Software Requirements Specification for EEGSourceLocalization software

Leila Mousapour

October 19, 2020

Contents

Revision History

Date	Version	Notes
October 2 2020	1.0	The first version

1 Reference Material

This section records information for easy reference.

1.1 Table of Units

Throughout this document SI (Système International d'Unités) is employed as the unit system. In addition to the basic units, several derived units are used as described below. For each unit, the symbol is given followed by a description of the unit and the SI name.

symbol	unit	SI
m	length	metre
hz	Frequency	herts
s	time	second
V	voltage	volt
W	power	watt ($W = J\,s^{-1}$)

1.2 Abbreviations and Acronyms

symbol	description
A	Assumption
DD	Data Definition
GD	General Definition
GS	Goal Statement
IM	Instance Model
LC	Likely Change
PS	Physical System Description
R	Requirement
SRS	Software Requirements Specification
EEGSourceLocalizer	Electroencephalogram Source Localization Software
T	Theoretical Model
EEG	Electroencephalogram
BCI	Brain-Computer Interface
SL	Source Localization
ML	Machine Learning
LCMV	Linearly Constraint Minimum Variance
DICS	Dynamic Imaging of Coherent Sources

2　Introduction

Electroencephalography (EEG), which is a method to record electrical activity of the brain, has a plethora of applications such as decoding mental imageries used in brain-computer interfaces (BCI). One of the big open problems in EEG signal processing is finding a good feature space after which we can apply machine learning and classification methods to the data. The standard in the field is to start with the electrode space and then extract features such as amplitude or latencies (time domain features), frequency power spectra (frequency features), common spatial patterns (spatial features) etc. A novel approach that we would like to investigate is to first map EEG signals from electrode space into spatial coordinates of the brain to achieve more useful features.

These techniques are known as "source localization " algorithms which can be applied to the signals recorded from the scalp and locate the underlying active sources generating the activity sensed on the electrodes expected to increase the signal to noise ratio. Additionally, mapping the activity from an n-channel space to fewer sources reduces data dimensionality immensely, helps avoid overfitting and redundancy, leads to better human interpretations and less computational cost with the simplification of models. Therefore, this scientific software will implement several techniques to map EEG signals from electrode space to source space.

In this introduction, first the purpose of the SRS document is discussed. Afterwards, the scope of requirements for the software are stated from the big picture view. Additionally, the skills and knowledge of the readers of this document are clarified. This section will end by illustrating a road map of the whole SRS document.

2.1　Purpose of Document

This document describes the features and behaviour of the EEGSourceLocalizer software application. It includes a variety of elements that attempts to define the intended functionality required by the customer. The purpose of the document is to provide detailed requirements of the software and to provide planning for the design stage.

Goals, assumptions, theoretical models, definitions, and other model derivation information are explained, allowing the reader to fully understand and verify the purpose and scientific basis of EEGSourceLocalizer. This SRS will remain abstract, describing what problem is being solved, but not how to solve it. This document will be used as a starting point for subsequent development phases, including writing the design specification and the software verification and validation plan.

2.2 Scope of Requirements

Source localization is a very general problem which can be in raised in various contexts such as telecommunication signals, brain signals etc. Also, there are numerous algorithms with different set of assumptions to solve this problem. However, the scope of this project is limited to source localization of EEG signals and it will not be applicable to other neuro imaging modalities such as magnetoencephalgggram (MEG), fNIRS etc. Moreover, this software will only use 2 different algorithms of source localization: Beamforming and sLORETA and will not cover other approaches.

2.3 Characteristics of Intended Reader

The intended readers are the people who will read, review and maintain the SRS. They are the people that will conceivably design the software that is intended to meet the requirements. Therefore, they are expected to have a general understanding of EEG neuroimaging technique and EEG signal processing. Also, a very good knowledge of linear algebra and matrix manipulation is required.

2.4 Organization of Document

The document follows the organizational scheme laid out by Smith *et al.* **??**, starting with general reference material for the reader. The introduction of the document details the purpose of the document and intended reader, and is followed by the General System Description. The Specific System Description reduces the problem under consideration to a set of clearly defined definitions and models. The Requirements section of the document clearly outlines both the Functional and Non-functional Requirements that the software must comply with. Finally, the document contains a list of Likely Changes to the software and Traceability matrices and graphs.

3 General System Description

This section provides general information about the system. It identifies the interfaces between the system and its environment, describes the user characteristics and lists the system constraints.

3.1 System Context

Figure 1 shows the system context which is an abstract view of the software. A circle represents an external entity outside the software, the user in this case. A rectangle represents the software system itself (EEGSourceLocalizer). Arrows are used to show the data flow between the system and its environment including the main inputs and outputs. The responsibilities of the user and the system are as follows:
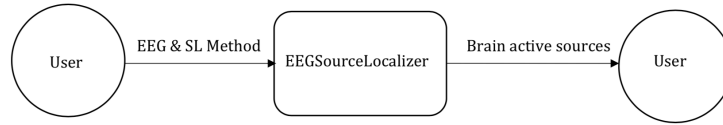
Figure 1: System Context

- User Responsibilities:

  - Clean and preprocess the EEG data
  - Provide all the input data to the system (including EEG recording)
  - Given two or more options by the system, decide to use which SL method

- EEGSourceLocalization Responsibilities:

  - Detect data type mismatch or missing inputs
  - Calculate and plot the required outputs

3.2 User Characteristics

The user is expected to be familiar with EEG cleaning and preprocessing to provide the software with cleaned data such as artifact removal as SL techniques are very sensitive to the data quality. Also, the user should have the knowledge about the different techniques that are provided as SL options in this software to be able to set the configurations of the methods. For this purpose, they should have an understanding of signal processing and linear algebra. Additionally, the end user is preferred to have a basic cognitive neuroscience knowledge so that they have a rough idea where the active sources during the recorded task should be expected.

3.3 System Constraints

There is no system constraints for this project.

4 Specific System Description

This section first presents the problem description, which gives a high-level view of the problem to be solved. This is followed by the solution characteristics specification, which presents the assumptions, theories, definitions and finally the instance models.

4.1 Problem Description

During rest state or while performing any physical or mental task, there are many activities occurring inside the brain which we can record a superficial effect of them sensed on the EEG electrodes. However, different parts of the brain have different functionalities and are more active while performing specific tasks. EEGSourceLocalization is intended to find the location and activity pattern of the underlying sources in the brain which are responsible for the EEG signal that we record from the scalp.

4.1.1 Terminology and Definitions

This subsection provides a list of terms that are used in the subsequent sections and their meaning, with the purpose of reducing ambiguity and making it easier to correctly understand the requirements:

- EEG : The electroencephalogram (EEG) is a recording of the electrical activity of the brain from the scalp. The recorded waveforms reflect the cortical electrical activity. Signal intensity: EEG activity is quite small, measured in microvolts.

- MRI: Magnetic resonance imaging (MRI) is a medical imaging technique that uses a magnetic field and computer-generated radio waves to create detailed images of the organs and tissues in your body.

- Coordinate system: in geometry, a coordinate system is a system that uses one or more numbers, or coordinates, to uniquely determine the position of the points or other geometric elements on a manifold such as Euclidean space.

- Covariance: In probability theory and statistics, a covariance matrix is a square matrix giving the covariance between each pair of elements of a given random vector.

- Rank: The maximum number of linearly independent vectors in a matrix is equal to the number of non-zero rows in its row echelon matrix. Therefore, to find the rank of a matrix, we simply transform the matrix to its row echelon form and count the number of non-zero rows.

- Rank-deficient: A matrix is said to be rank-deficient if it does not have full rank. The rank deficiency of a matrix is the difference between the lesser between the number of rows and columns, and the rank.

- Beamforming: Beamforming or spatial filtering is a signal processing technique used in sensor arrays for directional signal transmission or reception.

4.1.2 Physical System Description

The purpose of this section is to clearly and unambiguously state the physical system that is to be modelled. The complete physical system consist of two main components: the head and the EEG electrodes set up on the the scalp. Each of these 2 have several elements.

EEG equipment consist of a number of electrodes (could be anything between 6 to 256 electrodes) placed on predefined locations. These locations could be specified by a cap or by sectioning the head. Also, head can be physically modeled with 3 tissues with different conductivity: brain, skull and scalp.

The physical system of EEGSourceLocalization, as shown in Figure ?, includes the following elements:

PS1: EEG recording electrodes

PS2: Electrical activity sensed by the electrodes (EEG signal)

PS3: Exact location of the electrodes (coordination)

PS4: Head (Scull, scalp and brain)

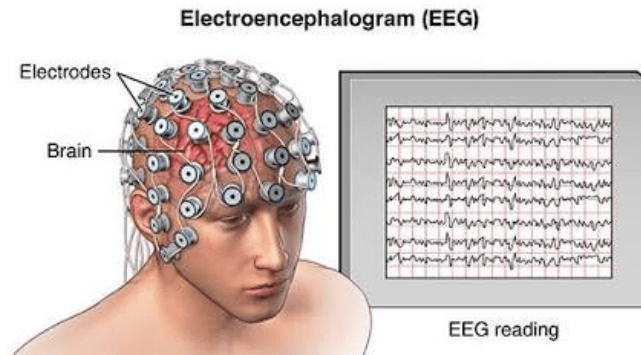PS5: Head MRI

PS6: Brain electrical activity

Figure 2: System Context

4.1.3 Goal Statements
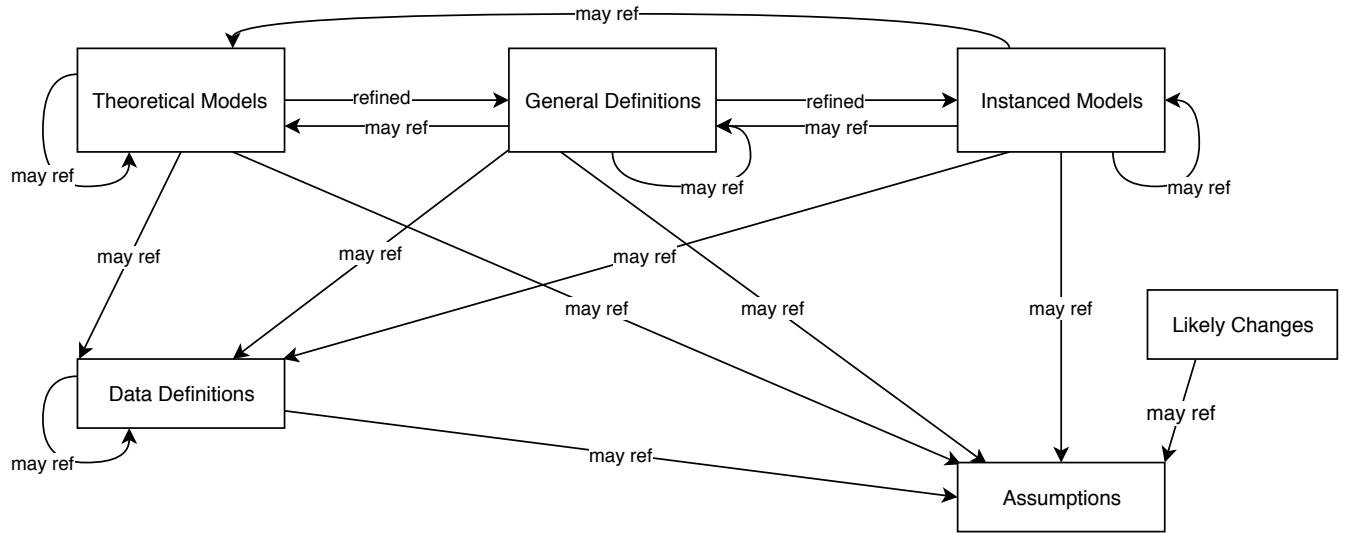
Given the inputs the goal statements are:

GS1: Locate the active sources and their orientation

GS2: Estimate the sources activity time course (time series)

GS3: Estimate the average source power (over time)

GS4: Plot the sources on a sliced MRI and 3D brain model

4.2 Solution Characteristics Specification

The instance models that govern EEGSourceLocalization are presented in Subsection **??**. The information to understand the meaning of the instance models and their derivation is also presented, so that the instance models can be verified.

4.2.1 Assumptions

This section simplifies the original problem and helps in developing the theoretical model by filling in the missing information for the physical system. The numbers given in the square brackets refer to the theoretical model [T], general definition [GD], data definition [DD], instance model [IM], or likely change [LC], in which the respective assumption is used.

A1: Each source contributes to all the channels

A2: Each source have different visibility on each channel (defines the contribution of each source at each channel)

A3: The activity on each channel is the linear combination of all source activities (we base our source modeling on this)

A4: Sufficient length of stationary data (only related to the function of interest) to get a good estimate of covariance matrix

4.2.2 Theoretical Models

This section focuses on the general equations and laws that EEGSourceLocalization is based on. Theoretical models are sets of abstract mathematical equations or axioms for solving the problem described in Section "Physical System Description". The fundamental idea of source localization is divided into two main problems: one is termed a forward problem and the other is an inverse problem. The forward problem is the prediction of potential differences from the current sources inside the brain, and the estimation of the location of the sources from the measured data is termed as inverse problem.

Number	T1
Label	**Current Dipole**
Equation	$d = \mathbf{I}pn'_d$
Description	The electrical activity inside the brain can be modeled as a current dipole. This dipole provides good approximation for a small source viewed from a relatively large distance, as the far field of a realistic source is mainly dipolar.
	The current dipole consists of the current source and the sink separated by distance p with an equal amount of current to be injected or removed. The magnitude and orientation of dipole are characterized by the dipole moment d vector, which is pointing from the source to the sink in the figure. Hence, the magnitude of the dipole is $|d| = \mathbf{I}p$, where p is the distance between the positive and negative poles, respectively. n' is the unit vector defining the direction from the source to the d sink. However, in the Cartesian coordinate system, the dipole can be represented by x, y, z unit vectors.
Source	Jatoi, M. A. (2019). Brain Source Localization Using EEG Signal Analysis. CRC Press.
Ref. By	GD**??**

Number	T2
Label	**Forward Problem/Model**
Equation	$x = \sum_{i=1}^{L} \mathbf{H}(\mathbf{q_i})\mathbf{m}(\mathbf{q_i}) + n$
Description	The forward problem is concerned with the computation of scalp potentials for a specific set of neural current sources. Meaning, if we know the active sources in the brain, what would be the EEG we record on the scalp. Hence, it needs to model the head that is used for calculation of scalp potentials.
	Let x be an $N \times 1$ vector composed of the potentials measured at the electrode sites at a given instant in time associated with a single dipole source. If this source has location represented by the 3×1 vector $\mathbf{q}$, then $x = \mathbf{H}(\mathbf{q})\mathbf{m}(\mathbf{q})$ where the elements of vector $\mathbf{m}(\mathbf{q})$ are x, y, z components of the dipole moment at an instant in time and the column of $N \times 3$ transfer matrix $\mathbf{H}(q)$ represent solution to the forward problem. The medium is linear so the potential at the scalp is the superposition of the potentials from many active neurons. Suppose x is composed of the potentials due to L active dipole sources at locations $\mathbf{q}_i$ and noise.
Source	Van Veen BD, van Drongelen W, Yuchtman M, Suzuki A. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. IEEE Trans Biomed Eng. 1997
Ref. By	GD??

Number	T3
Label	**Inverse Problem/Model**
Equation	$Var(\mathbf{q}_i) = tr[\mathbf{H}^T(q_i)C^{-1}(x)H(q_i)]^{-1}$
Description	Inverse problem intends to find the source q_i given the activity x on the scalp and is the problem of interest in this project. This problem is known as source localization or EEG inverse problem as the data (potentials) are given and one has to design the model from the available data. Mathematically, the EEG inverse problem is typically an optimization problem. The EEG inverse problem is an ill-posed optimization problem as unknown (sources) outnumbers the known (sensors). LCMV algorithm, which is one of the methods this software implements, gives an estimated variance or strength of the activity of the moment at location q_i which is represented in the equation above.
Source	Van Veen BD, van Drongelen W, Yuchtman M, Suzuki A. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. IEEE Trans Biomed Eng. 1997
Ref. By	GD**??**

4.2.3   Data Definitions

This section collects and defines all the data (given information) needed to build the instance models. The dimension of each quantity is also given.

Number	DD1
Label	**Electroencephalogram (EEG)**
Symbol	x
SI Units	μV
Dimmension	$nChannel \times \frac{t}{f_s}$
Description	EEG is the voltage recorded on the electrodes. Depending on the experiment and the system we use, the number of channels will be different (ranging from 5 to 256 (HDEEG)). Each electrode is recording a sampled signal with sampling frequency of f_s during time t.
Sources	Niedermeyer E.; da Silva F.L. (2004). Electroencephalography: Basic Principles, Clinical Applications, and Related Fields. Lippincott Williams & Wilkins.
Ref. By	IM**??**

Number	DD2
Label	**Magnetic Resonance Imaging (MRI)**
Symbol	MRI
SI Units	cm
Dimension	$Nslices \times res \times res$
Description	Magnetic resonance imaging (MRI) is a medical imaging technique used in radiology to form pictures of the anatomy and the physiological processes of the body. MRI uses a strong magnetic field and radio waves to create detailed images of the organs and tissues within the body. We need these images to extract the head shape of the subject in order to model their head as precise as possible. An MRI file contains several 2D images of the head, assigning a shade of grey (0-255) depending on the tissue. Also the resolution (res) of this imaging depends on the device.
Sources	Todd A. Gould, RT-(R)(MR)(ARRT) and Molly Edmonds. "How MRI Works." HowStuffWorks.com. 25 October 2010. Web. 20 December 2018..
Ref. By	IM**??**

Number	DD3
Label	**Electrode location**
Symbol	$loc_i(x, y, z)$
SI Units	cm
Dimension	$nChannel \times 3$
Description	In order to create a head model, we need to know the exact location of electrodes which are placed on the scalp. These locations are usually standardized in different systems of EEG recording and need to be aligned with the MRI of the subject based on biological landmarks (3 points: nasion, inion and asterion)
Sources	Citation here
Ref. By	IM??

4.2.4 Instance Models

This section transforms the problem defined in Section **??** into one which is expressed in mathematical terms. It uses concrete symbols defined in Section **??** to replace the abstract symbols in the models identified in Sections **??** and **??**.

Number	IM1
Label	**Head model (Volume conduction model)**
Input	MRI and the location of landmarks (Fiducial points)
	The input is constrained so that $MRI_{pixel} \leq 255$
	Also the MRI should be in the same coordinate system as the electrodes.
Output	The description of the head geometry (3 3D segmented surfaces (into triangles) on the borders of scalp, skull and brain.)
	And tissue connectivity
Description	A volume conduction model describes how currents flow through the tissue. For creating a head model, we first need to have an MRI image of the person's head which have the head's anatomy and the coordinate system the anatomy is represented in. First, we align the MRI to the head coordinate system by a transfer function mapping the fiducial points of the MRI into the electrodes'. Finally, the voxels of the anatomical MRI are segmented/separated into 3 different tissue types: scalp, skull, brain using Boundary Element Method (BEM). Each of these tissues have a conductivity constant. (A**??**, A**??**).
Sources	Jatoi, M. A. (2019). Brain Source Localization Using EEG Signal Analysis. CRC Press
Ref. By	IM**??**

Number	IM2
Label	**Lead filed (forward model)**
Input	Electrode locations and head model from IM**??**
	The input is constrained so that the electrodes should be aligned on the head model (IM**??**)
	Also, only the main electrodes should be feed as input here (remove all the externals and fiducial points)
Output	source model (the 3D brain grid - each voxel is a dipole/source point)
	lead field from IM**??**
Description	The lead field consists of the forward model for all the dipole locations on a 3D source model. All the channels should be included in the forward model calculation. The forward solution is expressed as the lead field matrix ($nChannel \times sources$) where each column corresponds with the potential of field distribution on all sensors for each orientation of the dipole. For this purpose, we first have to discretize the brain volume into a regular grid of dipoles (voxels). Then, for each grid point, the lead field matrix is calculated. (A**??**, A**??**).
Sources	Jatoi, M. A. (2019). Brain Source Localization Using EEG Signal Analysis. CRC Press
Ref. By	IM**??**

Number	IM3
Label	**Beamformer (inverse solution)**
Input	EEG data, lead filed, head model, MRI and the method
	The input is constrained so that $F(EEG) \leq 100$ (data should be low-pass filtered)
	Also, the covariance of the EEG data should not be rank deficient : $rank(Cov_{EEG}) \geq Maxlength(Cov_{EEG})$
Output	source activity in time for every brain voxel as a signal (matrix of the dimension $nSources \times \frac{t}{f_s}$
	Source activity map on an slices MRI based on source power
Description	The activity of the sources are calculated in this step based on the decided method (LCMV or sLORETA). (A**??**, A**??**).
Sources	Jatoi, M. A. (2019). Brain Source Localization Using EEG Signal Analysis. CRC Press
Ref. By	IM**??**

4.2.5 Input Data Constraints

Table **??** shows the data constraints on the input output variables. The column for physical constraints gives the physical limitations on the range of values that can be taken by the variable. The column for software constraints restricts the range of inputs to reasonable values. The software constraints will be helpful in the design stage for picking suitable algorithms. The constraints are conservative, to give the user of the model the flexibility to experiment with unusual situations. The column of typical values is intended to provide a feel for a common scenario. The uncertainty column provides an estimate of the confidence with which the physical quantities can be measured. This information would be part of the input if one were performing an uncertainty quantification exercise.

The specification parameters in Table **??** are listed in Table **??**.

Table 1: Input Variables

Var	Physical Constraints	Software Constraints	Typical Value	Uncer
EEG	$0 < EGG$	$EEG_{\min} \leq EGG \leq EEG_{\max}$	30 μV	
$F(EEG)$	$0 < F(EGG)$	$F(EEG)_{\min} \leq F(EGG) \leq F(EEG)\mathrm{max}$	5 -15 HZ μV	

Table 2: Specification Parameter Values

Var	Value
$EEG_{\min}$	$10\mu V$
$EEG_{\max}$	$100\mu V$
$F(EEG)_{\min}$	$0.1hz$
$F(EEG)\max$	$100hz$

4.2.6 Properties of a Correct Solution

A correct solution must exhibit a reasonable signal power for each source point relative to the EEG signal power. Also, the power plots should be close to the expected active functional regions based on the experiment paradigm (the task the participant was given).

5 Requirements

This section provides the functional requirements, the business tasks that the software is expected to complete, and the nonfunctional requirements, the qualities that the software is expected to exhibit.

5.1 Functional Requirements

R1: EEGSourceLocalizationshall verify that the input data are valid. Frequency component of the EEG data should be checked to see if the data is band-pass filtered. Also, EEGSourceLocalizationshould ask the user to confirm if the artifact removal has been done. An error message shall be displayed if input data are invalid.

R2: EEGSourceLocalizationshall verify that the covariance matrix of the input data is not rank-deficient. An error message shall be displayed if input data is invalid.

R3: A topographical plot of the electrode signals (EEG) would be output in addition to power plot for confirmation.

R4: EEGSourceLocalizationshall provide correct calculate according to Instance Models according to the user's choice of which method to use.

R5: EEGSourceLocalizationshall plot the sources on a sliced MRI (and shall guarantee that the output file is the same pixel size as the input file) and 3D brain model.

5.2 Nonfunctional Requirements

R6: EEGSourceLocalizationshall be easy to be checked or tested.

R7: EEGSourceLocalizationwill not crash when a user provides invalid input.

R8: EEGSourceLocalizationshall be easy and satisfying for users to learn and use.

R9: EEGSourceLocalizationshall be able to run on Windows 10, macOS 10.14, and Ubuntu Linux 18.04. environments.

6 Likely Changes

LC1: The algorithms used to solve the inverse problem might change. Currently LCMV and sLORETA are going to be implemented. However, we might modify this.

LC2: The option of plotting the source activity in a 3D format might be added.

7 Unlikely Changes

LC3: The headmodel will not be simplified any further to a single shell sphere model or similar models.

LC4: 3D estimation of electrode locations based on a 2D map of the electrodes will not be added to the software and the exact coordinations should be provided.

LC5: This software will not be upgraded to be able to process MEG signals.

8 Traceability Matrices and Graphs

The purpose of the traceability matrices is to provide easy references on what has to be additionally modified if a certain component is changed. Every time a component is changed, the items in the column of that component that are marked with an "X" may have to be modified as well. Table ?? shows the dependencies of theoretical models, general definitions, data definitions, and instance models with each other. Table ?? shows the dependencies of instance models, requirements, and data constraints on each other. Table ?? shows the dependencies of theoretical models, general definitions, data definitions, instance models, and likely changes on the assumptions.

[You will have to modify these tables for your problem. —TPLT]

[The traceability matrix is challenging to maintain manually. Please do your best. In the future tools (like Drasil) will make this much easier. —TPLT]

	T??	T??	T??	GD??	GD??	DD??	DD??	DD??	DD??	IM??	IM??	IM??	IM??
T??													
T??			X										
T??													
GD??													
GD??	X												
DD??				X									
DD??				X									
DD??													
DD??								X					
IM??					X	X	X					X	
IM??					X		X			X	X		X
IM??		X											
IM??		X	X				X	X	X		X		

Table 3: Traceability Matrix Showing the Connections Between Items of Different Sections

	IM??	IM??	IM??	IM??	??	R??	R??
IM??		X				X	X
IM??	X			X		X	X
IM??						X	X
IM??		X				X	X
R??							
R??						X	
R??					X		
R??	X	X				X	X
R??	X						
R??		X					
R??			X				
R??				X			
R??			X	X			
R??		X					
R??		X					

Table 4: Traceability Matrix Showing the Connections Between Requirements and Instance Models

	A??	A??	A??	A??	A??	A??	A??	A??	A??	A??	A??	A??	A??	A??	A??	A??	A??	A??
T??	X																	
T??																		
T??																		
GD??		X																
GD??			X	X	X	X												
DD??							X	X	X									
DD??			X	X						X								
DD??																		
DD??																		
IM??											X	X		X	X	X		X
IM??											X	X			X	X	X	
IM??														X				X
IM??														X			X	
LC??				X														
LC??								X										
LC??									X									
LC??											X							
LC??													X					
LC??															X			

Table 5: Traceability Matrix Showing the Connections Between Assumptions and Other Items

The purpose of the traceability graphs is also to provide easy references on what has to be additionally modified if a certain component is changed. The arrows in the graphs represent dependencies. The component at the tail of an arrow is depended on by the component at the head of that arrow. Therefore, if a component is changed, the components that it points to should also be changed. Figure **??** shows the dependencies of theoretical models, general definitions, data definitions, instance models, likely changes, and assumptions on each other. Figure **??** shows the dependencies of instance models, requirements, and data constraints on each other.

9 Values of Auxiliary Constants

NA

[The following is not part of the template, just some things to consider when filing in the template. —TPLT]
[Grammar, flow and LATEXadvice:

- For Mac users `*.DS_Store` should be in `.gitignore`

- LATEX and formatting rules

    - Variables are italic, everything else not, includes subscripts (link to document)
        * Conventions
        * Watch out for implied multiplication
    - Use BibTeX
    - Use cross-referencing

- Grammar and writing rules

    - Acronyms expanded on first usage (not just in table of acronyms)
    - "In order to" should be "to"

—TPLT]
[Advice on using the template:

- Difference between physical and software constraints

- Properties of a correct solution means *additional* properties, not a restating of the requirements (may be "not applicable" for your problem). If you have a table of output constraints, then these are properties of a correct solution.

- Assumptions have to be invoked somewhere

- "Referenced by" implies that there is an explicit reference

- Think of traceability matrix, list of assumption invocations and list of reference by fields as automatically generatable

- If you say the format of the output (plot, table etc), then your requirement could be more abstract

—TPLT]