic transformation.

The resulting data values that will be used in the estimations are in the BoxCox-space. Backtransformation depends on the optimal λ value that was chosen for the specific data set.

The Box-Cox transformation is only defined for positive data values. Since this transformation is applied on the detrended data set, it is highly likely that the detrended data will feature negative or 0 values. In the case of negative values,

SEKS-GUI adds a constant to the detrended set so that all values to be transformed are positive. This constant is later removed from the estimates before their back-transformation to the original-space. All of the above actions are performed automatically in the SEKS-GUI and are seamless to the user.

Finally, if under some scenario there exist values equal to 0 in the set, these must be approximated by an adequately small number before being subjected to transformation. This will prevent the event of having to calculate the logarithm of 0 that is unacceptable numerically. Towards this goal, SEKS-GUI initially maps automatically 0 as 10^{-3} to be used if the Box-Cox transformation is selected. The user is further allowed to map 0 using values in the range of 10^{-1} to 10^{-10} instead, depending on the data set processed. **Example:** If the user data set range is between 0 and 1, then it may be a wiser choice to map 0 as a smaller number (e.g., 10^{-6}) than the case where the data set ranges between 0 and 1000 (where a choice of mapping 0 as 10^{-3} may be adequate enough).

- ⇒ Note that none of the above options is the correct or the wrong one. In the current stage, the detrended data are further manipulated to comply with the estimation process requirements. It lies in the user's choice and experience to try alternatives that will lead to a successful estimation process. Different approaches can be attempted. Each transformation type forces the detrended data set in the original-space into a set of different values, and therefore different dynamics may develop within the data set as a result of the transformation. SEKS-GUI simply provides the user with options to optimally handle the available information in a manner that requires the least amount of effort on behalf of the user.
- ⇒ SEKS-GUI automatically suggests the use of a transformation if the maximum observed deviation of the detrended distribution from normality is larger than 10%.

B. Instructions

1. On the left-hand side of the screen statistics on the data set are displayed (Fig. III.B.6). Depending on the presence of hard and/or soft data, these statistics are based on the hard data and the soft data approximations. The user can choose from the drop-down menu the form of the data for which statistics will be displayed. In this screen, options include:

- Non-detrended data ("Non-D" in the menu, same as in previous screen)
- Detrended data ("Detr" in the menu, same as in previous screen)
- Detrended, N-score transformed data ("D-NscTr" in the menu)
- Detrended, Box-Cox transformed data ("D-BcxTr" in the menu).

Statistics on the transformed data are available as transformations take place before the screen appears. In this way, all different forms of the user data are available for comparison so that the user can make an optimal choice.

- ⇒ The data statistics boxes cannot be edited by the user. They are only available for display of values.

2. The message box on the upper part of the screen communicates useful messages to the user and cannot be edited.

3. On the right-hand side of the screen there is the Transformation menu. Click on the drop-down menu to decide what form of the data will be used for the estimation stage later. As analyzed earlier in this screen's Introduction part, the available options are:

- No transformation
 - N-score transformation (“Use N-scores” in the menu)
 - Box-Cox transformation (“Use BoxCox” in the menu)
- ⇒ This is an important choice for your investigation. As stated earlier, there is no single correct answer to your decision, but one should rather consult the data in each individual study to proceed. You can always repeat your study using a different transformation type and compare the estimation results in the end, if necessary.
4. If the Box-Cox transformation is selected for use, there may be an issue with 0 values in the detrended, non-transformed data set. The transformation may require calculation of the logarithm of 0 and this corresponds to $-\infty$. However, numerically this is not acceptable, so the user is asked to provide a correspondence of 0 to a suitably small value in the original-space. The approximation is left to user’s discretion, as discussed in the Introduction part of the current screen. A default approximation of 0 by the value 0.001 is used when the screen appears. The user can adjust the approximation value with the slider on the lower right-hand side (Fig. III.B.6).
- ⇒ If the user makes a choice other than the Box-Cox transformation, the investigation course is not affected at all by the 0 mapping slider value.

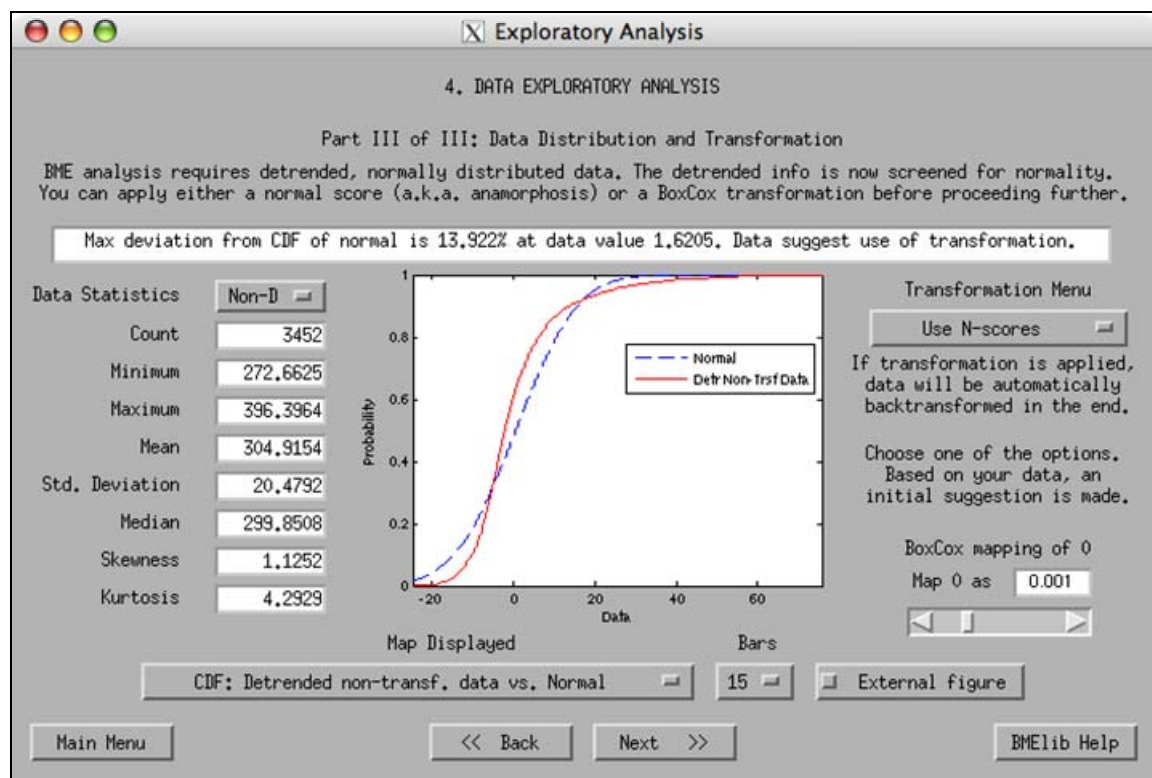


Fig. III.B.6

5. The user can plot a variety maps from the drop-down menu (under “Map Displayed”). At any time the user can choose to view maps of:
- Comparison of the detrended, non-transformed data CDF (red line in plot) against the normal CDF defined by the detrended, non-transformed data mean and variance (blue dashed line in plot).

- Comparison of the detrended, N-score-transformed data CDF (red line in plot) against the normal CDF of mean 0 and variance 1 (blue dashed line in plot).
 - Comparison of the detrended, Box-Cox-transformed data CDF (red line in plot) against the normal CDF defined by the detrended, Box-Cox-transformed data mean and variance (blue dashed line in plot).
 - Histogram of the detrended, non-transformed data distribution (as in the previous screen).
 - Histogram of the detrended, N-score-transformed data distribution (Fig. III.B.7).
 - Histogram of the detrended, Box-Cox-transformed data distribution.
- ★ Use the “Bars” drop-menu to display the histogram plots with more or fewer bars as desired.
 - ★ Use the “External figure” button to display the plot requests that will follow in an external, independent window when activated, or to return to the in-screen display when de-activated.
 - ⇒ When the “External figure” button is activated, it enables complete control of the plot by making use of Matlab tools (e.g., axes rotation, renaming, etc.). Also, this feature allows the user to print the particular figure using the independent window menu. For more information on handling plots in separate figure windows please consult Matlab Help.
 - ⇒ When plotting maps with legends, Matlab versions pre-dating the 7.1, R14, Service Pack 3 version may produce legend-related error messages in the main Matlab Command Window. These do not affect the SEKS-GUI execution, and can be ignored.

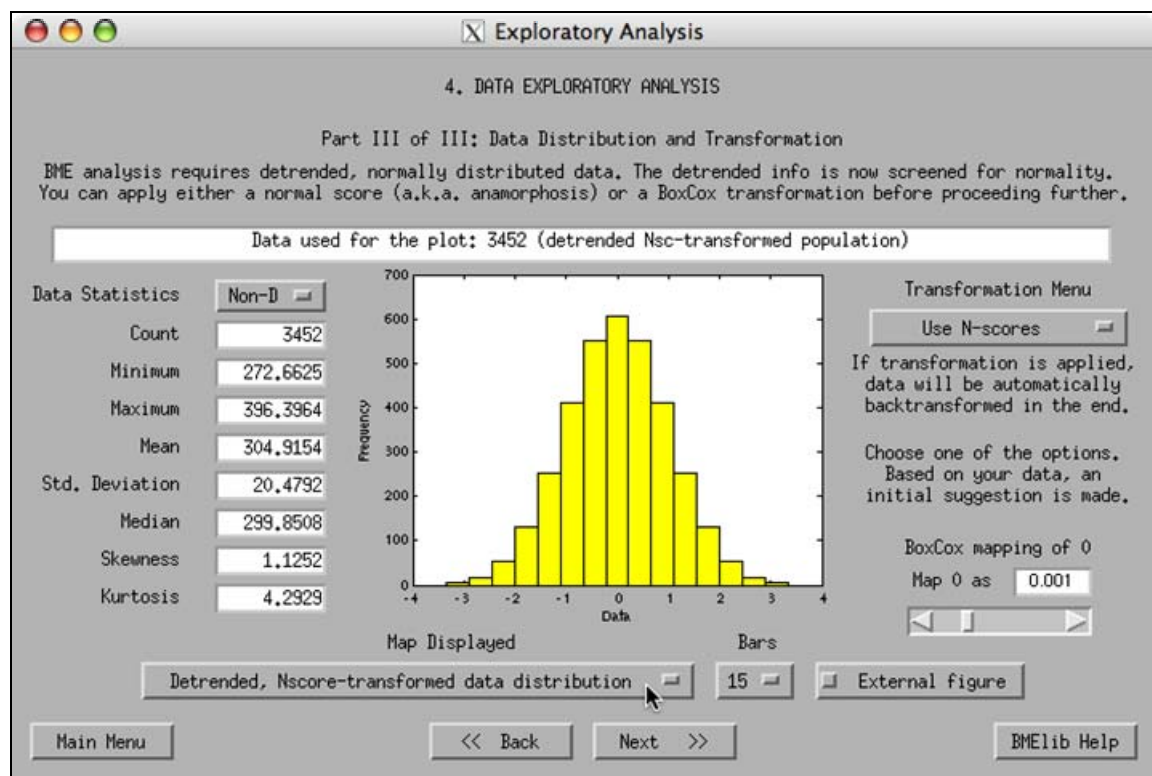


Fig. III.B.7

6. When done, push the “Next” button to proceed.

3. BME “Covariance Analysis” – Part I

A. Introduction

In covariance analysis section the goal is to describe potential spatiotemporal correlation patterns among the data.

- ⇒ The two screens in paragraphs III.B.3 and III.B.4 are combined in one screen in the purely spatial case.
- ⇒ The following refer to a spatiotemporal investigation. In a purely spatial case references to the temporal element are to be ignored.

It is crucial to stress that correlations may differ at different spatiotemporal neighborhoods due to the nature of the field under investigation. In that sense, for example, a pattern that is modeled through a particular covariance function and applies for a specific spatial neighborhood may be inappropriate for another neighborhood on the same output grid. In principle, the correlation patterns also differ when changing the scale of observation; for example, correlations among data in a large grid do not exhibit, in general, the same behavior as the correlations in a more localized scale within the same grid. The effects of upscaling or downscaling are very important in the prediction process, and the user must be very cautious if the investigation includes action in different scales (see also Christakos *et al.*, 2002).

BME analysis in SEKS-GUI allows for the above considerations by means of multiple investigations: Ideally, the user needs to have a good prior knowledge of the spatiotemporal correlations of the underlying field. In this manner, the user can split appropriately the prediction field in sub-grids, and repeat the BME-based analysis for each one of these neighborhoods by defining a different covariance function as necessary. Alternatively, if no good prior knowledge is available regarding the underlying spatiotemporal correlations, the user may try to obtain predictions on the desired estimation grid as one solid neighborhood. If the user believes that correlations may indeed differ within the grid neighborhoods, then the grid can be split appropriately and investigations can be repeated for each one of the sub-grids therein. The results from the two approaches [i.e., (i) one grid using one covariance function, and (ii) multiple sub-grids within the same grid, each with a covariance function of its own] can be compared by the user to reach a reasonable conclusion regarding the nature of the spatiotemporal correlations in the field.

- ⇒ The GBME analysis performs the above investigation phase automatically during the GBME estimations stage as it performs on a local level, rather than a global level as BME does. See also the introductory discussions in paragraphs III.C.1 and III.C.2.

Based on the previous, the SEKS-GUI BME covariance analysis takes place in the full extent of the output grid defined earlier (see paragraph III.A.8) in two stages, as follows:

- a. Explore all data at different distances between them (in particular, by considering classes of spatial and temporal distances – also known as lags), and obtain the experimental (or sample) covariance values calculated for each of the distance classes. The number of these classes in which the experimental covariances are calculated is a parameter set by the user.
- b. Fit a model over the individual experimental covariance points calculated in stage (a), so as to obtain an explicit expression for the covariance across space and time.

In a similar manner as seen at the detrending stage of the exploratory analysis, only individual values can be used in the covariance analysis, as well. However, based on Christakos *et al.* (2002), for the purposes of the covariance analysis of soft data, the covariance of the soft data means is equivalent with the covariance of the soft data themselves. To this end, SEKS-GUI conveniently uses at this stage again the set consisting of the – now detrended and possibly transformed – hard data and the soft data approximations (defined earlier as the means of the soft distributions or the middle point in case of soft intervals).

In stage (a) the user is offered 3 different directions along which one can explore the sample correlations: The East-West or 0° axis, the North-South or 90° axis and an all-directional analysis. The user must then define an estimate of spatial and temporal ranges to which the sample correlations extend. These numbers may be approximately known if the user has a thorough idea of the data that are used or the phenomenon/process under investigation. If these maximum correlation ranges cannot be approximated based on prior knowledge, the user can try to repeatedly calculate the experimental covariance for a series of range values that will reveal eventually how far in space/time the sample correlation extends. To this end, the user is also offered the convenience of setting a variable number of space and time lags. At each one of these distance classes the correlations are investigated by means of the number of data neighbors (to any other given datum) that are found within the class. One value of experimental covariance is calculated for each such class.

⇒ Based on the previous, the user can calculate more values of the experimental covariance if more distance classes are considered in space and time. However, the calculated covariance values depend on the amount of data neighbors found at each one of these lags. If no neighbors exist within a particular lag, then a warning message appears in the Matlab Command Window reporting that “some space/time classes do not contain pairs of points”. This warning does not affect the use of the SEKS-GUI as long as the user can eventually obtain a satisfactory number of experimental covariance points to fit a model on. The numbers of spatial and temporal lags to be used are therefore on a balance with the data sample size. As stated previously in different occasions, the user is advised to proceed with caution: A good knowledge of the data, and trying different values for the correlation length and lag number parameters can provide valuable insight in defining the experimental covariance.

For the creation of the covariance model in stage (b) the user chooses up to 2 individual models to nest in each other from a selection of acceptable covariance functions, and then adjusts covariance parameters for each of the selected nested models until an optimal visual fit is reached. These parameters are the covariance sill and range. The **sill** is the maximum value the covariance model can take, and it is the sum of the nested models sills (or is the sill of the single nested model, if only one such is used). The spatial or the temporal **range** are measures of how far the correlation spans in space or time, respectively. In case of nested models, the covariance model range is based on a combination of the nested components.

⇒ The user cannot proceed to the following screens unless there are covariance data available to the SEKS-GUI.

⇒ The covariance has a maximum at 0 space/time lag, as the maximum correlation of each datum is with itself. The covariance value at lag 0 equals the data set variance.

B. Instructions

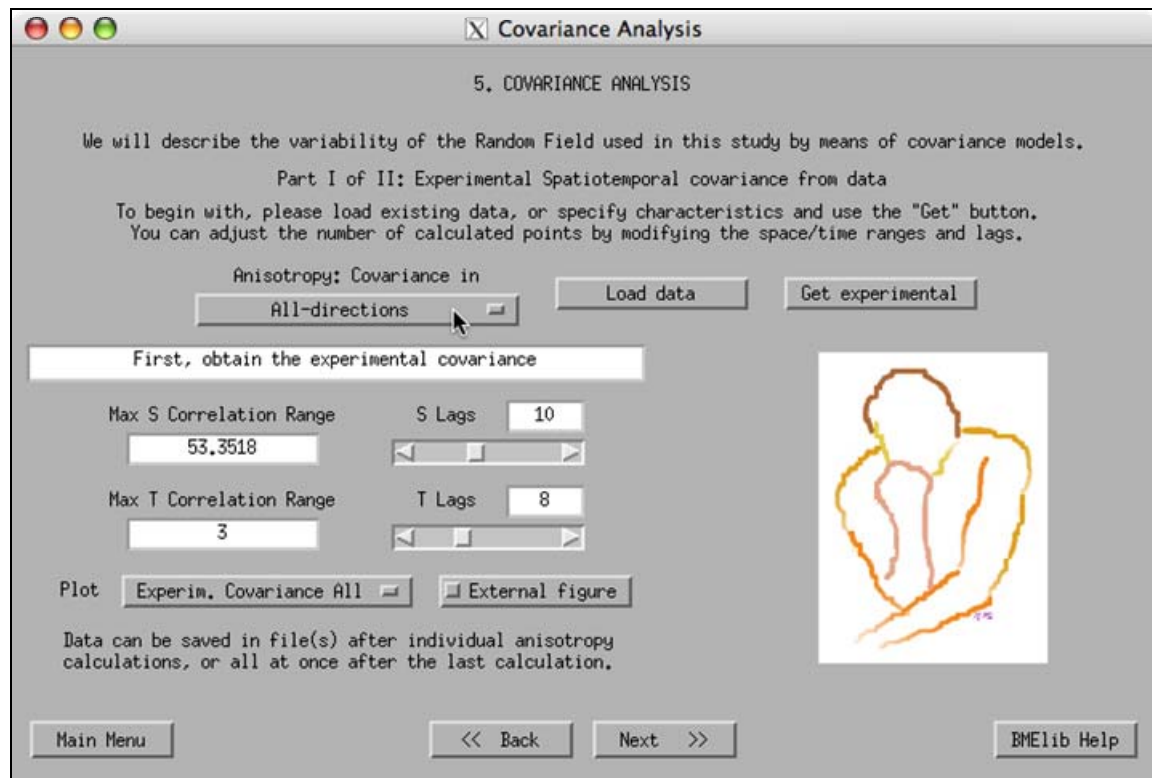


Fig. III.B.8

1. The drop-down menu “Anisotropy: Covariance in” allows the user to choose the direction in which the experimental covariance will be calculated (Fig. III.B.8). The user has the following 3 directions options:

- All-directions
- East-West (0° axis)
- North-South (90° axis)

This choice is the first step prior to calculations. Before proceeding to obtaining the experimental covariance, the user needs also to provide the maximum space/time correlation ranges (see 5 below) and lags (see 6 below), and finally initiate the calculations (see 3 below).

⇒ Once the experimental covariance has been calculated in any of the direction options, the user may choose to store the data immediately, or cancel saving and continue for the covariance in a different direction. After each calculation the user is prompted to save the results. Typically, you may want to calculate the experimental covariance in all of the offered direction options, and then save all the results after the last calculation.

2. If you are returning to this screen or re-running the same data set study, you may wish to load a previously saved version of the experimental covariance information (having followed step 3 below at an earlier time). In that case, when the screen appears push the “Load data” button (Fig. III.B.9). You will be asked to navigate in your filesystem to find the Matlab-formatted file ending in “.mat” that contains the experimental covariance data. Upon

successfully loading a pre-existing experimental covariance file for the current investigation, a message will appear acknowledging the event.

⇒ Please make sure you provide proper input, because SEKS-GUI cannot guess on the file contents.

3. The “Get experimental” button initiates calculations on the BMElib experimental covariance where the correlation range and lags parameters have been set in advance by the user (see following steps 5 and 6 for related instructions).

⇒ Be patient and wait until the calculations come to an end. Calculations may take sometime, depending on the volume of data in the data set and the hardware you are using. There is no indication on the screen regarding the calculation progress. Matlab and SEKS-GUI cannot respond during that time to any other commands. Please refrain from pushing buttons at this time, as these actions are queued and may result in unwanted events or errors after the calculations are done.

⇒ At the end of the calculations in any of the direction options you will be prompted to save the outcome data in a file. In case of multiple calculations, you can opt to save the data once at the end. If you consider the results satisfactory, it is strongly suggested to save them in a suitably named file, as you may later wish to return or re-run this study (step 2 above explains how to evoke previously saved data). The data are saved in a Matlab format which has the ending “.mat” in the folder you will specify, and cannot be viewed independently unless they are loaded within the Matlab environment. Once prompted to save the data, the user does not have an option to save them at a later point.

4. The message box under the anisotropy choice drop-menu communicates useful messages to the user and cannot be edited.

5. The maximum spatial correlation range box and the maximum temporal correlation range box (not available in the purely spatial case) can be edited by the user. These fields adjust the spatiotemporal ranges to which data correlation may extend, based on the user’s knowledge of the data or the user’s estimate of the ranges. These values can affect the experimental covariance outcome depending on what the actual correlations are in the spatiotemporal field. The default starting values that SEKS-GUI applies when the screen appears are half of the euclidian spatial distance between the most remote data in the set as the maximum correlation range in space, and half of the maximum data time span as the maximum correlation range in time (not applicable in the purely spatial case).

⇒ The importance of good knowledge of the data is stressed again here. In any case, a guess estimate for the ranges can be applied – or one can simply experiment by inspecting the results for different range values.

⇒ Negative entries in these boxes are prevented by SEKS-GUI. Entries of any size are allowed in this GUI version, as no cap is applied in guessed correlation values. However, in order to prevent entries by mistake, the user is informed with a message when the user-chosen entry exceeds beyond 150% the distance between the data furthest apart in any spatial direction, and beyond 150% the maximum data temporal span (the temporal feature is not applicable in the purely spatial case).

6. The lags sliders and edit boxes can be used to define the number of lags at which the experimental covariance values will be calculated. The user is allowed to ask for covariance estimates at any number between 2 and 22 lags in space and time, as applicable. SEKS-GUI automatically distributes these distances logarithmically across the preset correlation range, thus aiming to obtain a closer monitoring of the covariance behavior as the class sizes get smaller closer to the starting point.

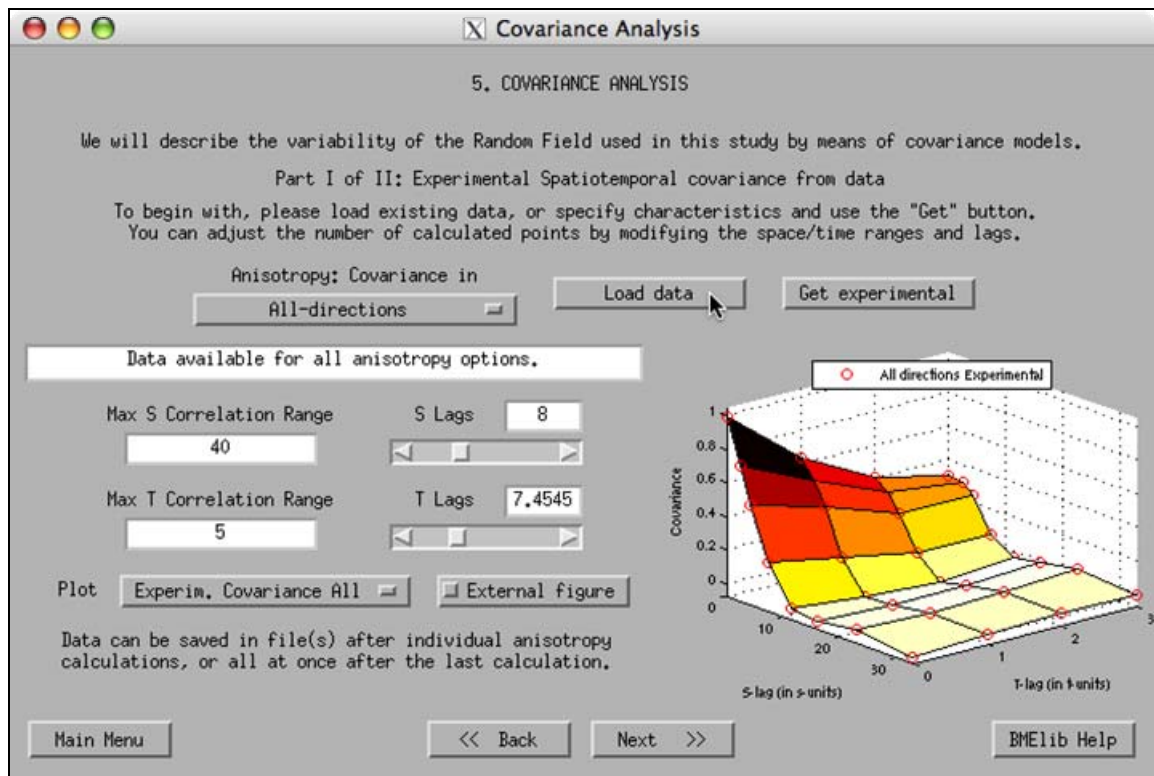


Fig. III.B.9

7. Once experimental covariance data are available to SEKS-GUI, the user can view experimental covariance plots from the drop-down menu (next to “Plot”) that create maps from the covariance information. The user can choose at any time to view maps of:

- Experimental covariance in all-directions (example shown in Fig. III.B.9).
- Experimental covariance in the East-West direction
- Experimental covariance in the North-South direction
- All experimental covariance points in one plot
- All-directions, East-West, or North-South experimental covariances at $s=0$
- All-directions, East-West, or North-South experimental covariances at $t=0$

The last two plot categories are not available in purely spatial cases. If the data for a particular request are not available, a message will appear in the message box on the screen.

★ Use the “External figure” button to display the plot requests that will follow in an external, independent window when activated, or to return to the in-screen display when de-activated (as first shown in paragraph III.B.1 earlier and Fig. III.B.5).

⇒ When the “External figure” button is activated, it enables complete control of the plot by making use of Matlab tools (e.g., axes rotation, renaming, etc.). Also, this feature allows the user to print the particular figure using the independent window menu. For more information on handling plots in separate figure windows please consult Matlab Help.

8. When done, push the “Next” button to proceed.

4. BME “Covariance Analysis” – Part II

- ⇒ The two screens in paragraphs III.B.3 and III.B.4 are combined in 1 screen in the purely spatial case.

In this screen the user is asked to fit a covariance model to the experimental covariance information (Fig. III.B.10). This fitting model may consist of more than one nested models. The current version of SEKS-GUI supports a model fitting approach based on visual inspection, rather than an automated one (e.g., least squares fit). This method depends on the user to perform a visual fit by adjusting the nested models sill and range parameters.

- ⇒ The user cannot proceed to the following screens unless there are covariance model data available to the SEKS-GUI. However, it may occur that you already have results from previous sessions and only want to advance to the visualization screen, i.e., without having to go through covariance modeling again. In this case, the following “acknowledgement” process is necessary: In order to proceed further, you will need to add at least 1 nested model based on step 2 below; there will be no need to make use of that model. Simply choose any model from the “Select a model to add” drop-menu so that the “Models present” counter is other than 0, and push the “Next” button to continue.
- ⇒ When plotting maps with legends, Matlab versions pre-dating the 7.1, R14, Service Pack 3 version may produce legend-related error messages in the main Matlab Command Window. These do not affect the SEKS-GUI execution, and can be ignored.

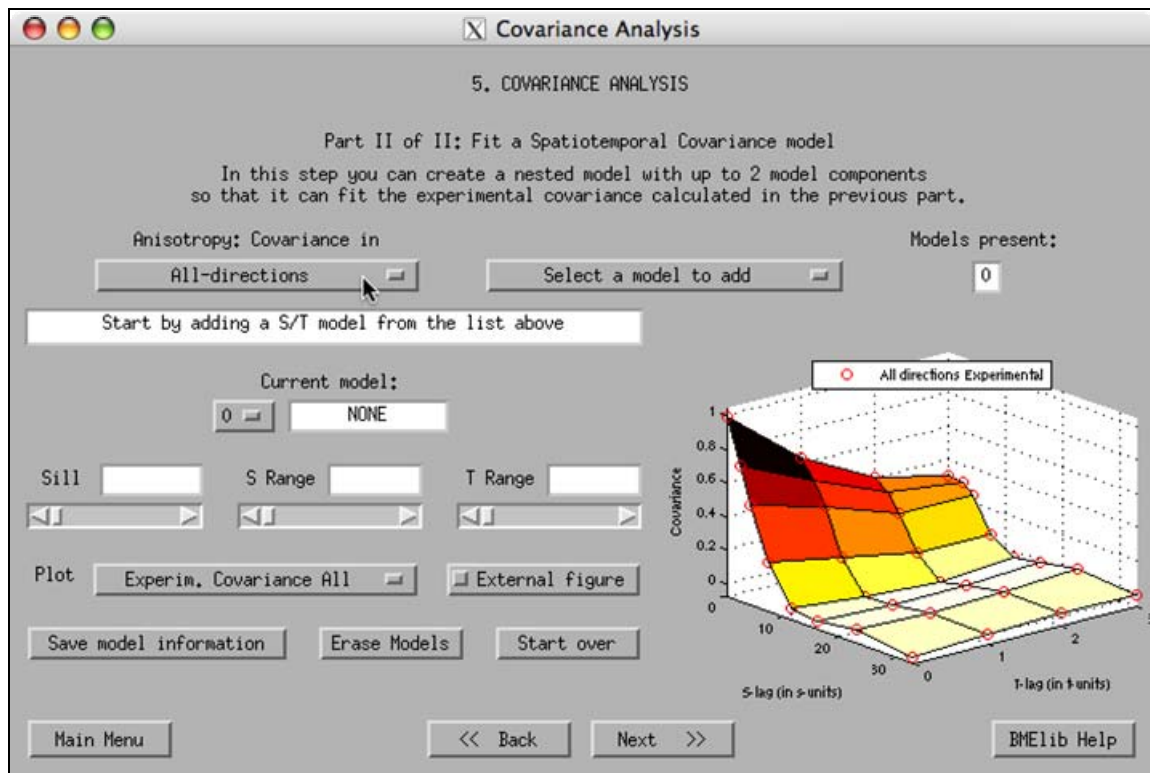


Fig. III.B.10

1. The drop-down menu “Anisotropy: Covariance in” allows the user to project the available experimental covariance information on a graph so that a model can be visually fit on top of it. As in the previous screen, the options allow for covariances in:

- All-directions
- East-West (0° axis)
- North-South (90° axis)

2. Push the “Select a model to add” drop-menu to add a spatiotemporal covariance nested model and proceed with adjusting its parameters. In the present version of SEKS-GUI, available models are separate in space and time, and the user may choose between any combination of Exponential-, Gaussian-, or Spherical-based models for the spatial and temporal components (9 spatiotemporal models available; in the purely spatial case no temporal components are included and there are 3 different spatial models available).

⇒ Every time step 2 is executed, the index of the “Models present” box increases by 1 (Fig. III.B.11). SEKS-GUI allows for the simultaneous presence of a maximum of 2 nested models. In any case, the sill of the fitted model is the sum of the nested ones, and equal to the variance of the data set.

3. The message box under the anisotropy choice drop-menu communicates useful messages to the user and can not be edited.

4. The drop-menu under “Current model:” displays the nested model for which the sill and spatiotemporal range information is currently shown underneath. The “Current model:” message box displays the type of the current nested model and can not be edited. Once you have added 1 or more models, choose from the drop-down menu the index of the nested model for which you would like to adjust the sill and range parameters.

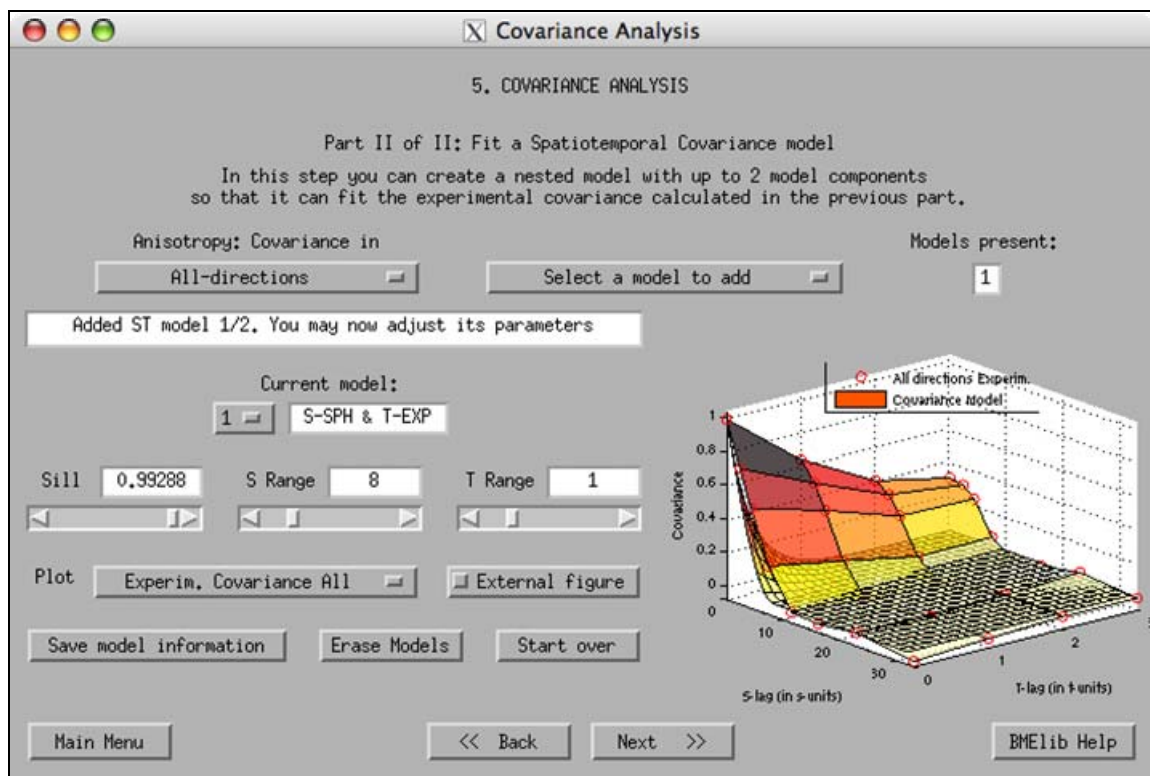


Fig. III.B.11

5. The “Sill” slider and edit box allow you to set sill values for any of the nested models you have in your outcome fitting model. Drag the slider or enter a desired value in the edit box, and instantly see the updated model in the covariance plot (Fig. III.B.12). Any adjustments will apply only to the current nested model, as defined by the user in step 4 above.

- ⇒ In any case, the sill of the outcome fitting model is the sum of the nested ones, and is equal to the variance of the data set. When 2 nested models are present and the sill of the current one is adjusted by the user, the sill of the other nested model is automatically adjusted by SEKS-GUI to comply with this rule.
- ⇒ In the presence of only 1 nested model SEKS-GUI automatically sets the sill to the variance value and the user is not allowed to modify it.

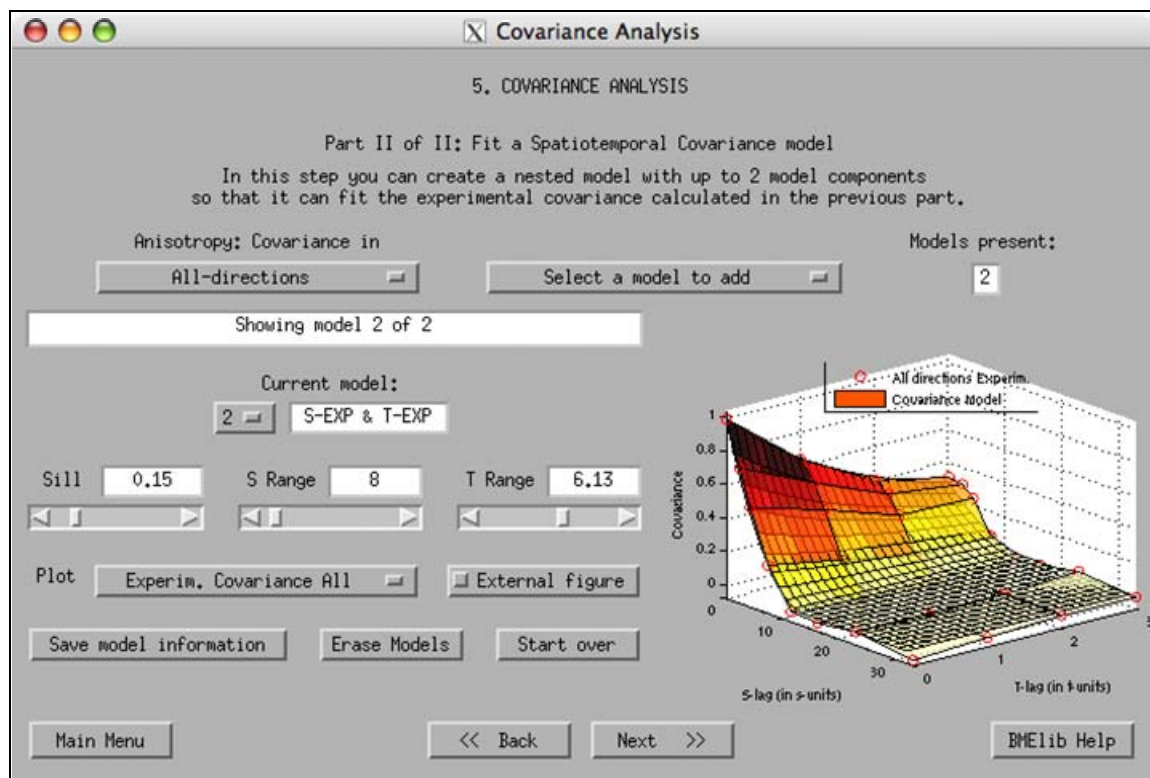


Fig. III.B.12

6. The “S-range” slider and edit box allow you to adjust the model spatial range parameter for any of the nested models you have in your outcome fitting model. Drag the slider or enter a desired value in the edit box, and instantly see the updated model in the covariance plot (Fig. III.B.12). Any adjustments will apply only to the current nested model, as set in step 4 above.

- ⇒ Use of negative entries in these boxes is prevented by SEKS-GUI. Also, this GUI version applies an upper cap in S-range values, and allows for entries up to 1.5 times the distance between the data furthest apart in any spatial direction.

7. (Not applicable in the purely spatial case) The “T-range” slider and edit box allow you to adjust the model temporal range parameter for any of the nested models in your outcome fitting model. Drag the slider or enter a desired value in the edit box, and instantly see the updated model in the covariance plot (Fig. III.B.12). Any adjustments will apply only to the current nested model, as set in step 4 above.

⇒ Use of negative entries in these boxes is prevented by SEKS-GUI. Also, this version of the GUI applies an upper cap in T-range values, and allows for entries up to 2 times the maximum data temporal span.

8. As the user shapes the fitting model, a variety of covariance maps are available from the drop-down “Plot” menu. The user can choose at any time to view maps of:

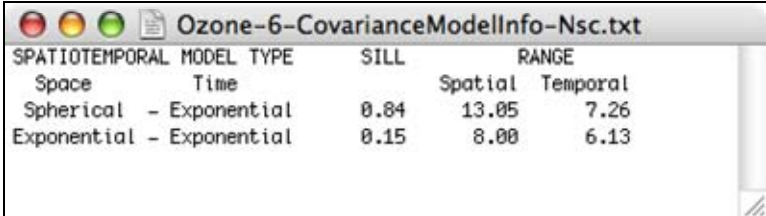
- Experimental covariance in all-directions
- Experimental covariance in the East-West direction
- Experimental covariance in the North-South direction
- Modelled covariance curve only, showing the curve that has been fitted
- Modelled covariance curve and the experimental covariance points
- Experimental and modelled covariances
- Modelled covariance plot at lag s=0
- Modelled covariance plot at lag t=0

The last two plots are not available in the purely spatial case. If the data needed for a particular request are not available, a message will appear in the message box on the screen.

★ Use the “External figure” button to display the plot requests that will follow in an external, independent window when activated, or to return to the in-screen display when de-activated (as first shown in paragraph III.B.1 earlier and Fig. III.B.5).

⇒ When the “External figure” button is activated, it enables complete control of the plot by making use of Matlab tools (e.g., axes rotation, renaming, etc.). Also, this feature allows the user to print the particular figure using the independent window menu. For more information on handling plots in separate figure windows please consult Matlab Help.

9. Push the “Save model information” button after you have fitted a model to the experimental covariance in order to save your nested models sill and range parameters in a text file. Pushing this button will prompt you to choose a location to save this file. The information that you save corresponds to one particular fitted model for future reference; different models can be fitted, and their nested components information can be saved in different text files, as well. An example of such a file’s contents is shown in Fig. III.B.13.



SPATIOTEMPORAL MODEL TYPE		SILL	RANGE	
Space	Time		Spatial	Temporal
Spherical	- Exponential	0.84	13.05	7.26
Exponential	- Exponential	0.15	8.00	6.13

Fig. III.B.13

10. Push the “Erase models” button if you would like to start the fitting process from the beginning (Fig. III.B.14). This action will erase any currently existing nested models and their characteristics, and will allow you to add new models. “No” is the default answer in the pop-up window that appears in this case.

⇒ If you confirm the “Erase models” action after you push the button, any unsaved information regarding existing nested models on the screen will be lost.

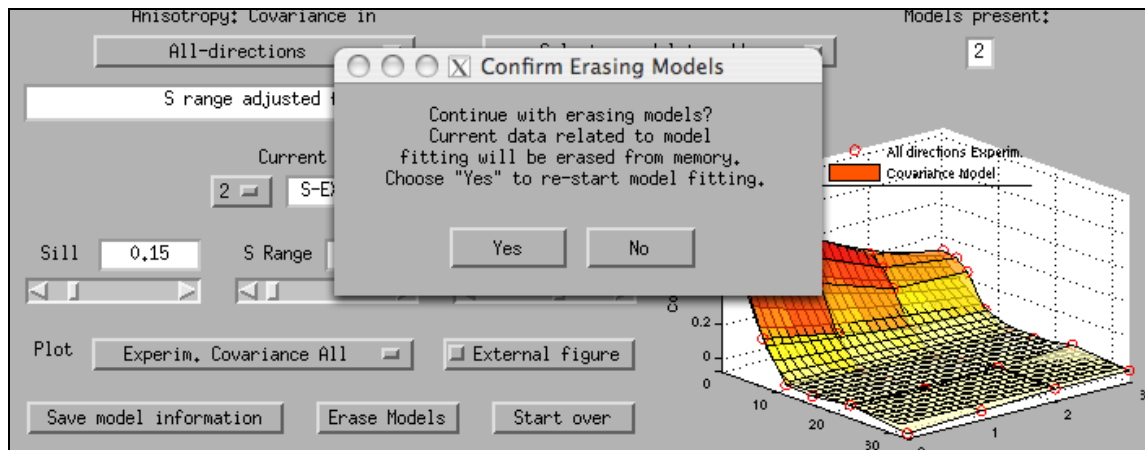


Fig. III.B.14

11. Push the “Start over” button if you would like to start the covariance analysis process from the beginning (Fig. III.B.15). This action will erase any currently existing covariance information (experimental and models) on the screens, and will take you to the previous screen III.B.3 (part I of covariance analysis) or will reload the screen in the purely spatial case. “No” is the default answer in the pop-up window that appears in this case.

⇒ If you confirm the “Start over” action after you push the button, any unsaved information regarding existing experimental covariance or nested models on the screen will be lost.

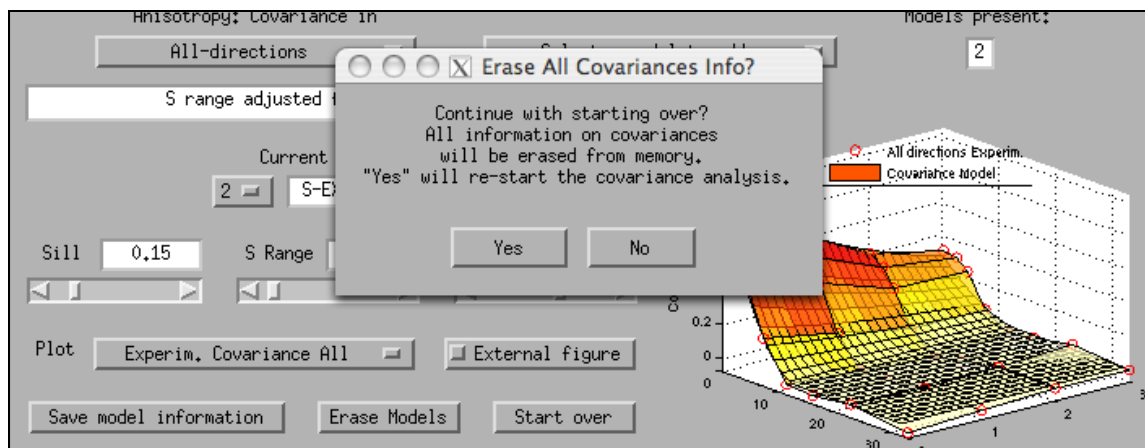


Fig. III.B.15

12. When done, push the “Next” button to proceed.

5. “BME Estimations Wizard”

In this screen the user can select and initialize the type of BME estimations to be performed by BMElib. If you are returning to this screen from the following visualization screen you may proceed forward by pushing the “Next” button.

1. Use the “Select an estimation type” drop-menu to define what type of results you want (Fig. III.B.16). The options are:

- a. BME Mode (the mode of the estimation posterior PDF at each output grid node)
- b. BME Moments (the mean, variance and skewness coefficient of the estimation posterior PDF at each output grid node)
- c. BME PDF (the estimation posterior PDF at each output grid node)
- d. BME Confidence Intervals (the estimation posterior PDF and the user-preset percentile confidence interval at each output grid node)

The above options are ranked with respect to the time required for the calculations, starting with the fastest one and ending with the most time-consuming.

- ⇒ Selection of an estimation type does not initiate the calculations immediately; the button “Begin Estimation” needs to be pushed first.
- ⇒ Note that with the exception of BME Mode, each one of the gradually slower selections (c) and (d) above provides all of the information given by the previous, faster ones. The BME Mode cannot be extracted from the information in (b), (c) or (d). One can have all possible BME estimation outcomes simply by performing the BME Mode and CI estimations.

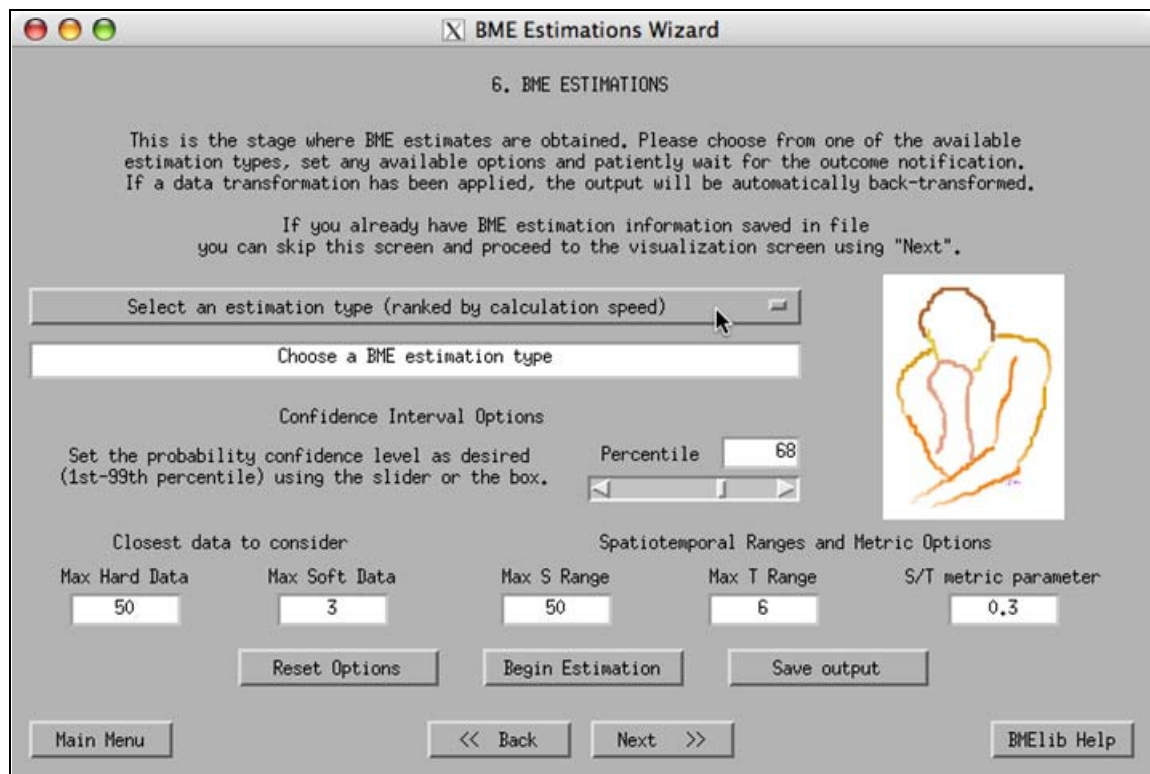


Fig. III.B.16

2. The message box under the estimation type selection drop-menu communicates useful messages to the user and cannot be edited.

3. Use the “Percentile” slider or edit box if you will be performing the BME Confidence Interval estimations, in order to adjust the percentile for which you want results. The default SEKS-GUI value is 68, which corresponds to the 68th percentile.

⇒ Only one percentile level of confidence can be selected for estimations in this version of SEKS-GUI. If you would like the confidence levels at additional different percentiles you will have to repeat the calculations after adjusting the slider or edit box as desired.

4. Adjust the number of closest neighboring data to consider for the estimation. Depending on the availability of hard and soft data, the user can edit the boxes under “Max Hard Data” and “Max Soft Data”. Each one displays the current parameter value for the maximum number of neighboring hard and soft data, respectively, to be taken into account during the calculations at each estimation node. SEKS-GUI sets as default starting values a maximum of 50 hard data and 3 soft data for these parameters (Fig. III.B.16).

⇒ Either of the edit boxes for the “Max Hard Data” and “Max Soft Data” parameters can be edited in the presence of hard and soft data, respectively. If one of these data categories is not present in an investigation, the corresponding box displays “N/A” and cannot be further edited.

⇒ One would obviously like to consider as many data as possible when estimating a value at an unsampled location. However, considering too many data may significantly slow down the estimation process. This is particularly evident when the maximum number of soft data is increased, in the presence of soft data. Estimations may become significantly slower if more than 4-5 soft data are considered at a time, when compared to using fewer soft data. The user may adjust the default values as desired prior to the beginning of the estimation process, or can compare the results by trying repeated estimations using different parameter values.

⇒ The edit boxes for the “Max Hard Data” and “Max Soft Data” parameters accept positive integer numbers. Any different entry is unacceptable and produces an error message window.

⇒ Any positive integer number is accepted in the “Max Hard Data” edit box. However, if the user sets the parameter to more than 500, the user is warned about the seemingly large number as a precaution to prevent potential typing errors.

⇒ Any positive integer number is accepted in the “Max Soft Data” edit box. However, if the user sets the parameter to more than 4, the user is warned about the seemingly large number as a precaution to prevent potential typing errors.

5. Adjust the spatial range parameter for the estimation. This parameter shows in the edit box under “Max S Range”, and regulates the maximum spatial distance from the current estimation location within which BME will search for data neighbors to contribute to the estimation. SEKS-GUI sets as a default starting value for this parameter a number equal to 125% of the maximum spatial correlation value, as set earlier in paragraph III.B.3.5.

⇒ The “Max S Range” edit box accepts positive integer numbers. Any different entry is unacceptable and produces an error message window.

⇒ Any positive integer number is accepted in the “Max S Range” edit box. However, if the user sets the parameter to more than 150% of the output grid largest side size, the user is warned about the seemingly large number as a precaution to prevent potential typing errors.

6. (Not applicable in the purely spatial case) Adjust the temporal range parameter for the estimation. This parameter shows in the edit box under “Max T Range”, and regulates the maximum temporal distance from the current estimation instance within which BME will search for data neighbors to contribute to the estimation. SEKS-GUI sets as a default starting value for this parameter (Fig. II.C.32) a number equal to 125% of the maximum temporal correlation value, as set earlier in paragraph III.B.3.5. In the purely spatial case the corresponding box displays “N/A” and cannot be further edited.

- ⇒ The “Max T Range” edit box accepts positive integer numbers. Any different entry is unacceptable and produces an error message window.
- ⇒ Any positive integer number is accepted in the “Max S Range” edit box. However, if the user sets the parameter to more than the time span of the data set, the user is warned about the seemingly large number as a precaution to prevent potential typing errors.

7. (Not applicable in the purely spatial case) Adjust the spatiotemporal metric parameter for the estimation. This parameter shows in the edit box under “S/T metric parameter”, and is used to establish a spatiotemporal distance between two spatiotemporal coordinates by means of the relationship:

$$[\text{Spatiotemporal distance}] = [\text{Spatial distance}] + [\text{S/T Metric Parameter}] * [\text{Temporal distance}]$$

There are no guidelines for setting this parameter. The user may experiment with different values, if desired, in order to define the case-specific spatiotemporal distance as a function of the distances in space and time. SEKS-GUI sets a default starting value for this parameter as 0.3 (Fig. III.B.16). In the purely spatial case the corresponding box displays “N/A” and cannot be further edited.

- ⇒ The “S/T metric parameter” edit box accepts positive integer numbers. Any different entry is unacceptable and produces an error message window.

8. Once you have selected the BME estimation type from step 1 above and set the estimation parameters as applicable and described in steps 3-7 earlier, push the “Begin Estimation” button to start the calculations.

- ⇒ Be patient and wait until the calculations come to an end. Matlab and the SEKS-GUI cannot respond during the calculation time to any other commands. Please refrain from pushing other SEKS-GUI buttons at this time, as these actions are queued and may result in unwanted events or errors after the calculations are done. Calculations may take a lot of time, depending on the volume of data in the data set and the hardware you are using. The message box displays for your convenience at any instance the calculations progress (Fig. III.B.17). However, no results can be viewed before the successful termination of the calculations and prior to advancing to the following visualization screen.
- ⇒ If you need to terminate the calculations before the completion of the estimation process, you can click on the Matlab Command Window outside the GUI window, and subsequently press the sequence of keyboard keys: <CTRL-C>. The Matlab Command Window will then display some error messages related to the premature termination of the calculations. Back in the GUI window you will notice the progress counter in the message box has stopped, which indicates that the estimations are halted. You are free to continue from there.
- ⇒ The counter that displays the current estimation point may come to a halt before the “Estimations completed” message appears in the message box. If error messages appear in the main Matlab Command Window, or the estimations do not seem to proceed by inspecting the counter indications, this suggests that the estimation process has been interrupted. SEKS-GUI cannot recognize this event explicitly. If such an

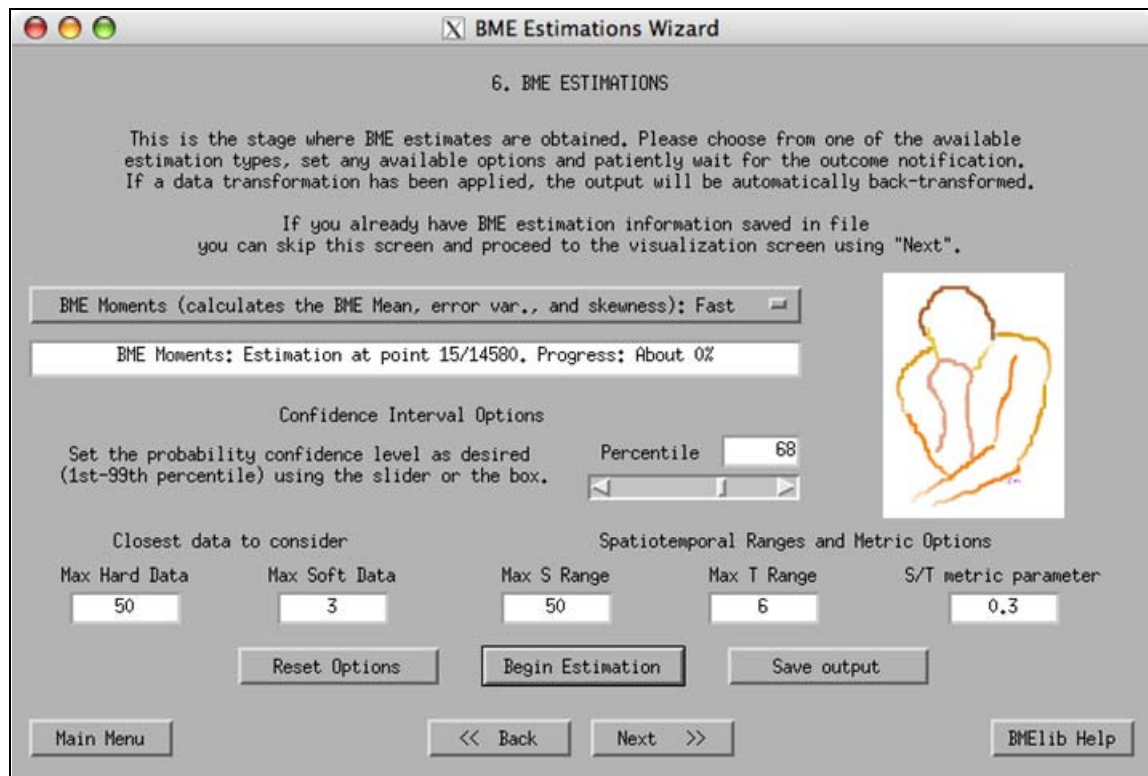


Fig. III.B.17

interruption occurs, this could be an indication that some estimation parameters need to be re-adjusted. Typical reasons for interruption would include:

- Singularities in the covariance matrices, in which case a revision of the covariance model is suggested by returning to the previous screen 12.
 - Limited amount of data for the estimation, in which case the use of more hard data is suggested. You can do this by increasing the existing number in the “Max Hard Data” edit box and restarting the estimations.
- ⇒ After the BME estimations have taken place at all of the output nodes, SEKS-GUI calculates the moments (mean, variance and skewness) of the estimation PDFs, if applicable. Consequently, SEKS-GUI back-transforms automatically all the estimation information into the original-space, if transformation has been applied to the initial data set. All of the material is eventually arranged in variables to be saved in an output file or to be used in the following screen for the generation of maps.
- ⇒ If a transformation has been applied to the data and moments have been calculated using any of the choices (b), (c) or (d) of paragraph 1, the following apply for each estimation node:
- The first moment (mean) of the raw results in the original space is the backtransformed mean of the BME mean in the transformed space.
 - The second moment (variance) in the original space is based on the BME variance in the transformed space, but it is not the direct backtransform of the BME variance. In particular, the variance value in the original space is rather a measure of the variance in the transformed space. This measure is obtained by backtransforming the standard deviation value from the transformed into the original space.

- The third moment (skewness of the PDF) cannot be meaningfully backtransformed into the original space, and it is therefore not available in the original space when the user calculates only the BME Moments – choice (b) in paragraph 1. If the user has selected to obtain the BME PDF or Confidence Intervals results – choices (c) or (d) in paragraph 1, respectively – then the BME output PDF is present and can be backtransformed into the original space. In this case, skewness values are available in the original space and they are calculated based on the backtransformed PDF.
- ⇒ At the end of the calculations you are prompted to save the outcome data in a file. You can decline and choose to save all the results at the end after you run additional estimations, or you can proceed to save the results from all estimations to that point by pushing the “Save output” button.

9. Push the “Save output” button to store the BME estimation results in a file. It is strongly suggested to save them in a suitably named file, as you may later wish to return or re-run this study. The data are saved in a Matlab format which has the ending “.mat” at the folder you specify, and cannot be viewed independently unless they are loaded within the Matlab environment.

- ⇒ Once the user leaves the screen without saving, BME estimation results may be lost upon returning to this screen. The user is not offered the option to save these results at a later point or on a different screen.

10. When done, push the “Next” button to proceed.

Section C – GBME Analysis Screens in SEKS-GUI

1. GBME “Exploratory Analysis” – Parts II and III

A. Introduction

GBME analysis is based on a different approach than BME; instead of estimating the spatiotemporal process from the global aspect, GBME performs the estimation at each local neighborhood separately. At each estimation point, the spatial and temporal orders, ν and μ , respectively, spatiotemporal increments and generalized covariance models will be automatically assessed based on its local information. In this sense, GBME analysis is a more data-oriented approach than BME analysis.

⇒ The quality of the estimation of GBME analysis at each locality highly depends on the quantity and quality of the data in its neighborhood.

Due to its nature of global spatiotemporal mean trend and covariance estimations, BME analysis can borrow the information further far from the estimation, which may introduce a smoothing effect on the estimates. In that sense GBME may perform better estimation without exhibiting the effect of smoothing; however, BME can perform more reasonable estimation in case of a data scarcity in the neighborhood.

Given the above outline of the GBME conceptual approach, an important factor for GBME is data availability. GBMElib performs optimally in a uniform distribution of the data in space and time. The present exploratory stage serves so that the SEKS-GUI users can check the spatial and temporal distribution of the hard and soft data in their set.

Visualization of the data population in this screen relies on the representation of data as point values. Hard data are point values, but this is not the case with soft data. For the purposes of this screen SEKS-GUI represents any soft data by considering the means of the soft distributions (or the middle points of the soft intervals). This approximation is independent from the GBME estimation stage where soft information is considered in full as provided by the user.

B. Instructions

1. On the left-hand side of the screen statistics on the data set are displayed (Fig. III.C.1). Depending on the presence of hard and/or soft data, these statistics are based on the hard data and the soft data approximations. The user can ask for the statistics of subsets at individual temporal instances in the spatiotemporal case, and in that case the statistics at a given instance t will correspond to the data population of t .

⇒ If the user asks for statistics that are unavailable, then the data statistics boxes display the message “N/A”.

⇒ The data statistics boxes cannot be edited by the user. They are only available for display of values.

2. The message box on the upper part of the screen communicates useful messages to the user and cannot be edited.

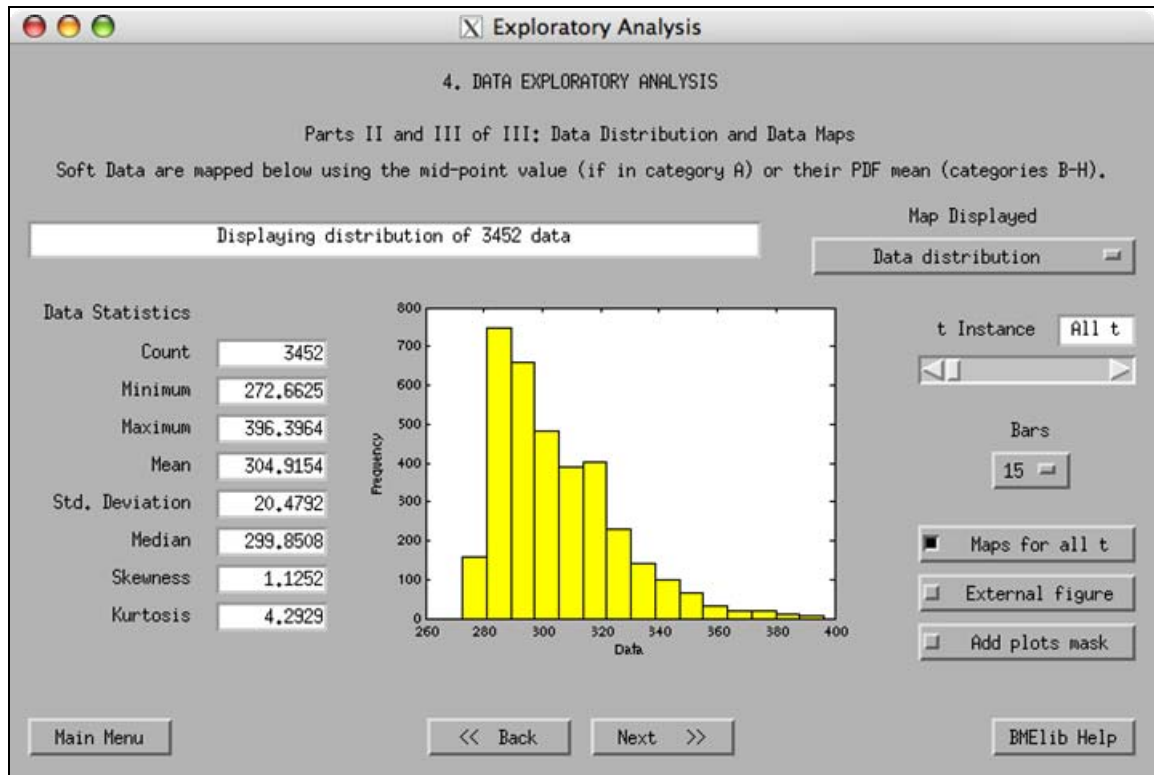


Fig. III.C.1

3. The user can plot a variety maps from the drop-down menu (under “Map Displayed”). At any time the user can choose to view maps of:

- All data locations
- Hard data locations
- Soft data locations
- Markerplots of all data (hard data and soft data approximations)
- Colorplots of all data (hard data and soft data approximations)
- Data distribution

If data necessary for a particular request are not available, a message will appear in the message box on the screen.

★ Use the “Bars” drop-menu to display the histogram plots with more or fewer bars as desired.

★ Use the “t-Instance” slider (or write a suitable number in the “t-Instance” box) to view maps at a particular instance in time, as in Fig. III.C.2 (not available in purely spatial cases). In spatial-only investigations the “t Instance” slider and box are disabled.

★ Activate or de-activate as desired the “Maps for all t” button to see time-aggregated maps of data locations and distributions or maps at user-selected instances, respectively. In spatial-only investigations the button is disabled.

★ Use the “External figure” button to display the plot requests that will follow in an external, independent window when activated (as in fig. III.B.5), or to return to the in-screen display when de-activated.

⇒ When the “External figure” button is activated, it enables complete control of the plot by making use of Matlab tools (e.g., axes rotation, renaming, etc.). Also, this feature allows

the user to print the particular figure using the independent window menu. For more information on handling plots in separate figure windows please consult Matlab Help.

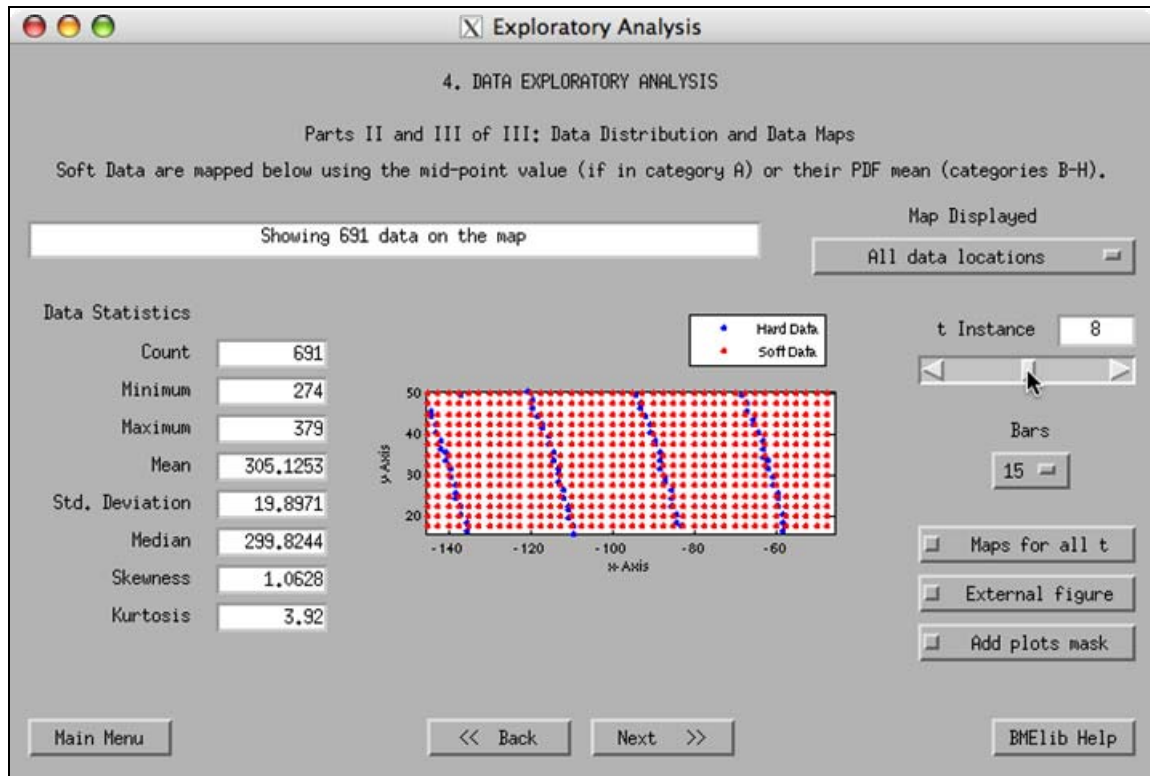


Fig. III.C.2

4. A more advanced feature for map presentations is plot masking, and requires some knowledge of Matlab programming. This is a useful feature if you would like to show the results that appear in part of your output area (e.g., by masking out the portion of a map outside a country's borders). It is more advanced because it requires that the user provides a Matlab ".m" code file with the masking information. If you can program in Matlab, you can create a map that produces the desired mask over the output grid area using suitable coordinates (see also paragraph V.3 in this manual for a basic utility provided with SEKS-GUI to assist you in the mask creation). Upon activation of the "Add mask to plots" button the user is prompted to locate such a masking code Matlab file within the user's filesystem. If you push this button accidentally you can cancel the file search. If the button is activated you can push it again to de-activate it.

5. When done, push the "Next" button to proceed.

2. “GBME Estimations Wizard”

A. Introduction

Based on the GBME conceptual approach, at each estimation point the GBME analysis produces an estimate based on the information within the local neighborhood. Steps that are explicit stages in the BME analysis prior to the estimation are integrated in this case within the GBME type of analysis. In particular, these steps are:

- a. Estimation of the spatial and temporal heterogeneity orders ν and μ , respectively, for each output node location. These are derived based on the availability of local information in a neighborhood, and represent the variability of the local spatiotemporal trend. The ν and μ orders are also used in the GBME covariance analysis, as different covariance models are defined using different combinations of the ν and μ values – see Tables III.C.1 and III.C.2.
- b. Estimation of the spatiotemporal increments, which are derived using the ν and μ values. The GBME theory uses the spatiotemporal increments as means to reduce the variability of the random field we investigate.
- c. Estimation of the generalized covariance. The difference between this and the covariances used in the BME analysis is that the former are estimated locally and using the random field spatiotemporal increments, whereas the BME covariance analysis yields the global covariance from the random field itself.
- d. Estimations at the output grid nodes using the above information in addition to the of the BME kernel function.

GBME operates on a local level within the data, therefore the spatial and temporal orders, the spatiotemporal increments, and generalized covariance models can vary throughout the domain of interest. To achieve accuracy in GBME analysis, the user has to provide enough data for each estimation neighborhood. If the amount of data is not enough for a certain neighborhood, the availability of space-time models that are used internally for the GBME analysis will be reduced. The following tables III.C.1 and III.C.2 show in the spatiotemporal case and the spatial-only case, respectively, the minimum amount of data required in a neighborhood for different spatial and temporal orders combinations.

Spatial order ν	Temporal order μ	Minimum number of data required
0	0	2
0	1	3
0	2	4
1	0	4
1	1	7
1	2	10
2	0	7
2	1	13
2	2	19

Table III.C.1: Minimum number of data required for S/T GBME analysis (2-D in space, 1-D in time)

Spatial order ν	Minimum number of data required
0	2
1	4
2	7

Table III.C.2: Minimum number of data required for purely spatial GBME analysis (2-D in space).

In GBME, mean trends in random processes are represented by space-time polynomials. An order of n (in either space or time) means that it is required to apply $n+1$ times a differential operator (in the spatial or temporal domain, respectively) to remove the trend from the original random process. As an example, let us consider a spatiotemporal process that is found to have an order of $\nu/\mu=0/0$ for a given location. This means that the local trend in the process is linear in both space and time. Upon removal of the linear trend using a suitable differential operator once, the process becomes homogeneous in space and stationary in time, and is now characterized by the order pair $\nu/\mu=-1/-1$. These steps are performed implicitly by GBME in the present stage.

B. Instructions

1. The message box under the estimation type selection drop-menu communicates useful messages to the user and cannot be edited.
2. Adjust the number of closest neighboring data to consider for the estimation. Depending on the availability of hard and soft data, the user can edit the boxes under “Max Hard Data” and “Max Soft Data”. Each one displays the current parameter value for the maximum number of neighboring hard and soft data, respectively, to be taken into account during the calculations at each estimation node. SEKS-GUI sets as default starting values a maximum of 50 hard data and 10 soft data for these parameters (Fig. III.C.3).
 - ⇒ As analyzed earlier in this section, the amount of data in each neighborhood plays a very important role in GBME analysis. The user should set carefully the “Max Hard Data” and “Max Soft Data” parameters in conjunction with the “Min Data” one that follows in step 3 below, in order to avoid data scarcity in neighborhoods that can lead to potential numerical divergence.
 - ⇒ Either of the edit boxes for the “Max Hard Data” and “Max Soft Data” parameters can be edited in the presence of hard and soft data, respectively. If one of these data categories is not present in an investigation, the corresponding box displays “N/A” and cannot be further edited.
 - ⇒ One would obviously like to consider as many data as possible when estimating a value at an unsampled location. However, considering too many data may significantly slow down the estimation process. This is particularly evident when the maximum number of soft data is increased, in the presence of soft data. Estimations may become significantly slower the more soft data are considered at a time, when compared to using fewer soft data. The user may adjust the default values as desired prior to the beginning of the estimation process, or can compare the results by trying repeated estimations using different parameter values.

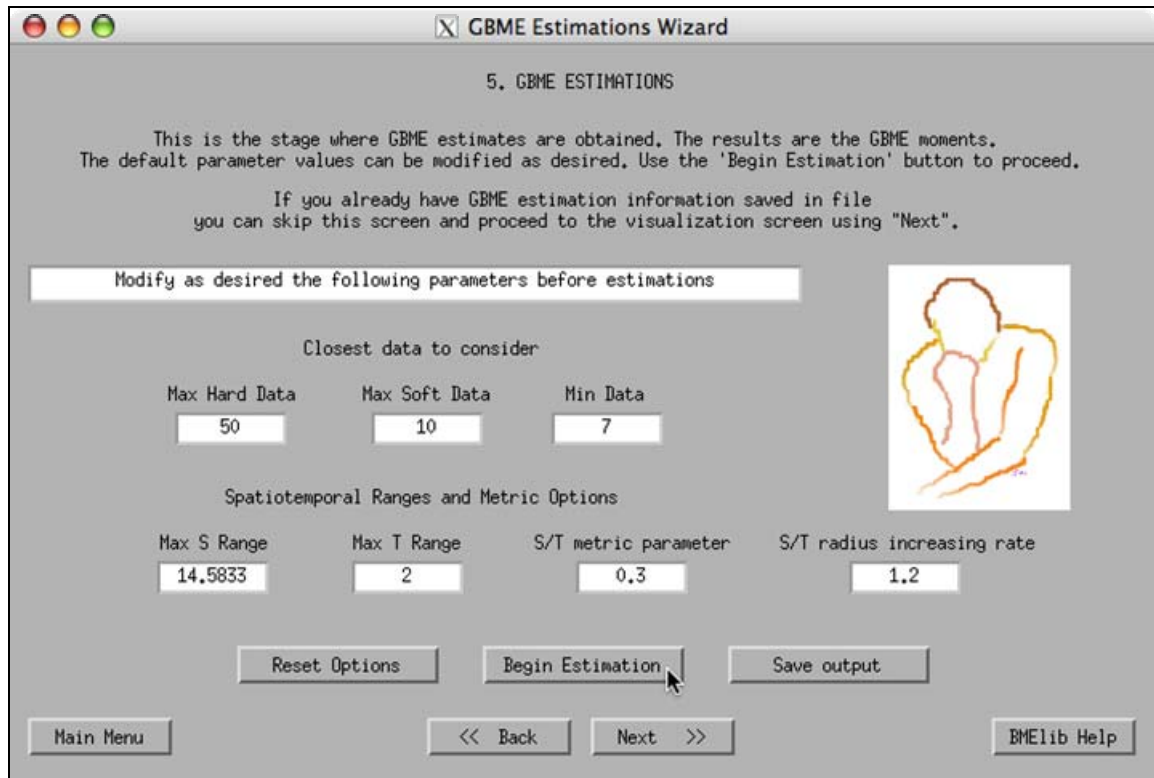


Fig. III.C.3

- ⇒ The edit boxes for the “Max Hard Data” and “Max Soft Data” parameters accept positive integer numbers. Any different entry is unacceptable and produces an error message window.
- ⇒ Any positive integer number is accepted in the “Max Hard Data” edit box. However, if the user sets the parameter to more than 500, the user is warned about the seemingly large number as a precaution to prevent potential typing errors.
- ⇒ Any positive integer number is accepted in the “Max Soft Data” edit box. However, if the user sets the parameter to more than 15, the user is warned about the seemingly large number as a precaution to prevent potential typing errors.

3. Adjust the number of minimum data to be considered for the GBME estimation. The tables III.C.1 and III.C.2 above suggest the minimum data requirements in the spatiotemporal and the purely temporal cases, respectively. The default value is 7, which, based on Table III.C.1 allows for 6 different models (based on different combinations of heterogeneity orders) to be considered in the spatiotemporal GBME analysis.

- ⇒ As analyzed earlier in this section, the amount of data in each neighborhood plays a very important role in GBME analysis. The user should set carefully the “Min Data” parameter in conjunction with the “Max Hard Data” and “Max Soft Data” ones from step 2 above, in order to avoid data scarcity in neighborhoods that can lead to potential numerical divergence.
- ⇒ The edit box for the “Min Data” parameter accepts any positive integer number above the default value of 7. Any different entry is unacceptable and produces an error message window.

4. Adjust the spatial range parameter for the estimation. The data for the local estimation should be inside the specified spatial range. This parameter shows in the edit box under “Max S Range”, and regulates the maximum spatial distance from the current estimation location within which GBME will search for data neighbors to contribute to the estimation. SEKS-GUI defines the default maximum spatial correlation range semi-arbitrarily, as being half of the euclidian spatial distance between the most remote data in the user data set. Based on this definition, the default Max S Range is a number equal to 125% of the maximum spatial correlation value.

- ⇒ The “Max S Range” edit box accepts positive integer numbers. Any different entry is unacceptable and produces an error message window.
- ⇒ Any positive integer number is accepted in the “Max S Range” edit box. However, if the user sets the parameter to more than 150% of the output grid largest side size, the user is warned about the seemingly large number as a precaution to prevent potential typing errors.

5. (Not applicable in the purely spatial case) Adjust the temporal range parameter for the estimation. The data for the local estimation should be inside the specified temporal range. This parameter shows in the edit box under “Max T Range”, and regulates the maximum temporal distance from the current estimation instance within which GBME will search for data neighbors to contribute to the estimation. SEKS-GUI defines the default maximum temporal correlation range semi-arbitrarily, as being half of the maximum data time span in the user data set. Based on this definition, the default Max T Range is a number equal to 125% of the maximum temporal correlation value.

In the purely spatial case the corresponding box displays “N/A” and cannot be further edited.

- ⇒ The “Max T Range” edit box accepts positive integer numbers. Any different entry is unacceptable and produces an error message window.
- ⇒ Any positive integer number is accepted in the “Max S Range” edit box. However, if the user sets the parameter to more than the time span of the data set, the user is warned about the seemingly large number as a precaution to prevent potential typing errors.

6. (Not applicable in the purely spatial case) Adjust the spatiotemporal metric parameter for the estimation. This parameter shows in the edit box under “S/T metric parameter”, and is used to establish a spatiotemporal distance between two spatiotemporal coordinates by means of the relationship:

$$[\text{Spatiotemporal distance}] = [\text{Spatial distance}] + [\text{S/T Metric Parameter}] * [\text{Temporal distance}]$$

There are no guidelines for setting this parameter. The user may experiment with different values, if desired, in order to define the case-specific spatiotemporal distance as a function of the distances in space and time. SEKS-GUI sets a default starting value for this parameter as 0.3. In the purely spatial case the corresponding box displays “N/A” and cannot be further edited.

- ⇒ The “S/T metric parameter” edit box accepts positive integer numbers. Any different entry is unacceptable and produces an error message window.

7. Adjust the spatiotemporal radius increasing rate in the edit box “S/T radius increasing rate”. This parameter is the multiplier for the original spatial and temporal radii, i.e., the Max S range and Max T range parameters set in steps 4 and 5 above, respectively. The purpose of this parameter is to set an automatic increase of these ranges when necessary, in order to include the required amount of minimum data (i.e., the Min Data number defined in step 3 earlier) into the GBME analysis.

- ⇒ The spatiotemporal radius increasing rate must be a positive real value greater than 1. Any different entry is unacceptable and produces an error message window.

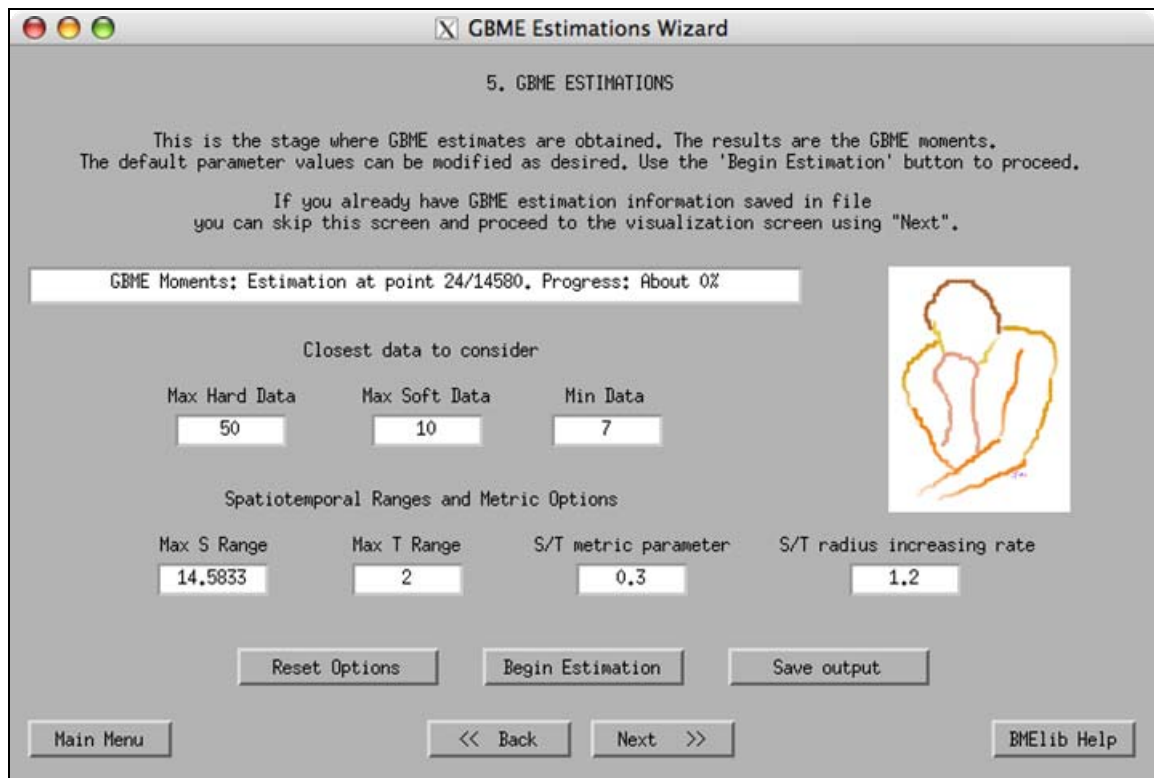


Fig. III.C.4

8. Once you have set the estimation parameters as desired and described in steps 2-7 earlier, push the “Begin Estimation” button to start the GBME calculations.

- ⇒ Be patient and wait until the calculations come to an end. Matlab and the SEKS-GUI cannot respond during the calculation time to any other commands. Please refrain from pushing other SEKS-GUI buttons at this time, as these actions are queued and may result in unwanted events or errors after the calculations are done. Calculations may take a lot of time, depending on the volume of data in the data set and the hardware you are using. The message box displays for your convenience at any instance the calculations progress (Fig. III.C.4). However, no results can be viewed before the successful termination of the calculations and prior to advancing to the following visualization screen.
- ⇒ If you need to terminate the calculations before the completion of the estimation process, you can click on the Matlab Command Window outside the GUI window, and subsequently press the sequence of keyboard keys: <CTRL-C>. The Matlab Command Window will then display some error messages related to the premature termination of the calculations. Back in the GUI window you will notice the progress counter in the message box has stopped, which indicates that the estimations are halted. You are free to continue from there.
- ⇒ The counter that displays the current estimation point may come to a halt before the “Estimations completed” message appears in the message box. If error messages appear in the main Matlab Command Window, or the estimations do not seem to proceed by inspecting the counter indications, this suggests that the estimation process has been interrupted. SEKS-GUI cannot recognize this event explicitly. If such an

interruption occurs, this could be an indication that some estimation parameters need to be reviewed. In this case it is strongly suggested to review the data set for scarcity of data in neighborhoods, and to retry the estimations by moderately re-adjusting upwards the parameters in this screen.

- ⇒ After the GBME estimations of the moments have taken place at all of the output nodes, SEKS-GUI arranges the results in variables to be saved in an output file or to be used in the following screen for the generation of maps.
- ⇒ At the end of the calculations you are prompted to save the outcome data in a file. You can decline and choose to save all the results at the end after you run additional estimations, or you can proceed to save the results from all estimations to that point by pushing the “Save output” button.

9. Push the “Save output” button to store the GBME estimation results in a file. It is strongly suggested to save them in a suitably named file, as you may later wish to return or re-run this study. The data are saved in a Matlab format which has the ending “.mat” at the folder you specify, and cannot be viewed independently unless they are loaded within the Matlab environment.

- ⇒ Once the user leaves the screen without saving, GBME estimation results may be lost upon returning to this screen. The user is not offered the option to save these results at a later point or on a different screen.

10. When done, push the “Next” button to proceed.

Section D – Visualizations with SEKS-GUI

1. “Visualization Wizard”

In this screen SEKS-GUI offers a bundle of map options to display the BME estimation results after the mean trend has been restored (where applicable). This screen can be accessed either from the estimation wizard screens (paragraph III.B.5 for the BME analysis and paragraph III.C.2 for the GBME analysis), or directly from the “Choose a Task” screen (paragraph III.A.2) if an estimation output file is available from previous investigations.

1. The message box at the top communicates useful messages to the user and cannot be edited. Upon loading this screen, the message box will indicate whether there are any estimation output data available in memory. If no such output is available, the user is prompted to load a previously saved estimations output Matlab file (created in steps 8 or 9 of the estimations screens in either BME or GBME analysis), as shown in Fig. III.D.1.

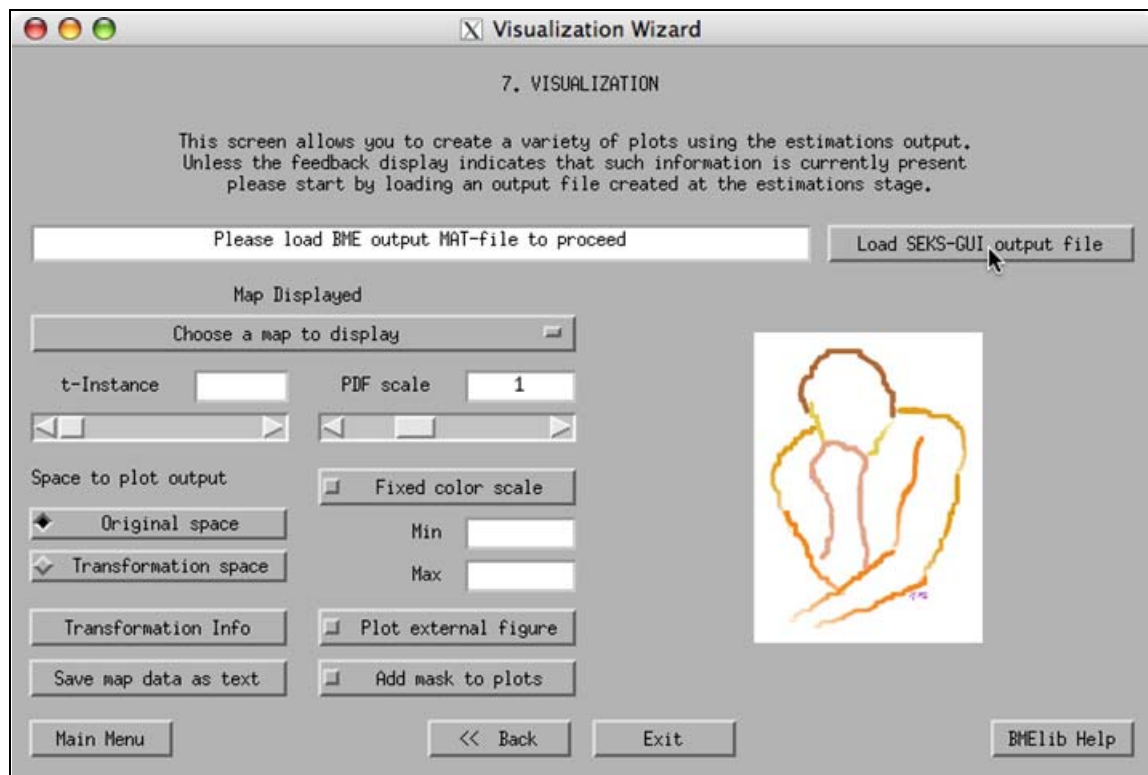


Fig. III.D.1

2. By pushing the “Load SEKS-GUI output file” the user will be asked to locate in the filesystem a Matlab-formatted file ending in “.mat” with previously saved estimations output (created in steps 8 or 9 of the estimations screens in either BME or GBME analysis). The action can be cancelled from the pop-up window. The user can repeatedly load files with estimation information from different investigations. When doing this, the data most recently loaded replace the corresponding data stored in the SEKS-GUI memory up to that point.

- ⇒ If the loaded data correspond to GBME analysis results, then the screen buttons “Transformation space”, “Transformation Info”, and the PDF scale slider and edit box are automatically disabled.

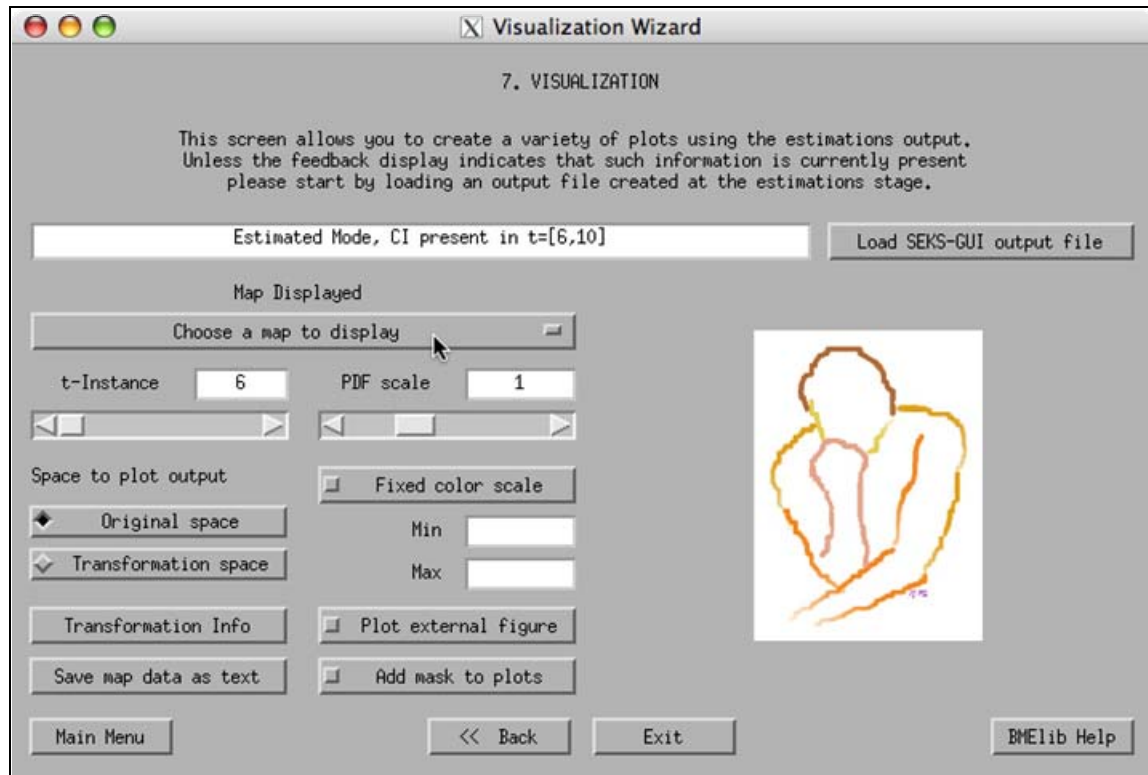


Fig. III.D.2

3. Once estimation output data are available to the SEKS-GUI, the user can choose from a series of maps from the drop-down menu under “Map Displayed” (Fig. III.D.2). Depending on the available results, the user can choose at any time to view an abundance of different maps including: (the type of results necessary for the particular map is printed in *italic*)

- The mean of the estimation posterior PDF at each output grid node – see Fig. III.D.3. *BME (Moments, PDF, or Confidence Interval) or GBME results required.*
- The mode of the estimation posterior PDF at each output grid node. *BME Mode results required.*
- The estimation error variance (the variance of the estimation posterior PDF at each output grid node). *BME (Moments, PDF or Confidence Interval) or GBME results required.*
- The estimation standard deviation (the standard deviation of the estimation posterior PDF at each output grid node). See Fig. III.D.7 for an example. *BME (Moments, PDF or Confidence Interval) or GBME results required.*
- The skewness of the estimation posterior PDF at each output grid node. *BME (Moments, PDF or Confidence Interval) or GBME results required.*
- The actual probability density functions (PDFs) produced by the BME estimations at pre-selected output locations. The PDFs are projected vertically on a map of the output grid (Fig. III.D.4). A selection has been made among all the estimation points

- to present a map of estimation PDFs without cluttering the graph. This version of SEKS-GUI supports display of PDFs at these pre-selected locations throughout the output grid. *BME PDF or Confidence Interval results required.*
- The size of the BME estimations confidence intervals (at the user-selected interval level as chosen on Screen 13, where 68% is the default). In particular, at each output grid node the value mapped is the difference between the variable values at the limits of the confidence interval. This map displays for each estimation location the width of variable values within which the estimate is expected to be found at the selected confidence level. *BME Confidence Interval results required.*
 - The map of the lower limit values of the BME estimations confidence intervals (at the user-selected interval level as chosen on Screen 13, where 68% is the default) at each output grid node. *BME Confidence Interval results required.*
 - The map of the upper limit values of the BME estimations confidence intervals (at the user-selected interval level as chosen on Screen 13, where 68% is the default) at each output grid node. *BME Confidence Interval results required.*
 - The map of the BME estimation PDF value at the confidence interval limits at each output grid node. *BME Confidence Interval results required.*
 - The GBME map of the spatial order ν at each estimation node. ν is an integer that can take the values 0, 1, or 2. *GBME results required.*
 - The GBME map of the temporal order μ at each estimation node – see Fig. III.D.7. μ is an integer that can take the values 0, 1, or 2. *GBME results required.*
 - The GBME map of the difference $\nu - \mu$ at each estimation node (fig. III.D.9). The difference $\nu - \mu$ can take integer values in the range [-2, 2]. *GBME results required.*
- If data necessary for a particular request are not available, a message will appear in the

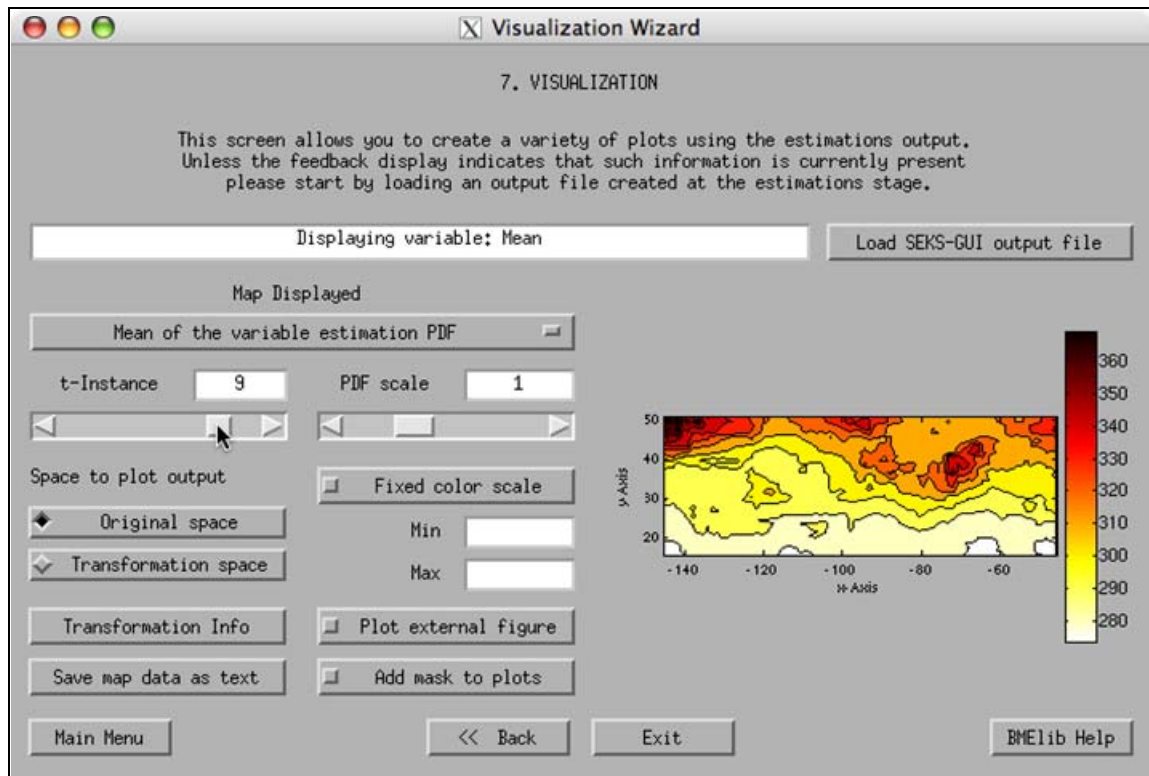


Fig. III.D.3

message box on the screen.

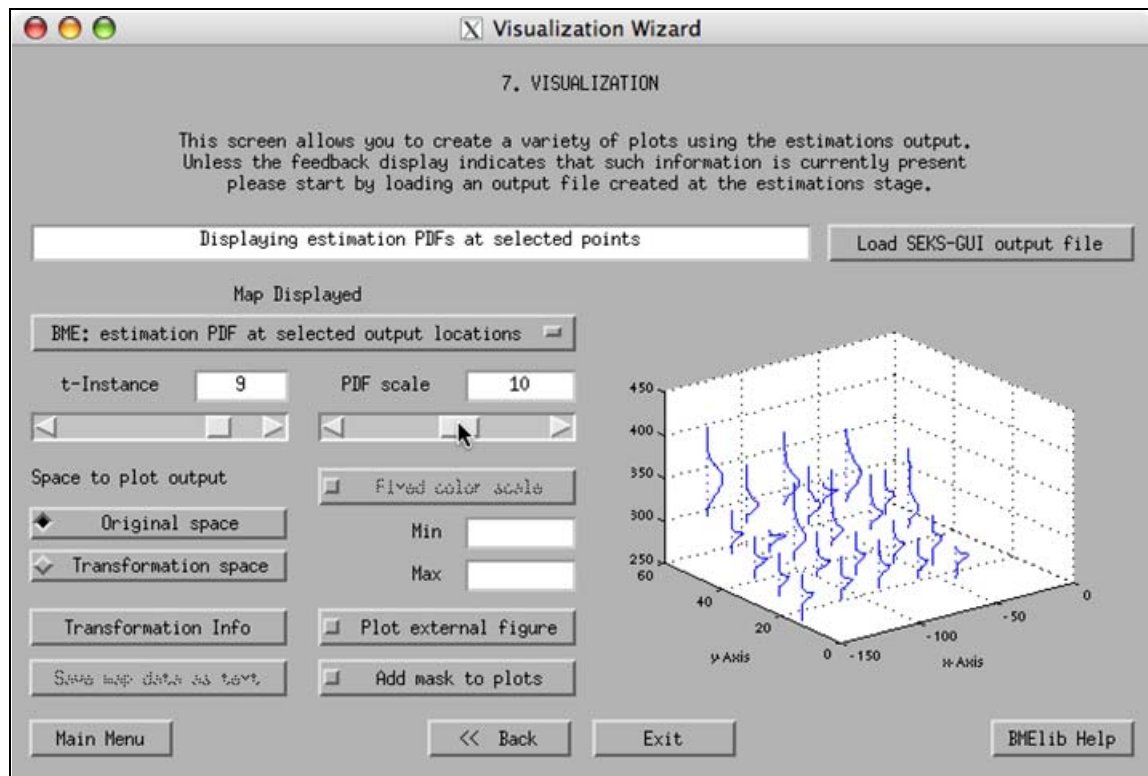


Fig. III.D.4

- ★ Use the “t-Instance” slider (or write a suitable number in the “t-Instance” box) to view maps at any instance in time from the ones designated in the output grid specifications (Fig. III.D.3).
 - ⇒ The “t-Instance” slider and the edit box are disabled in the purely spatial cases.
- ★ Use the “PDF scale factor” slider (the “PDF scale factor” box is read-only) to scale the actual size of the displayed PDFs on the graph (Fig. III.D.4). PDFs are projected on the map in a way that may cause the individual PDF plots to interfere due to their size. The scale factor feature allows you to achieve the optimal visual result for presentation purposes. The scaling levels available are in the broad range of factors at 10^{-7} , 10^{-6} , 10^{-5} , 10^{-4} , 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1, 5, 10, 50, 100, 500, 1000, 10^4 , 10^5 , 10^6 , and 10^7 .
- ★ Use the “Plot external figure” button to display the plot requests that will follow in an external, independent window when activated, or to return to the in-screen display when deactivated (as first shown in paragraph III.B.1 earlier and Fig. III.B.5).
 - ⇒ When the “Plot external figure” button is activated, it enables complete control of the plot by making use of Matlab tools (e.g., axes rotation, renaming, etc.). Also, this feature allows the user to print the particular figure using the independent window menu. For more information on handling plots in separate figure windows consult Matlab Help.
 - ⇒ When plotting maps with legends, Matlab versions pre-dating the 7.1, R14, Service Pack 3 version may produce legend-related error messages in the main Matlab Command Window. These do not affect the SEKS-GUI execution, and can be ignored.

4. Use the “Fixed color scale” button if you would like to plot the eligible maps in a particular color scale, so that, e.g., maps of the same quantity can be compared at different temporal instances. The button toggles between activation and de-activation when pushed each time. Upon activating the button the first time, you need to set the lower and upper bounds of the color scale. Inspect your maps and once you make your choice regarding these bounds enter the lower bound in the box next to the “Min” tag, and the upper bound in the box next to the “Max” tag. The next map you will ask to plot will display within the color scale you just defined (Fig. III.D.5, compare with Fig. III.D.3).

- ⇒ When the button is de-activated (by pushing it again when activated) any bounds indications disappear from the boxes, and the boxes are disabled. However, if bounds have been previously defined they remain in the memory, and the last bounds to have been set will reappear in the boxes if the “Fixed color scale” button is re-activated. The bounds values can be modified as desired when the boxes are enabled (i.e., when the button is activated).
- ⇒ It may occur that the button is activated and one (or both) of the bounds boxes do not contain any values. In this case, if the user asks for a map that makes use of the fixed color scale then an error message appears indicating the issue, but the map is created and the color scale for the map is automatically set to the default. The user has the option to either de-activate the button or define properly the color scale bounds in the boxes.

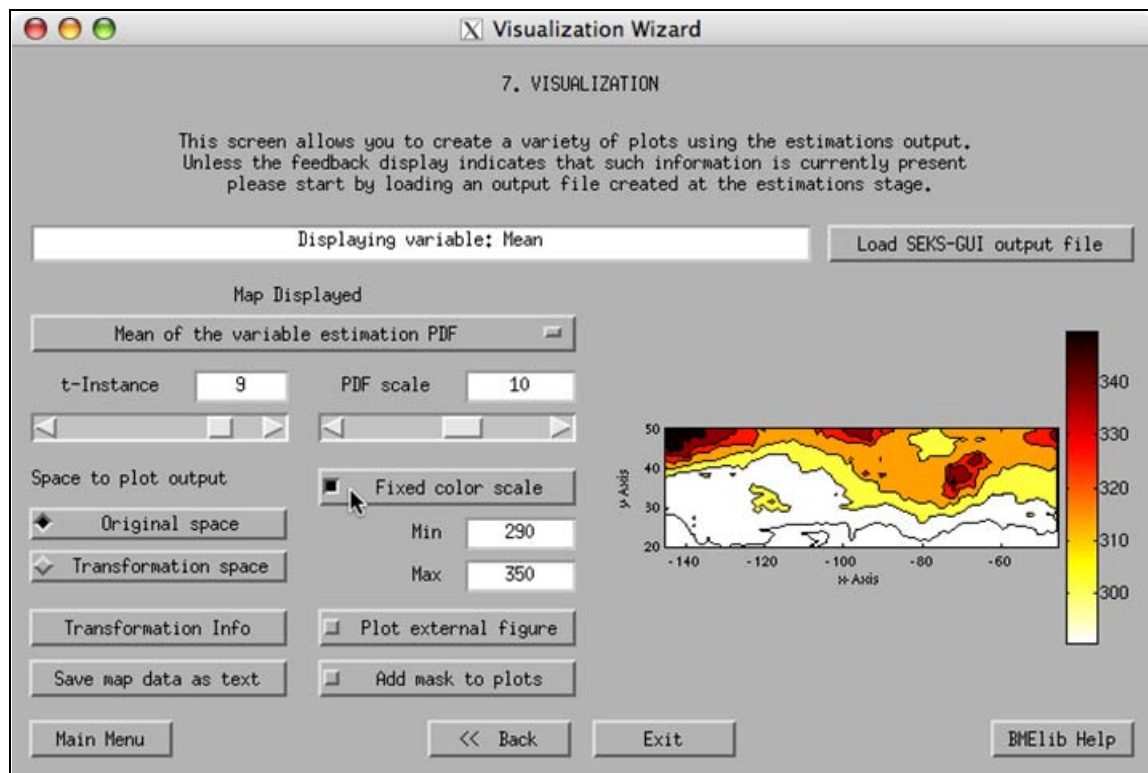


Fig. III.D.5

5. Use the “Save map data as text” button in the lower left-hand side corner to export the currently showing map data in a text file for use outside SEKS-GUI. When pushing the button, the user is prompted to specify a filename and a location for a file to save the information. There will be 4 columns in the text file with the following format:

- Column 1 will contain the X coordinates.
- Column 2 will contain the Y coordinates.
- Column 3 will contain the current temporal instance (same number in all lines), or the number “1” if this is a spatial-only case.
- Column 4 will contain the value of the selected quantity on the map at the corresponding coordinates.

The number of lines in the file is equal to the number of output nodes on the spatial estimation grid. In a spatiotemporal case, the user can export a time series of the output in text files by repeating the exporting process for the desired amount of temporal instances and assigning suitable names to each one of the text files.

- ⇒ The “Save map data as text” button is available for all types of maps in the visualization screen except for the one showing BME estimation PDFs at pre-selected output locations.
- ⇒ When Matlab cannot obtain an estimate at a node it gives out a result for that node that is called a NaN (acronym for “not a number” quantity). If NaNs are detected in the exporting process at specific nodes, SEKS-GUI replaces these NaNs with 0.
- ⇒ If for any reason the output at an estimation node is an imaginary (non-real) number and the user asks to export the map that contains it, SEKS-GUI exports only the real part of that number in the text file.

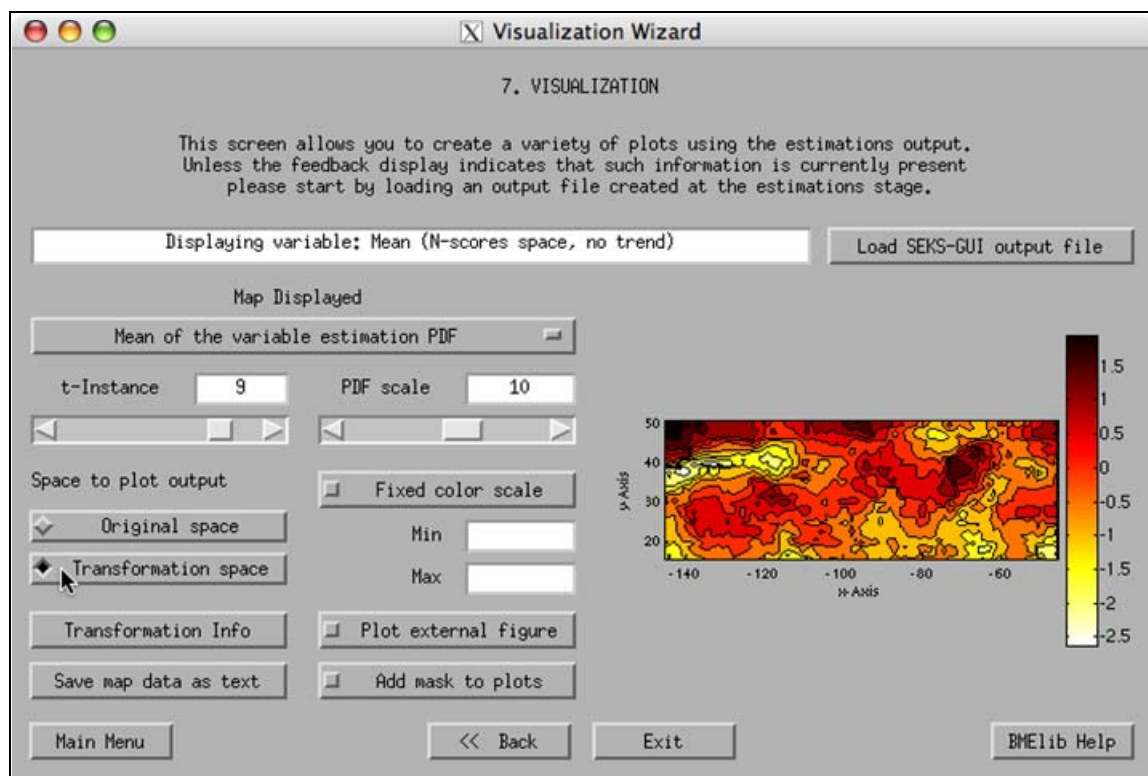


Fig. III.D.6

6. Use the 2 buttons under the tag “Space to plot BME output” to choose whether to display the eligible maps in the original-space (click on “Original space”) or the transformation-space (if a transformation was used to obtain the estimates – click on “Transformation space”). The default choice when this screen appears is display in the original-space.

- ⇒ The buttons “Original space” and “Transformation space” are mutually exclusive: Only one of these two can be chosen at a time.
- ⇒ The maps in the original-space are the back-transformed estimates with the mean trends restored at the estimation locations. The moments that are shown are based on the raw estimated PDF moments which have been themselves back-transformed and had consequently the mean trends at these locations restored. On the other side, the maps in the transformation-space feature detrended data that come directly from the raw estimated PDFs. Thus, in these maps the mean trend has not been added back, in addition to displaying the variable in the transformed-space – see Fig. III.D.6 and compare it with Fig. III.D.3.
- ⇒ The transformation buttons refer only to the BME analysis. If the present output data are derived from GBME analysis, then the screen buttons “Transformation space”, “Transformation Info”, and the PDF scale slider and edit box are automatically disabled (see Fig. III.D.7).

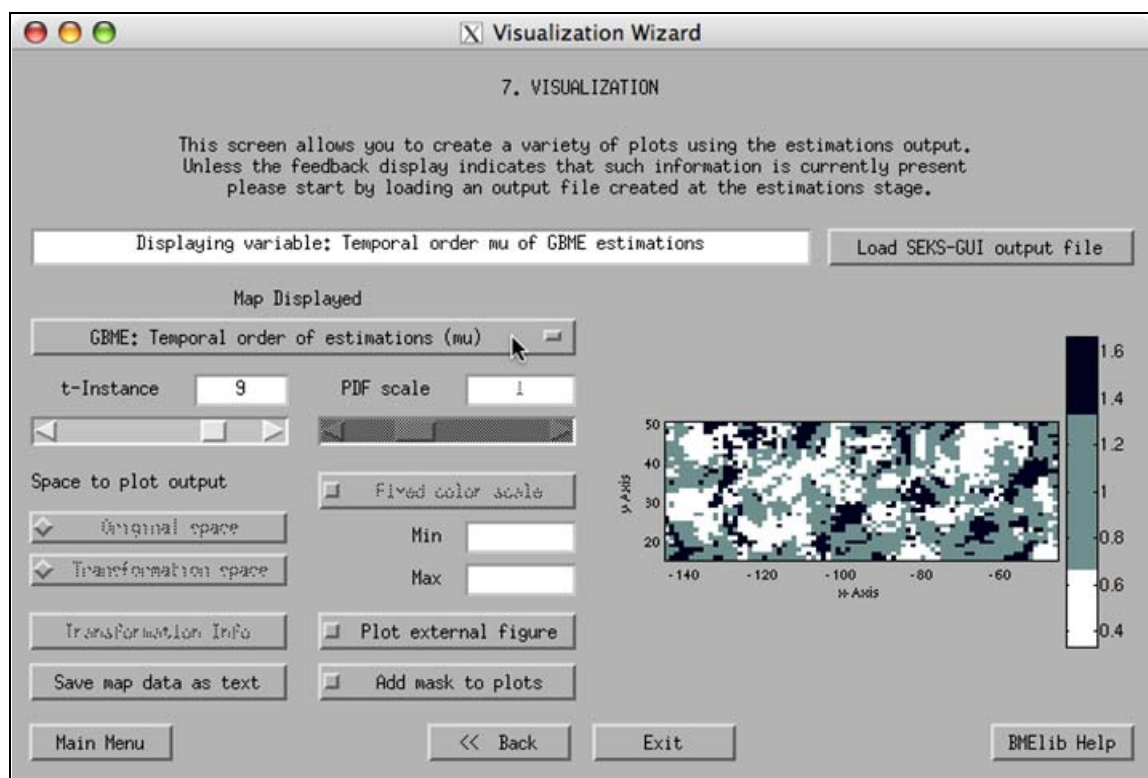


Fig. III.D.7

7. Use the button “Transformation Info” to review the transformation type and its characteristics, if any has been applied on the initial data set that was used to obtain the estimates. This button provides only information in the message box and is not used for visualization.

- ⇒ If an N-score transformation has been used, the type will show, as well as the lower and upper transformation bounds the detrended data are allowed to range in (see Fig. III.D.8, and discussion in the Introduction part A of paragraph III.B.2).
- ⇒ If a Box-Cox transformation has been used, the type will show, as well as the value of the parameter λ that was selected for the transformation (see discussion in the Introduction part A of paragraph III.B.2).
- ⇒ The transformation info button refers only to the BME analysis. If the present output data are derived from GBME analysis, then the screen buttons “Transformation space”, “Transformation Info”, and the PDF scale slider and edit box are automatically disabled (see Fig. III.D.7).

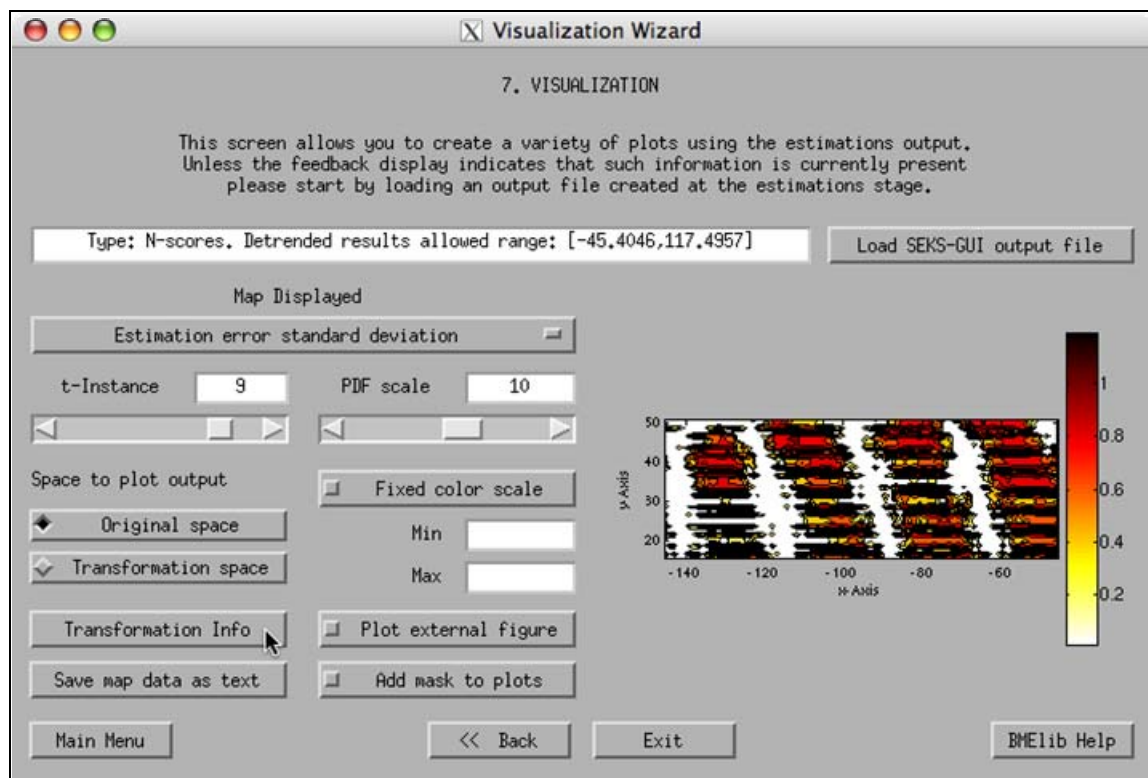


Fig. III.D.8

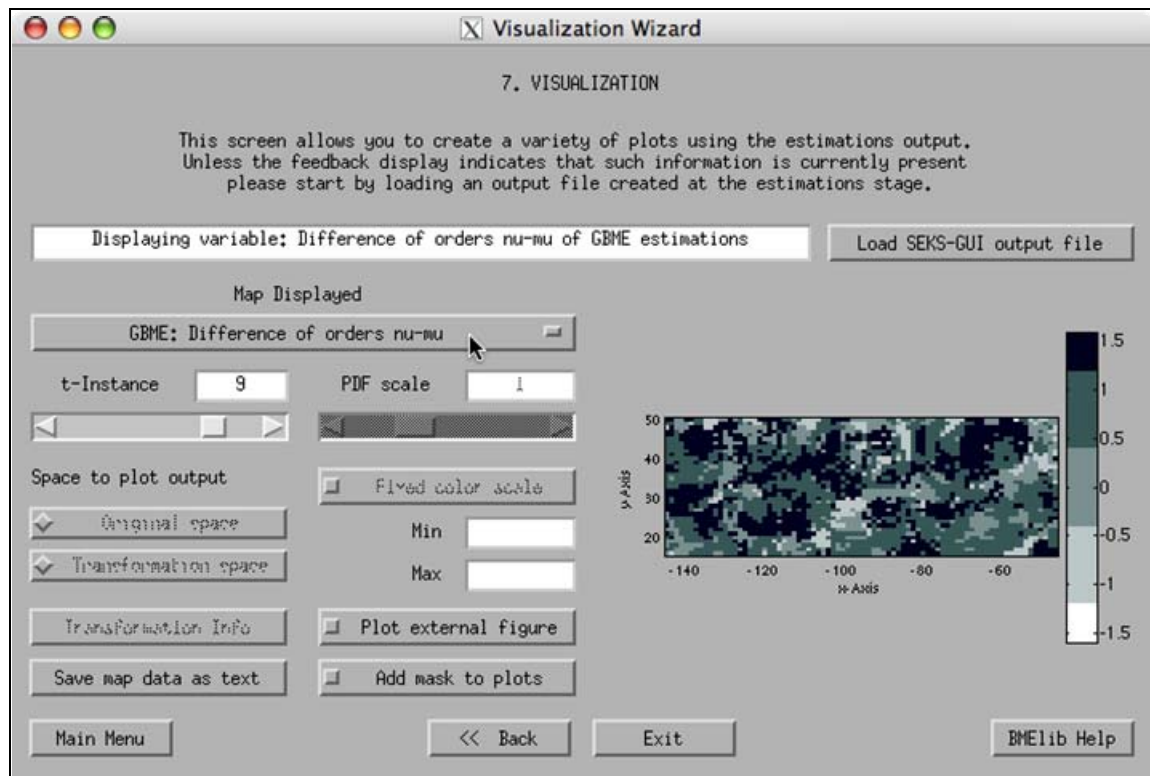


Fig. III.D.9

8. A more advanced feature for map presentations is plot masking. This is a useful feature if you would like to show the results that appear in part of your output area (e.g., by masking out the portion of a map outside a country's borders). It is more advanced because it requires that the user provides a Matlab ".m" code file with the masking information. If you can program in Matlab, you can create a map that produces the desired mask over the output grid area using suitable coordinates (see also paragraph V.3 in this manual for a basic utility provided with SEKS-GUI to assist you in the mask creation). Upon activation of the "Add mask to plots" button the user is prompted to locate such a masking code Matlab file within the user's filesystem. If you push this button accidentally you can cancel the file search. If the button is activated you can push it again to de-activate it.

9. The visualization stage concludes the SEKS-GUI functions, and at any point, you can push the "Back" button to return to the previous screen, or push the "Exit" button to exit the GUI.

⇒ In the case when a user has reached this screen after estimations, and, in addition, has loaded output from some other investigation results, the "Back" button will not send the user back to the estimation screen. The user will be rather prompted to return to the "Choose a Task" screen of the SEKS-GUI main menu.

SEKS-GUI EXAMPLES INCLUDED IN PACKAGE

You can run example test cases to familiarize with the SEKS-GUI environment. Two examples are presented in the following, namely a spatiotemporal study of Total Ozone concentrations over the United States (of which snapshots at various stages have been used as figures in this guide), and a spatial-only investigation of Arsenic concentrations in the Bangladesh groundwater.

⇒ When plotting maps with legends, Matlab may produce legend-related error messages in the main Matlab Command window. These do not affect SEKS-GUI execution, can be ignored at this stage, and will be dealt with in a coming SEKS-GUI version.

1. BME and GBME S/T study of Total Ozone concentrations over the United States

1. All the information regarding this study is located in the folder:

“guiExamples/001-TotalOzoneUSstudy-ST”

Please navigate Matlab into this folder first when requesting the input/output files in the following.

2. Start Matlab and SEKS-GUI, and when prompted for information at the appropriate screens provide the following files and input:

BME analysis

The task choice in the “Choose a Task” screen is “BME Spatiotemporal Analysis”. Highlight the task and push the “Start” button to continue.

Hard Data:

“Ozone-1-Input-HD.txt” in ASCII text format. This study is in the space/time domain.

Longitude (x-Axis, in degrees) is in data file column 1

Latitude (y-Axis, in degrees) is in data file column 2

Day (temporal reference: date in July, 1988, in z-Axis) is in data file column 3

Ozone concentrations (in ppm) are in data file column 4

Soft Data:

We use gaussian PDF soft data stored in **“Ozone-2-Input-SDgaussianPDF.txt”** in ASCII text format, and appropriately formatted with

Longitude (x-Axis, in degrees), in data file column 1

Latitude (y-Axis, in degrees), in data file column 2

Day (temporal reference: date in July, 1988, in z-Axis), in data file column 3

Mean of Ozone concentrations soft PDF (in Dobson Units), in data file column 4

Variance of Ozone concentrations soft PDF (in [Dobson Units]²), in data file column 5

Output Grid:

We use an estimation grid that is described in terms of its grid limits and node spacing stored in "**Ozone-3-Input-OutGrid.txt**". Note in the file that we ask for an output temporal span of 5 consecutive days, from July 6 to July 10, 1988.

The mean trend is stored in the Matlab file "**Ozone-4-MeanTrend.mat**". The default parameter values that appear on the detrending stage of the exploratory analysis screen have been used to obtain the trend.

The Total Ozone data in the present example have been subjected to an N-score transformation prior to proceeding to the covariance analysis.

The experimental (sample) covariance information is stored in the Matlab file "**Ozone-5-SampleCovariance-Nsc.mat**". This information has been obtained using the default parameter values that appear on the experimental covariance stage of the covariance analysis screen.

The covariance model that has been fitted comprises 2 nested model, whose characteristics are stored and can be read in the text file "**Ozone-6-CovarianceModelInfo-Nsc.txt**". Feel free to play with the values provided here and the sill/range adjustment tools of the GUI to familiarize better with the interface.

The BME estimation output is based on running the BME Mode, as well as the Confidence Interval (at 68% level) estimation types, and is stored in the Matlab file "**Ozone-7-Output-BmeModCI-Nsc.mat**". The default parameter values that appear on the BME estimation screen have been used. You can use the output file contents directly by advancing to the "Visualization Wizard" screen, or by loading the file at any time within the visualization screen to produce the Total Ozone BME study maps.

GBME analysis

The task choice in the "Choose a Task" screen is "GBME Spatiotemporal Analysis". Highlight the task and push the "Start" button to continue.

Hard Data (as in BME analysis):

"**Ozone-1-Input-HD.txt**" in ASCII text format. This study is in the space/time domain.

Longitude (x-Axis, in degrees) is in data file column 1

Latitude (y-Axis, in degrees) is in data file column 2

Day (temporal reference: date in July, 1988, in z-Axis) is in data file column 3

Ozone concentrations (in ppm) are in data file column 4

Soft Data (as in BME analysis):

We use gaussian PDF soft data stored in "**Ozone-2-Input-SDgaussianPDF.txt**" in ASCII text format, and appropriately formatted with

Longitude (x-Axis, in degrees), in data file column 1

Latitude (y-Axis, in degrees), in data file column 2

Day (temporal reference: date in July, 1988, in z-Axis), in data file column 3

Mean of Ozone concentrations soft PDF (in Dobson Units), in data file column 4

Variance of Ozone concentrations soft PDF (in [Dobson Units]²), in data file column 5

Output Data (as in BME analysis):

We use an estimation grid that is described in terms of its grid limits and node spacing stored in “**Ozone-3-Input-OutGrid.txt**”. Note in the file that we ask for an output temporal span of 5 consecutive days, from July 6 to July 10, 1988.

The GBME estimation output is based on using the default parameter values that appear as soon as the user reaches the GBME estimation screen. The output is stored in the Matlab file “**Ozone-8-Output-Gbme.mat**”. You can use the output file contents directly by advancing to the “Visualization Wizard” screen, or by loading the file at any time within the visualization screen to produce the Total Ozone GBME study maps.

Note that due to a few extreme values in the GBME estimates, you will need to set fixed bounds for the display of some maps – as explained in Section III.D.1, paragraph 4 – in order to compare the BME and GBME results.

2. BME and GBME Spatial study of Arsenic in Bangladesh drinking water

1. All the information regarding this study is located in the folder:
“**guiExamples/002-ArsenicBangladeshStudy-S**”

Please navigate Matlab into this folder first when requesting the input/output files in the following.

For the Arsenic study, masks are available for maps of the area that display the borders of Bangladesh and the surrounding countries. You can use masking information for the mean trend maps, and later for the estimation maps. This information is stored within the folder “**Arsenic-MapsMask**” in the Matlab file “**applyMask.m**”. If the path has been provided once earlier, you will not need to re-locate the necessary file. Please note that the information therein relates only to the particular case study. The script relies on knowledge of the individual coordinates of the borders – represented as a series of points – for the countries shown on the map. You can create similar scripts based on this one to create masks for your own case studies; some modest Matlab programming skills are required for this task. In addition to the files in the folder “**Arsenic-MapsMask**”, a simple start-up set of files for that purpose is provided for you in the folder “**SEKS-GUI0.6/guiLibs/Utilities/applyMaskFiles**”, and related information is provided in paragraph V.3 in this manual.

2. Start Matlab and SEKS-GUI, and when prompted for information at the appropriate screens provide the following files and input:

(continued on next page)

BME analysis

The task choice in the “Choose a Task” screen is “BME Spatiotemporal Analysis”. Highlight the task and push the “Start” button to continue.

Hard Data:

“**Arsenic-1a-Input-HD.txt**” in ASCII text format. Alternatively, you can see the structure of an Excel input file by equivalently using the Excel file “**Arsenic-1b-Input-HD.xls**”, instead. In any case, it is important to mark on the Hard Data Wizard screen the appropriate box indicating this is a spatial-only study.

Northing (x-Axis, in Km) is in data file column 1

Easting (y-Axis, in Km) is in data file column 2

Arsenic concentrations (in $\mu\text{g/L}$) are in data file column 3

Soft Data:

We use interval data stored in “**Arsenic-2-Input-SDintervals.txt**” in ASCII text format, and appropriately formatted with

Northing (x-Axis, in Km), in data file column 1

Easting (y-Axis, in Km), in data file column 2

Arsenic concentrations interval lower bounds (in $\mu\text{g/L}$), in data file column 3

Arsenic concentrations interval upper bounds (in $\mu\text{g/L}$), in data file column 4

Output Grid:

We use an estimation 2-D spatial grid that is described in terms of its grid limits and node spacing stored in “**Arsenic-3-Input-OutGrid.txt**”.

The mean trend is stored in the Matlab file “**Arsenic-4-MeanTrend.mat**”. The default parameter values that appear on the detrending stage of the exploratory analysis screen have been used to obtain the trend.

The Arsenic concentrations in the present example have been subjected to an N-score transformation prior to proceeding to the covariance analysis.

The experimental covariance information is stored in the Matlab file “**Arsenic-5-SampleCovariance-Nsc.mat**”. This information has been obtained using the default parameter values that appear on the experimental covariance stage of the covariance analysis screen.

The covariance model that has been fitted comprises 2 nested model, whose characteristics are stored and can be read in the text file “**Arsenic-6-CovarianceModelInfo-Nsc.txt**”. Feel free to play with the values provided here and the sill/range adjustment tools of the GUI to familiarize better with the interface.

The BME estimation output is based on running the BME Mode, as well as the Confidence Interval (at 68% level) estimation types, and is stored in the Matlab file “**Arsenic-7-Output-BmeModeCI-Nsc.mat**”. The default parameter values that appear on the BME estimation screen have been used. You can use the output file contents directly by advancing to the “Visualization Wizard” screen, or by loading the file at any time within the visualization screen to produce the Arsenic concentration study maps.

GBME analysis

The task choice in the “Choose a Task” screen is “GBME Spatiotemporal Analysis”. Highlight the task and push the “Start” button to continue.

Hard Data (as in BME analysis):

“**Arsenic-1a-Input-HD.txt**” in ASCII text format. Alternatively, you can see the structure of an Excel input file by equivalently using the Excel file “**Arsenic-1b-Input-HD.xls**”, instead. In any case, it is important to mark on the Hard Data Wizard screen the appropriate box indicating this is a spatial-only study.

Northing (x-Axis, in Km) is in data file column 1

Easting (y-Axis, in Km) is in data file column 2

Arsenic concentrations (in $\mu\text{g/L}$) are in data file column 3

Soft Data (as in BME analysis):

We use interval data stored in “**Arsenic-2-Input-SDintervals.txt**” in ASCII text format, and appropriately formatted with

Northing (x-Axis, in Km), in data file column 1

Easting (y-Axis, in Km), in data file column 2

Arsenic concentrations interval lower bounds (in $\mu\text{g/L}$), in data file column 3

Arsenic concentrations interval upper bounds (in $\mu\text{g/L}$), in data file column 4

Output Grid (as in BME analysis):

We use an estimation 2-D spatial grid that is described in terms of its grid limits and node spacing stored in “**Arsenic-3-Input-OutGrid.txt**”.

The GBME estimation output is based on using the default parameter values that appear as soon as the user reaches the GBME estimation screen. The output is stored in the Matlab file “**Arsenic-8-Output-Gbme.mat**”. You can use the output file contents directly by advancing to the “Visualization Wizard” screen, or by loading the file at any time within the visualization screen to produce the Total Ozone GBME study maps.

Note that due to a few extreme values in the GBME estimates, you will need to set fixed bounds for the display of some maps – as explained in Section III.D.1, paragraph 4 – in order to compare the BME and GBME results.

SEKS-GUI UTILITIES INCLUDED IN PACKAGE

1. Data exchange with shapefiles: Converting shapefiles into text

The shapefile format is the standard way to manipulate maps in the proprietary ArcGIS software by ESRI. The SEKS-GUI software requires that such information be in text format. Below are directions for converting polygon shapefiles and related data for a single attribute into a text file that SEKS-GUI can work with using a Python script called 'shape2text.py'.

Requirements:

1. A binary ESRI/ArcGIS shapefile. The points in the shapefile are converted by the script into entries (points with spatial/spatiotemporal coordinates) in a text file.
2. An attribute text file must be space delimited and must have rows referring to cross-sectional units and columns to time periods (if there are multiple time periods). The attribute data can either be cross-sectional or panel data (time-series). The coordinates in the shapefile are consolidated with the attribute data into a single file that SEKS can use as input.
3. To run the script you must have Python (www.python.com) installed and within your system path.

Instructions:

The following instructions are given in the form of an example where a shapefile and an attribute text file are selected using the script.

1. Place all key files in a single folder and navigate your way to it (in this example: “**\$local**”).
2. Run the python script provided in the SEKS-GUI package. The script is the file “**shape2text.py**” and is located in the folder “**SEKS-GUI/guiLibs/Utilities**”. You can run the script after you copy the script file to the folder with the other key files; then type the executable for Python in your system and use the script file as an argument:

```
$local\python shape2text.py
```

3. Choose the shapefile to be converted from the list of shapefiles (with a postfix: *.shp) in current directory:

```
kansas.shp  
newyork.shp  
california.shp
```

Enter the shape file name to be converted to X, Y (but **do not include** .shp):

```
$california
```

4. Choose the attribute or data file to be merged with shapefile information from a list:

```
2          bmepy.py          california.xyz.xyz  shape2text.py
Gis.py     california.csts   fileIOprac.py     shape2text2.py
Gis.pyc    california.shp   junk.txt          shapereader.py
addition   california.txt    saybye.py
arcv2stars.py  california.xyz  sayhi.py
```

```
Enter file holding attribute (Z) values: california.csts
```

5. Enter 'CS' if the data is just for one year or 'CSTS' for multiple years:

```
'CS'      = cross-sectional data
'CSTS'    = cross-sectional, time-series data
```

```
Enter 'CS' or 'CSTS': CSTS
```

The output from the above procedure is shown below:

Sample data from final file:

```
-116.0556175 33.752874 4 33.752874
-119.7219885 34.01836 4 34.01836
-120.3837255 34.045747 4 34.045747
-120.109372 33.9667205 4 33.9667205
-119.399344 34.009645 4 34.009645
-117.764832756 33.6671330053 4 33.6671330053
-116.838523677 33.0195905 4 33.0195905
-118.453602 33.3885865 4 33.3885865
-115.28455427 33.0261927617 4 33.0261927617
-119.5042505 33.2501805 4 33.2501805
-118.4800315 32.918273 4 32.918273
```

Name of output file ends in '.xyz' :

```
output x y points to:
california.xyz
```

Miscellaneous Summary Information on Shapefile:

```
=====
Shape File Name: california.shp
Type: Polygon
Number of records: 68
Polygons of multiple parts:[]
Bounding box:
  Xmin, Ymin:  -124.40959100,32.53415600
  Xmax, Ymax:  -114.13442654,42.00951800
=====
Number of Attributes or Z values:
1
```

The file structure of the 'output.xyz' file, based on time-series data, is simply as follows:

```
The first column = x coordinate
The second column = y coordinate
The third column = time period
The fourth column = attribute value
```


2. Data exchange with shapefiles: Converting text into shapefiles

Using this tool enables you to convert an ASCII textfile containing point data (such as the input or output of SEKS-GUI) into ESRI ArcMap's shapefile format. Essentially, this tool first creates a point layer based on X and Y coordinates, and then transforms this layer to a shapefile.

Requirements:

For the conversion of a text file into a shapefile you will need the following:

1. A text file that contains at least X and Y spatial coordinates of data that you want to put in the shapefile. Each line must contain an entry for 1 datum, and the coordinates must be consistent in all entries (e.g., longitudes in 1st column, latitudes in 2nd column, values in 3rd column). Spatiotemporal data may contain also temporal coordinates; the users need to know where this information is located in the text file, and whether they want any such additional information to be included in the shapefile.
2. The file named **text2shape.mxd** packaged with the SEKS-GUI in the folder "**guiLibs/Utilities**".
3. The ESRI ArcGIS software or any other software that can perform this task. The following example describes the process using ArcMap within ArcGIS.

Instructions:

1. Create a folder on your computer where you wish to place the shapefile and associated files.
2. To set things up, open ArcMap, select "File→Open" from the menu bar, and then locate and open the "**text2shape.mxd**" file packaged with the SEKS-GUI in the folder: "**SEKS-GUI/guiLibs/Utilities**"
"text2shape.mxd" is an ArcMap project file that is empty except for the addition of the SEKS-GUI Conversion Tool "Text to shapefile."

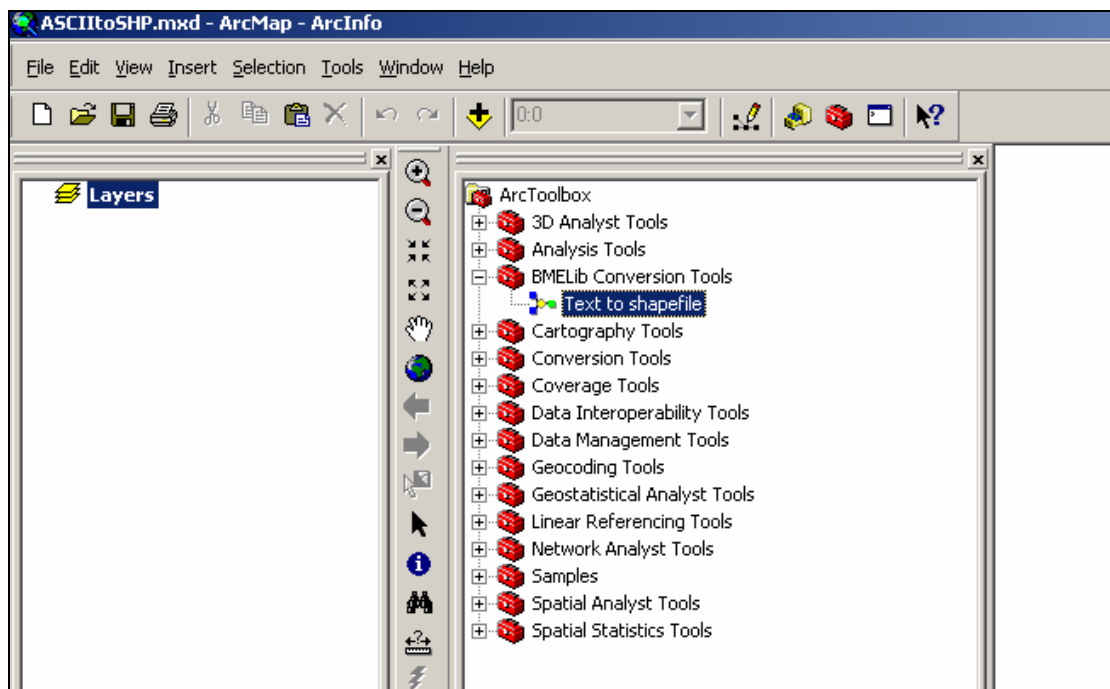


Fig. V.1

3. If the ArcToolbox is not open, select “Window→ArcToolbox” from the menu bar. Underneath “BMELib Conversion Tools” you should see “Text to shapefile” (Fig. V.1). Double-click this tool to open up the options window.
4. A new window opens that provides a brief description of the tool and asks the user to enter particulars for the conversion requested (Fig. V.2).
 - a. Enter the directory and name of your textfile, or click on the folder icon to browse for your file within your directory structure.
 - b. Select the column upon which your X-coordinate resides. If your textfile does not contain variable names in its first row, you will need to know beforehand where this column is located. If your textfile contains variable names in its first row, you may select the appropriate variable name from the drop-down menu.
 - c. Repeat Step (b) above for the column upon which your Y-coordinate resides.
 - d. Because shapefiles consist of both the shapefiles themselves and associated files, you need to specify a folder where all these files will be placed. Click on the folder icon to browse for your folder of choice. This folder must have been created prior to this step.
 - e. Choose a name for your shapefile and its associated files. The created files will all have the same names with different extensions.
 - f. Click OK.

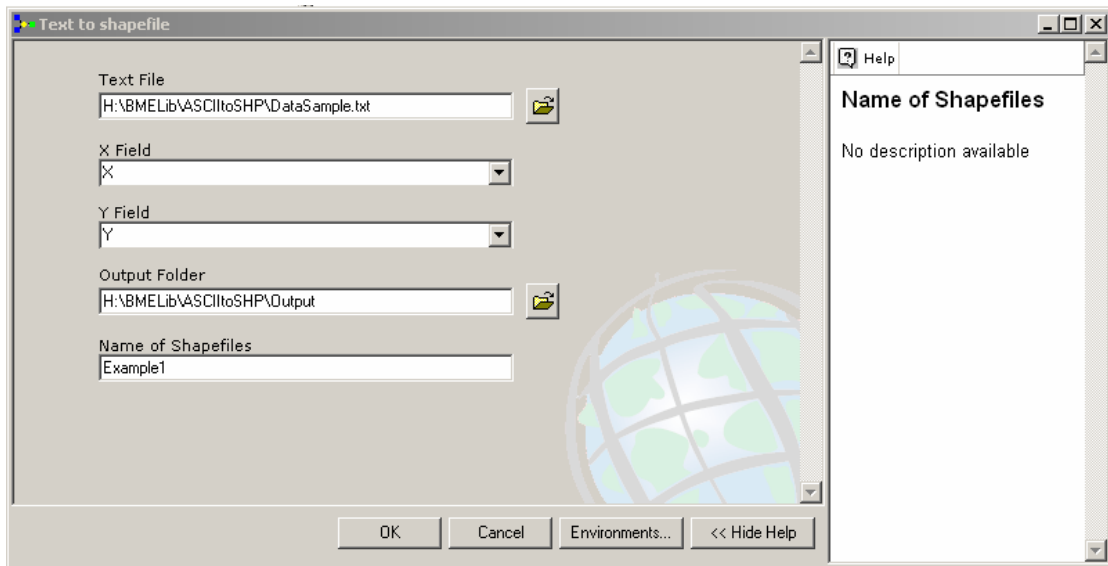


Fig. V.2

5. The process dialog will appear and inform you of the progress. Close this process dialog after ensuring that the process was successful.
6. Your data points should automatically appear as a layer in ArcMap (If not, you can open the shapefile by selecting File→Add data).
 - ⇒ If you check the contents within your output folder, you will see six newly created files with extensions **.dbf**, **.sbn**, **.sbx**, **.shp**, **.shp.xml**, and **.shx**.
7. Optionally, you may want to ensure that all your data was transferred to the shapefile. Check for x, y and attribute data (plus temporal coordinates when applicable) by:
 - a. Right clicking on the newly created layer name located on the left side of your ArcMap screen.
 - b. Selecting “Open Attribute Table”.
 - ⇒ A table containing all your data (and newly created ID values) will open.

3. A start-up set of files for the creation of masking files

All the information regarding this section is located in the folder:

“SEKS-GUI/guiLibs/Utilities/applyMaskFiles”

Please navigate Matlab into this folder first when requesting the input/output files in the following.

⇒ Some basic Matlab programming experience is required for the tasks in this section.

It was shown in the SEKS-GUI screens instructions that if desired, a mask can be applied on top of the GUI-generated maps. The above folder contains a simple example of such a mask that can be relatively easily modified to fit your needs.

1. First, you will need a text file that contains the masking element. For example, the masking folder contains a file of the borders of the state of California. The file is called **“californiaBorders.txt”** and is a collection of points, each of them fully described in a separate line using its longitude and latitude coordinates in the 1st and 2nd columns, respectively.

⇒ You must provide the points in a sequence, as Matlab will plot them by joining the coordinates in any two consecutive entries with a line.

⇒ The last entry in the file must be the same as the first one, so that there is a closed polygon to plot.

★ More than 1 polygons may be defined in separate files and can be all plotted within the Matlab **“applyMask.m”** file.

⇒ The file must also contain some reference to the output grid corner coordinates. In case you wish to mask out the surroundings of an area, as is the case with the California borders, the outer grid coordinates will be used alongside with the masking element coordinates to form a closed area that surrounds the actual part of the map you want to show. This closed area can be filled with some color (e.g., white) in the **“applyMask.m”** file in order to allow only the desired area to show in your final maps.

In the **“californiaBorders.txt”** example file, we assume that the output grid ranges within [-125,-114] longitude and [32,42] latitude. Notice how the coordinates of the 4 grid corners have been incorporated in the state borders information in the file lines 53-57.

Similar approaches can be used in a variety of cases.

2. The second step is to properly modify the Matlab M-file **“applyMask.m”** that you will indicate from within the SEKS-GUI as the masking file to use in your application. The following is a brief tour of the file structure:

- Line 11 is used to load the masking element text file.
- Line 12 defines the output grid map corners, as discussed in the previous paragraph.
- Line 16 explicitly asks to take actions on an existing map (in our case, the ones already created by SEKS-GUI). If this line is not used prior to any of the other following plotting commands, the actions that follow will overwrite the pre-existing map.
- Line 17 plots the contents of the masking element text file as a solid line.
- Line 18 fills the polygon defined by the above solid line with white color.

The remaining lines add labels and define the plot limits based on the corner coordinates provided earlier.

- ★ More than 1 polygons can be added for plotting in the “applyMask.m” file. If the additional files reside in the same folder as the “californiaBorders.txt” file, their content must be loaded in a similar manner as shown in the example line 11 of “applyMask.m” (or their location must be appropriately defined first, prior to loading their content). Any additional content can be plotted **on top** of the existing map by repeating the sequence currently shown in lines 16-18 in “applyMask.m”. Just copy and paste these lines and adjust them suitably according to the additional content.

PART VI

BIBLIOGRAPHY

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C.: *Time Series Analysis, Forecasting and Control*, 3rd ed. Prentice Hall, Englewood Cliffs, NJ, 1994.

Christakos, G.: *Random Field Models in Earth Sciences*. Academic Press, San Diego, CA, 474 p., 1992. (Out of Print.) New edition (2005), Dover Publ. Inc., Mineola, NY.

Christakos, G. and D.T. Hristopulos: *Spatiotemporal Environmental Health Modelling*. Kluwer Academic Publ., Boston, Mass., 423 p., 1998.

Christakos, G.: *Modern Spatiotemporal Geostatistics*. Oxford Univ. Press, New York, N.Y., 304 p., 2000.

Christakos, G., P. Bogaert, and M.L. Serre: *Temporal GIS*. Springer-Verlag, New York, N.Y., 220 p., With CD-ROM, 2002.

Deutsch, C. V., and Journel, A. G.: *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press, New York, 369 p. and 1 compact disk, 1998.

ESRI, link on shapefiles (current as of May 2006): "<http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>".

Olea, R.A.: *Geostatistics for Engineers and Earth Scientists*. Kluwer Acad. Publ., Boston, MA, 303 p., 1999.

Olea, R.A.: *A Six-Step Practical Approach to Semivariogram Modeling*. SERRA, In Press, 2006.