

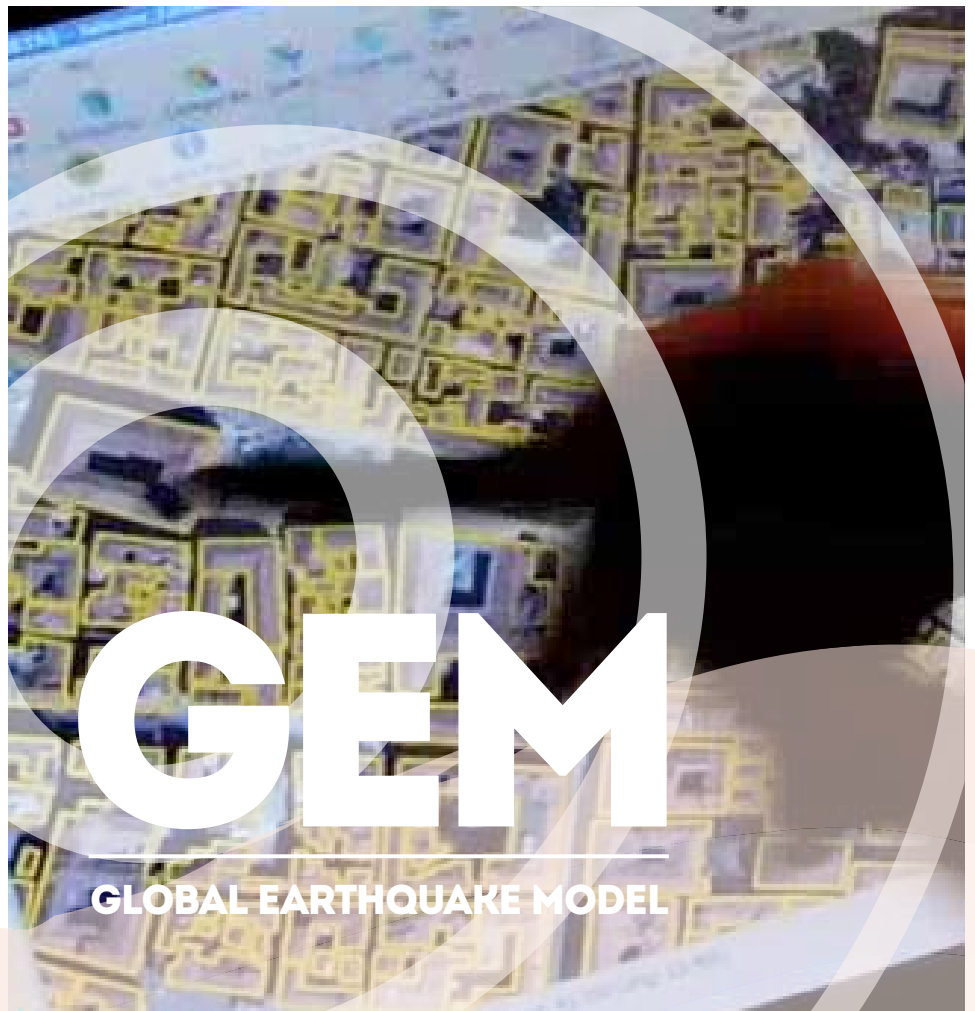
User guide: Tool for spatial inventory data development

GEM TECHNICAL REPORT
2014-05 V1.0.0



Hu, Z., C. Huyck, M. Eguchi, J. Bevington

DATA CAPTURE
TOOLS



User guide: Tool for spatial inventory data development

Technical Report 2014-05

Version: 1.0.0

Date: January 2014

Author(s)*: Hu, Z., Huyck C., Eguchi, M., Bevington, J.

(*) Authors' affiliations:

ZhengHui Hu, ImageCat, Long Beach CA

Charlie Huyck, ImageCat, Long Beach CA

Mike Eguchi, ImageCat, Long Beach CA

John Bevington, ImageCat, UK

Rights and permissions

Copyright © 2014 GEM Foundation, Hu Z., C. Huyck, M. Eguchi, J. Bevington

Except where otherwise noted, this work is licensed under a Creative Commons Attribution 3.0 Unported License.

The views and interpretations in this document are those of the individual author(s) and should not be attributed to the GEM Foundation. With them also lies the responsibility for the scientific and technical data presented. The authors have taken care to ensure the accuracy of the information in this report, but accept no responsibility for the material, nor liability for any loss including consequential loss incurred through the use of the material.

Citation advice

Hu, Z., C. Huyck, M. Eguchi and J. Bevington (2014), User guide: Tool for spatial inventory data development, GEM Technical Report 2014-05 V1.0.0, 60 pp., GEM Foundation, Pavia, Italy, doi: 10.13117/GEM.DATA-CAPTURE.TR2014.05.

Photo credits

Large cover image: © James Brown, GEM Foundation

Small cover image: © Chiara Casotto, GEM Foundation

<http://www.globalquakemodel.org/>

ABSTRACT

The Global Earthquake Model (GEM) aims to provide a set of tools and models for hazard and risk analysis. The fundamental aim of the GEM Inventory Data Capture Tools (IDCT) risk global component is to provide a set of software and accompanying user protocols to enable users to collect and modify building exposure information, which can be input into the Global Exposure Database and the Global Earthquake Consequences Database.

This report provides a background and user manual for the Spatial Inventory Data Developer (SIDD). SIDD serves as a conduit between base data sets derived from both ground observations and remotely sensed sources processed in large part by the IDCT, and a GEM Exposure File (GEF). SIDD allows users to process exposure data largely through a process of assigning “mapping schemes”, or statistical summaries of construction type, occupancy, era, and story height to “homogenous zones”, or areas with sufficiently similar structure type distribution. SIDD is designed as a tool to simplify complex processes that structural engineers and GIS analysts must undertake to develop building exposures, but does not simplify the decision-making process itself. Therefore, it is recommended for advanced users with a background in developing building exposure databases for catastrophe (CAT) modelling and loss estimation.

Keywords

inventory; exposure; spatial inventory data development; capture tool

TABLE OF CONTENTS

	Page
ABSTRACT	iii
TABLE OF CONTENTS	iv
LIST OF FIGURES	vi
LIST OF TABLES	vii
GLOSSARY	viii
1 Introduction.....	1
1.1 SIDD Overview.....	1
1.2 Outline of Report	3
2 SIDD Functionality and Workflow.....	4
2.1 SIDD Workflow	4
3 Loading Data and File Management.....	11
3.1 Creating and Opening Existing SIDD Projects.....	11
3.1.1 Creating a SIDD project.....	11
3.1.2 Opening an existing SIDD project	11
3.2 Homogenous Zones Data	14
3.3 Building Footprint Data	14
3.4 Population Grid	15
3.5 Survey and field data.....	16
3.6 Aggregation	16
3.7 Verify input data	17
4 Mapping Schemes and Modifiers	19
4.1 Mapping Scheme Overview	19
4.2 Creating a Mapping Scheme	19
4.3 Creating an Empty Mapping Scheme	20
4.4 Editing Mapping Schemes	21
4.5 Creating a Distribution of Heights and Date of Construction.....	23
4.6 Editing the Distribution of Attributes within a Given Homogenous Class.....	26
4.7 Creating a Mapping Scheme Using Building Survey Data	27
4.8 Accessing the Mapping Scheme Library	27
4.9 Modifiers	31
4.10 Saving Mapping Scheme	33

4.11 Modifying Average Dwelling Size and Replacement Cost	34
5 GIS Mapping	35
5.1 GIS Use	35
5.2 Toolbar	36
5.3 Thematic mapping.....	39
5.4 Adjusting Point Symbols.....	45
REFERENCES.....	46

LIST OF FIGURES

Figure 1.1 Overview of SIDD	1
Figure 2.1 SIDD Workflow, part 1.	4
Figure 2.2 Haiti Mapping Scheme – Tree Representation	6
Figure 2.3 SIDD workflow, part 2.....	7
Figure 2.4 IDCT data flowchart	8
Figure 3.1 Processing Options	12
Figure 3.2 The Data Input Tab	13
Figure 3.3 Homogenous zones data	14
Figure 3.4 Building footprint data.....	15
Figure 3.5 Population Grid.....	15
Figure 3.6 Survey and Field Data	16
Figure 3.7 Aggregation	17
Figure 3.8 Verifying data.....	17
Figure 4.1 Managing Mapping Schemes.....	19
Figure 4.2 Creating a Mapping Scheme	20
Figure 4.3 Initial Mapping Scheme for each Homogenous Zone	21
Figure 4.4 Editing the Mapping Scheme Branch, Drop Down	22
Figure 4.5 Editing the Mapping Scheme Branch, Sum of Weights	22
Figure 4.6 Adding Branches to the Mapping Scheme Dialog.....	23
Figure 4.7 Selecting the Level for Height Distribution	24
Figure 4.8 Editing the mapping scheme branch, height	24
Figure 4.9 Attribute Value Groups.....	25
Figure 4.10 Editing Mapping Scheme Branch, All height Entries.....	25
Figure 4.11 Selecting the Attribute for Editing	26
Figure 4.12 Editing mapping scheme beach, adjusting values	26
Figure 4.13 Create Mapping Scheme from a Survey	27
Figure 4.14 Managing Mapping Scheme, the Mapping Scheme Library	28
Figure 4.15 Managing Mapping Schemes, Assigning the Saved Mapping Scheme to a Branch	29
Figure 4.16 Managing Mapping Scheme, Final Distributions	30
Figure 4.17 Faulty Data in the Mapping Scheme Pane	31
Figure 4.18 Managing Mapping Scheme, Final Distributions	32
Figure 4.19 Editing building distribution modifier	32

Figure 4.20 Edit Modifier Value	33
Figure 4.21 Saving the mapping schemes.....	33
Figure 4.22 Edit Average Building Size and Cost.....	34
Figure 5.1 Sample output in the preview tab.	35
Figure 5.2 The preview toolbar.....	37
Figure 5.3 Zoom to Feature	37
Figure 5.4 Query feature for results	38
Figure 5.5 Query feature for input data	38
Figure 5.6 Layer Rendering Settings	39
Figure 5.7 Layer Rendering Setting, Selecting the thematic option	40
Figure 5.8 Layer Rendering Setting, making a colour ramp.....	40
Figure 5.9 Gradient colour Map	41
Figure 5.10 ColorBrewer ramp	41
Figure 5.11 Select Colour.....	42
Figure 5.12 Layer Rendering Setting, Adding the class values.....	42
Figure 5.13 Symbol Selector	43
Figure 5.14 Symbol Properties.....	43
Figure 5.15 Style Manger.....	44
Figure 5.16 Adjust map on the preview tab	44
Figure 5.17 Layer rendering setting, points.....	45
Figure 5.18 Layer rendering setting, change point symbol	45

LIST OF TABLES

Table 2.1 USGS PAGER mapping scheme based on dwelling fractions for Haiti (urban and rural, residential and non-residential) with GED hierarchy.	5
Table 2.2 Use cases for SIDD	9
Table 3.1 Valid Entry for the Data Input Tab, with use case and caveats.....	18
Table 5.1 Sample CSV export results	36
Table 5.2 Feature of the preview toolbar	37

GLOSSARY

Homogenous zones- areas with sufficiently similar structure type distribution to be characterized by a mapping scheme. These could be geographic regions of relatively uniform use such as a central business district, a manufacturing district, or a residential neighbourhood. Typically, these will correspond with use and occupancy classifications at the block scale but also consider known construction patterns. For example, an area of single family residential construction may be subdivided by era of development.

Mapping Scheme- A statistically-inferred distribution of building stock developed through field surveys or expert opinion. Used in conjunction with homogenous zones to construct an exposure database.

Mapping Scheme elements:

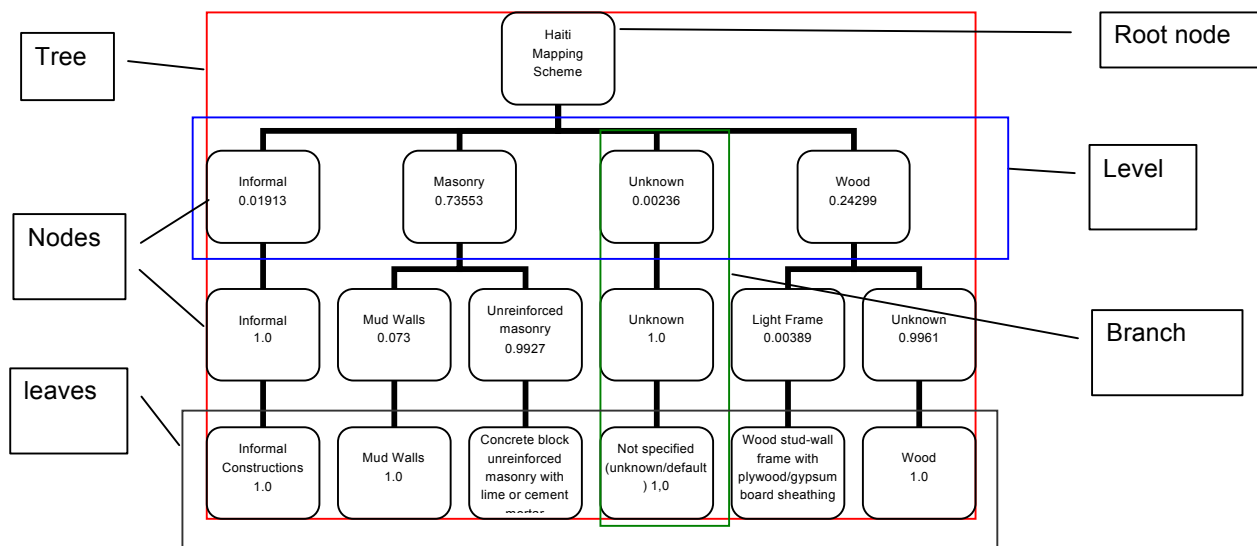
Node: contains an attribute value and its weight

Tree: a set of nodes order as a hierarchy under a single root node. it stores the complete mapping scheme

Branch: consists of any node and all nodes attached under it, also called subtree,

Level: nodes containing values and weights for same attribute. The sum of weights for all nodes with same parent must be 1.

Leaf/Leaves: node(s) with no additional nodes attached



PAGER STR- the taxonomy used in the USGS PAGER program.

PAGER mapping schemes- the mapping schemes used in the USGS PAGER program. These mapping schemes are based on assumptions about the distribution of dwelling units, not the distribution of the number of buildings. However, many engineers may want to use these assumptions as a starting point.

Secondary modifiers- attributes that are either optional to collect, or are not part of the primary structural hierarchy. Examples include soft-stories, roof type, and structural irregularity.

UNICEDE- a file format for CAT portfolio data developed by AIR Worldwide.

1 Introduction

This document provides background information and user documentation for running the Spatial Inventory Data Developer (SIDD) tool. SIDD is part of the GEM IDCT (Inventory Data Capture Tools) software for developing building exposure data in a GEF format (GEM Exposure File), suitable for analysis with the OpenQuake engine (Figure 1.1). The GEF is simply a comma-separated variable file (CSV) that conforms to the GEM exposure format. SIDD serves as a critical intermediary between raw sample data collected in the field, building footprint data extracted through remote sensing, and the final estimate of regional exposure contained within a GIS data set that can be loaded into the OpenQuake platform.

1.1 SIDD Overview

From a high level, SIDD develops, through a statistical process, regional estimates which are highly dependent on the development of “mapping schemes” and “homogenous zones.” Within SIDD, mapping schemes refer to allocating the number of buildings by occupancy type, detailed structure type, number of stories, era, or any number of “branches” defined by the user. A mapping scheme is then associated with homogenous zones within a GIS database through a user-defined class key. Several tools are provided for the review and modification of the mapping schemes, after which the user ultimately applies the building ratios to the aggregate building counts by homogenous zone to yield an exposure estimate consistent with the GEM taxonomy. The system is designed as a flexible workbench- allowing users to create exposure data with or without the aid of sampled data, and with or without the use of remote sensing data. An enhanced exposure database can be created with as little as expert modification of the default mapping schemes.

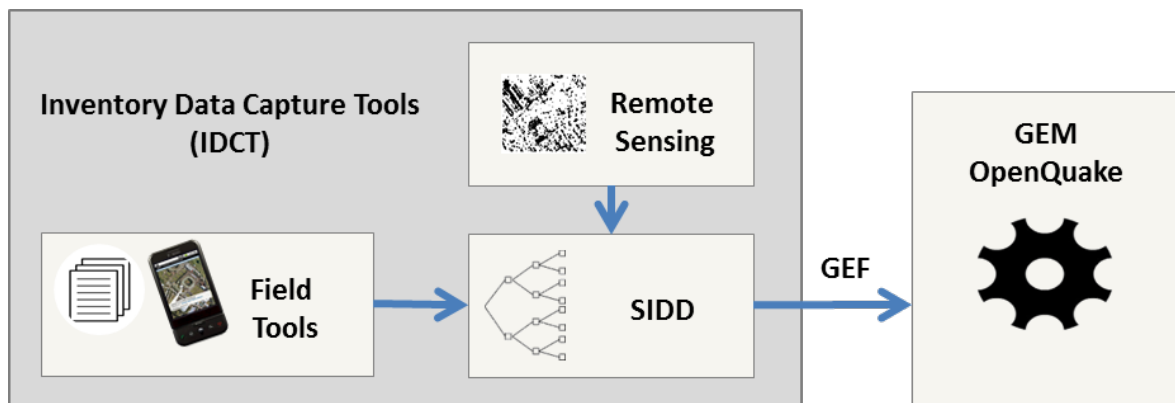


Figure 1.1 Overview of SIDD

The following example illustrates one possible use of the SIDD tool. For a given city with earthquake hazard, the public works department plans a moderate effort to develop a GEM-suitable exposure file. Working with the GIS department, aerial photography is used to apportion the city into 20 homogenous zones based on development patterns. These classes are defined by the city based on an assessment of the development patterns. The class names are for reference only and typically classify general occupancy or structural

patterns. For example, a given city engineer may designate 10 of the 20 zones as “Res- pre 1900”, 5 as “Res- post 1900”, 3 as “Com”, 1 is classified as “Ind” and, 1 as “Special.” In developing the homogenous zones, the building engineer weighs the benefits of increasing the number of classes of homogenous zone types (differentiation) against the need for a statistically robust sampling data set given the resources available. Referring to the GEM guidelines outlined in the IDCT protocol: In-field Sampling Protocol the engineer then designs a sampling strategy to collect building data. City staff use the android tools provided by IDCT to collect the data. This data is collected from all the staff members and aggregated into a single file. In a separate effort, GIS staff exports the building footprints from the city GIS into a “Shapefile” suitable for the tool.

The city engineer now has 3 key data sets for developing an exposure data set: 1) the footprint database, which will ultimately provide the number and square meter area of buildings; 2) the zones delineating land use and development patterns; and 3) the building samples consistent with the GEM taxonomy. These data are then loaded into SIDD. For each type of user zone defined in the homogenous zone dataset (i.e.: Res-pre 1900, Res-post 1900, Com, Ind, and special), SIDD creates a preliminary mapping scheme. The mapping scheme is simply a statistical summary of the percentage of sampled buildings in each category defined by the GEM Basic Building Taxonomy (Brzev et al., 2012). The building engineer then uses the SIDD tool in an iterative fashion to examine the mapping schemes, adjust the mapping schemes, review and adjust average building size and replacement cost estimates, and add additional “secondary modifiers” to characterize the personalized homogenous zone classes based on knowledge of development patterns throughout the city. Finally, when the mapping schemes appear satisfactory, they are applied to the exposure. The resulting file- the GEF- allocates the footprints aggregated for each individual homogenous zone into each GEM taxonomy class in the mapping scheme. In addition, the GEF file includes a file retaining the mapping schemes used to allocate the data.

The general process outlined in the example above is typical for the development of an exposure database for loss estimation. Without SIDD, the process of developing, applying, and verifying mapping schemes is quite complex. SIDD is designed specifically to simplify this process and allow experts to share and expand the assumptions inherent in developing exposure data through the GEFs. It is important to recognize that the process of creating these files requires a high level user with experience in the complications of creating exposure data through a sampling strategy. For example, there is a trade-off between the size of the homogenous zones used and the number of “remainders,” or allocations of less than one, which result from the application of the mapping scheme. A smaller zone leads to a more accurate loss assessment from a spatial perspective, but also results in more remaindered structures. Another trade-off is between the complexity of building classification and allocation of structures. With a more detailed building classification comes the ability to characterize the vulnerability with finer precision (given damage functions). However, with more classification categories there are more remaindered structures and a potentially less robust assessment in aggregate. SIDD is not designed to alleviate or override these critical decisions inherent in the process of developing building exposure, but to streamline the process of making them. Likewise SIDD is not designed to create a GEF from a fixed set of inputs, but empowers the user to develop exposure using many different inputs at different resolutions. This requires users that understand the benefits and pitfalls of using various data sets for exposure development. In addition to a general background in exposure development, users will also need to be familiar with the IDCT tools, the IDCT In-field Sampling Protocol, the GEM Building Taxonomy, and the OpenQuake platform in general.

1.2 Outline of Report

This document provides a basic presentation of SIDD and a user manual. Section 2 discusses further what can be done with the SIDD tool and provides illustrative examples. Section 3 provides documentation of file type for all inputs, as well as documentation of the resulting GEF. Section 4 provides a screen-by-screen description of the SIDD functionality. For the interested user, an early Preliminary Product Requirements Document for SIDD is available¹, although several iterations of comments during development significantly expanded the scope of the tool.

¹ <http://www.nexus.globalquakemodel.org/gem-idct/posts>

2 SIDD Functionality and Workflow

This chapter contains a complete flow chart illustrating the various uses of SIDD, followed by a discussion of sample applications. Referencing the workflow in Figure 2.1 and starting from the top- if building footprint data is not available, default building stock data is loaded from GEM. This data, developed by the Global Exposure Database (GED) consortium, will utilize a mapping scheme approach as well.

2.1 SIDD Workflow

GED assumes building stock characteristics largely from population statistics and statistics concerning numbers of occupants per dwelling or average floor area per person. This data will be difficult to obtain globally and it is thus highly recommended that a remote sensing data set is provided.

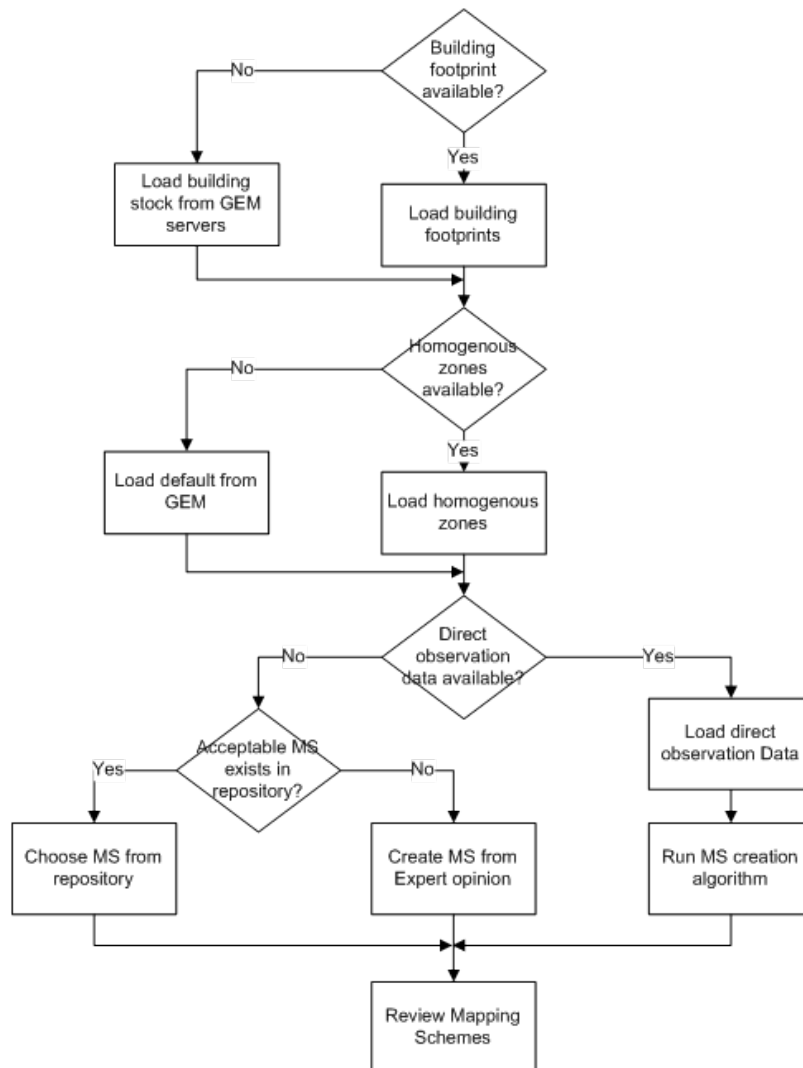


Figure 2.1 SIDD Workflow, part 1.

If direct observation data has been collected in the field, the data is loaded and used to create mapping schemes. Mapping schemes are created for each land use or texture class within the homogenous data. Users are then presented with a dialog to review the mapping schemes. If no mapping schemes are available, users can load mapping schemes from the repository, or from future mapping scheme libraries to be hosted by GEM. Default mapping schemes will be available for all countries with a significant earthquake risk. Table 2.1 presents a sample mapping scheme for Haiti derived as part of the GED effort using data from the USGS PAGER project (Jaiswal and Wald, 2008). It is important to note these estimates are based on the percentage of people dwelling in a given type of building, rather than the percentage of buildings. However, if other users had conducted studies in Haiti or a country with comparable building stock, users could load the mapping scheme as a starting point. In the event there was not a suitable alternative and the default mapping scheme was not deemed sufficient, a mapping scheme would have to be created from scratch based on expert opinion.

Table 2.1 USGS PAGER mapping scheme based on dwelling fractions for Haiti (urban and rural, residential and non-residential) with GED hierarchy.

Material	Ratio	First Level	Ratio	Second_Level	Ratio
Informal	0.01913	Informal	1.00000	Informal constructions.	1.00000
Masonry	0.73553	Mud walls	0.00730	Mud walls	1.00000
		Unreinforced masonry	0.99270	Concrete block unreinforced masonry with lime or cement mortar	1.00000
Unknown	0.00236	Unknown	1.00000	Not specified (unknown/default)	1.00000
Wood	0.24299	Light frame	0.00389	Wood stud-wall frame with plywood/gypsum board sheathing.	1.00000
		Unknown	0.99611	Wood	1.00000

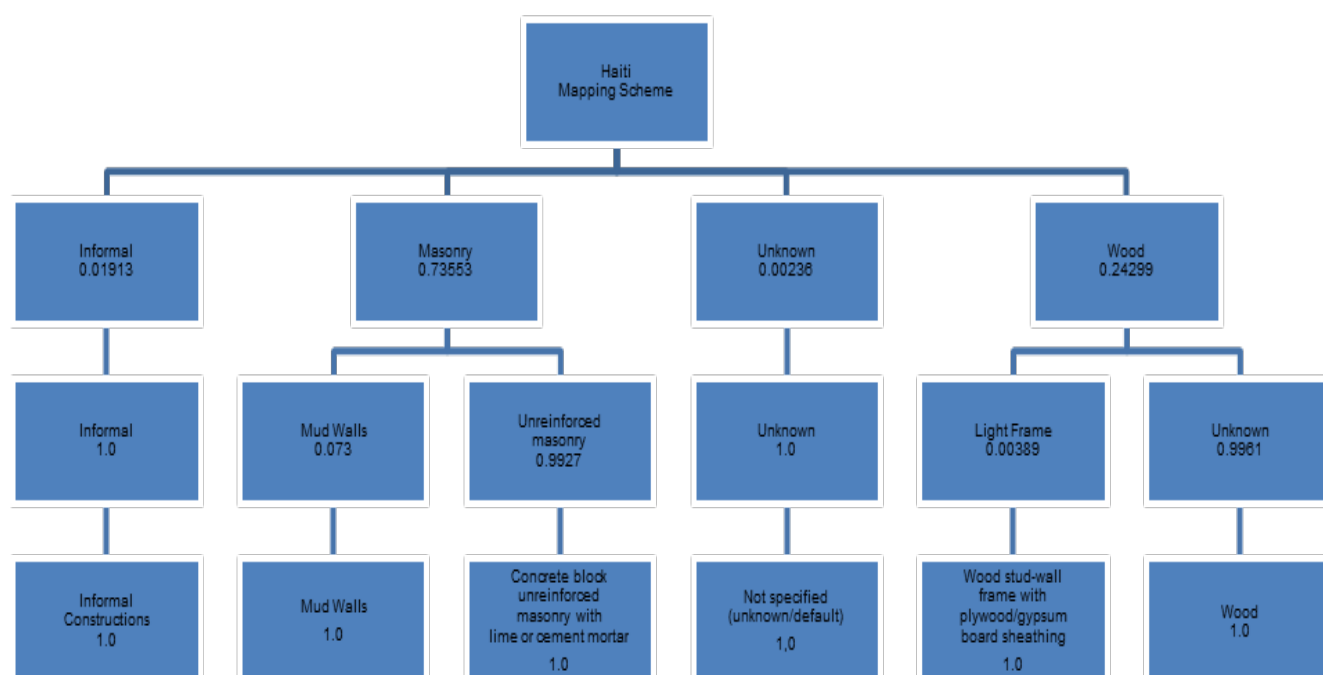


Figure 2.2 Haiti Mapping Scheme – Tree Representation

Figure 2.3 presents a continuation of the workflow. At this stage, a preliminary mapping scheme is available aggregated from field data, downloaded from a library, or entered from expert opinion. The user must review the mapping scheme carefully. It is anticipated that the first mapping scheme presented will not be sufficient. For example, while applying the mapping scheme approach in Indonesia, an 8-story hotel was surveyed (ImageCat, 2011). Although 8-story hotels are very rare in the area, because of the number of sites surveyed, a highly disproportionate number of 8-story buildings were represented in the mapping scheme. Users should visually check the numbers for anomalies and where appropriate “pare” the mapping scheme tree, or delete branches with a less than significant number of samples. In addition, users may want to combine various sources of mapping schemes to make a more robust set of assumptions. For example, in the USGS PAGER-based dwelling fraction mapping scheme for Haiti illustrated in Table 2.1, 24.299% of the development is wood frame and of that wood frame development, 99% is “unknown,” or has no additional classification provided. Given the additional information that has become available in Haiti following the earthquake in 2010, users may be able to designate the type of wood frame buildings in the region without significant effort.

The term “homogenous zones” is meant to convey to the user the importance of delineating zones which are sufficiently homogenous as to be characterized by a single mapping scheme. Even in highly heterogeneous areas where assigning a single scheme is difficult, these areas will ultimately represent a class for building allocation. For example, a historic central business district that also contains new development may be identified by the user as a single zone with a special type, even though it is not particularly homogenous. Homogeneity is an ideal, but often not feasible. It was chosen over the more common term “land use zones” in order to emphasize the ability to delineate areas based on any criteria that may be important for the characterization of risk- including, but not exclusive to land use.

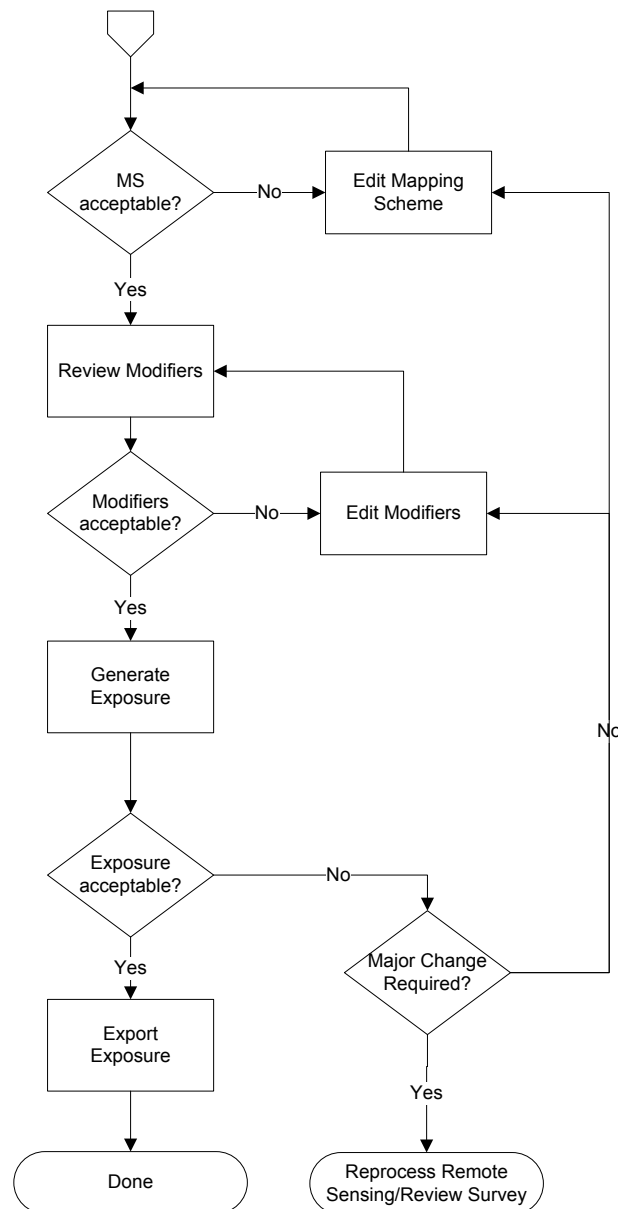


Figure 2.3 SIDD workflow, part 2.

Once the primary structural attributes are reviewed, modifiers are reviewed. Modifiers are seen as attributes that are either optional to collect or for some other reason not part of the primary structural hierarchy. Examples include soft-stories, roof type, and structural irregularity. If these data have been collected by the field team, they will be available to edit from within SIDD. In addition, users will have the option to add their own assumptions on modifiers. Although modifiers will not be a part of the mapping scheme tree, users will have the ability to link the secondary attributes to a specific branch of the tree—effectively applying the distribution to all the leaves on that branch.

Users will need to review the mapping scheme graphically in a mapping interface to assure that the results are reasonable. This is anticipated to be an involved process, where users “sanity check” homogenous zones to make sure that the number and distribution of buildings is appropriate. Users can compare between areas to make sure that areas with similar densities and construction types yield similar results. Users will want to note situations where a mismatch between surveyed data and homogenous zones can yield suspicious

results. For example, with the 8-story hotel in Indonesia referenced above, the photo of the structure was taken of the structure from across the road and the road was used as a boundary for a homogenous zone (between primarily commercial and industrial). Reviewing the survey data carefully to determine the source of the discrepancy made a considerable difference in the results. Adjustments of the various zones should be anticipated as a part of the process. When changes to the data are required, users will need to edit the base layers in a GIS and recreate the mapping schemes. Where users have specific “branches” they have edited and want to apply again, they will have to retrieve these numbers from the library or on disk and enter the numbers again. This iterative approach can take several days of application until a region is acceptable. Once the exposure is acceptable, the user generates the final GEF and it can be analysed in OpenQuake- either on disk or in an online environment. Significant efforts should be submitted to the GEM Secretariat to be considered for the GEM Global Exposure Database (GED). In the interim period before the OpenQuake platform is available, there is a form on GEM Nexus where users can upload level 2 data: <http://www.nexus.globalquakemodel.org/gem-idct/idct-pages/SIDD-data-submission-form>.

Figure 2.4 illustrates how SIDD rests within the IDCT group of tools. Data is loaded into SIDD from the remote sensing module (task 1) and the field data collection tool (task 2) and processed into a GEM suitable file. In addition, the user will typically load a GIS file with homogenous zones which roughly delineate land use patterns, development patterns, or any region which the user deems appropriate for assigning a unique allocation of structural attributes embedded in the mapping scheme.

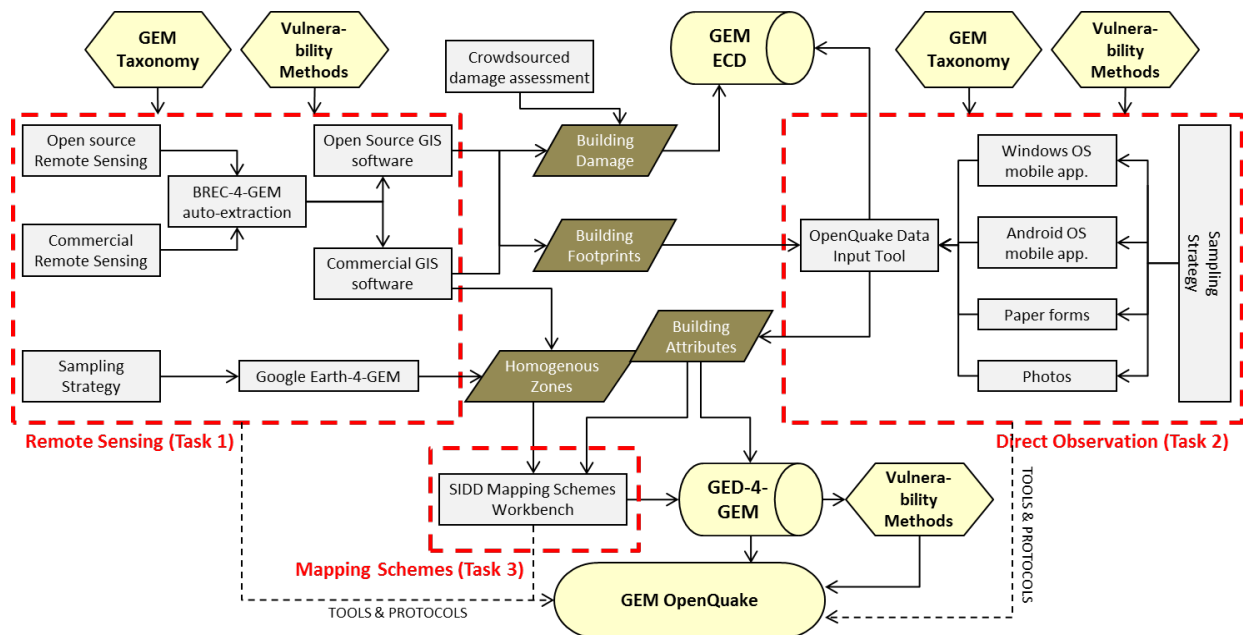


Figure 2.4 IDCT data flowchart

As described in the sections above, exposure data can be created from a very comprehensive study or cursory adjustments of mapping schemes for the purposes of sensitivity analysis. This flexibility was incorporated as a design requirement to accommodate as many users as possible. The table below presents several use cases that were envisioned when designing the system. Each use case has a short description and an overview of the SIDD methods, as well as other tools and techniques that might need to be incorporated.

Table 2.2 Use cases for SIDD

Use Case	Description	SIDD methods
Update national default mapping schemes	Updating default assumptions of national-level construction practices inherent in the GED	Users can download default exposure data from the GEM GED database along with the default mapping schemes. Mapping schemes can then be updated and applied. May take a considerable amount of time for large countries. Updates at this level need to be made using expert knowledge, unless used for sensitivity analysis.
Update regional mapping schemes	Default mapping schemes applied in the GEM GED are at the national level, although construction practices may vary in known ways throughout a country.	Users can download the default exposure and mapping schemes from the GEM GED. Mapping schemes can be replicated, adjusted, and applied to various regions in the country. If regional zonation is adequate using political boundaries, this data can be used as a base, otherwise, users will need to create custom zones to delineate region. Updates at this level need to be made using expert knowledge, unless used for sensitivity analysis.
Incorporate building data sets (attributes)	Building data is typically available for select buildings within an area, but sometimes data may be available representing every building. Examples may include U.S. tax assessor data or postal code databases. These will typically lack attribute information to assess on a building-specific level.	Complete data can be aggregated into zones by building count and square meter area, and ultimately allocated using the mapping schemes. Although typically there will not be enough information to assign structure type, some information may be available such as a fire-code classification or land use that can be assessed in aggregate by zone and used to adjust the mapping schemes. Updates such as these require users with significant database expertise to characterize the mapping schemes outside of the SIDD platform.
Incorporate building data sets (footprints)	Building footprints are often available for local municipalities based on digitization of aerial photographs. In addition, these data can be automatically extracted from the same data sets with varying degrees of success depending on the built environment. Although building footprints are seldom accompanied with attribute data to assign structure type, in aggregate, they significantly increase the accuracy of estimated number of buildings in the default GED database.	Building footprint data is directly imported into SIDD and aggregated to any zonal data provided. If accompanied by height in stories, building footprint data can be used as a guide in the creation of homogenous zones and assignment of mapping schemes. This will require database expertise.
Incorporate aggregate exposure data	In some cases, building data may be available at the national level, but not distributed in raw form due to privacy or security concerns. Typically, this data is aggregated to a grid or	Aggregate building data can be directly imported into SIDD. If structural data is supplied, users may be able to create data for OpenQuake directly. Typically, minimal

Use Case	Description	SIDD methods
	postal code.	structural data will be supplied that can be used to guide creation or assignment of mapping schemes. These methods can also be used with census data.
Calculate replacement cost	The GEM GED provides basic estimates of replacement costs that are only applicable to residential construction, and do not vary by construction type. These data are developed from global assumptions.	Replacement cost estimates can be updated by collecting square meter area and replacement cost during the sampling phase. This data will be used to create average square meter area and replacement cost for each structure type, which can be adjusted by the user. Another way to develop total replacement cost for a region is to collect square meter estimates in the field, and apply a per-unit cost based on published estimates for the area of interest.
Characterize local building patterns	It is common for development patterns to be strong indicators of construction practices, either due to land use, era of development, or other factors.	With local knowledge, the city engineer or similar expert can typically work with a GIS department to delineate homogenous zones to characterize patterns of development that can directly be tied to mapping schemes within SIDD.
Incorporate land use data	Construction practices can vary significantly by land use practices, often available in vector datasets at the national, regional, or local level. Urban density may even be inferred from moderate or low resolution satellite data- which can be used to develop crude land use zones.	Land use data may be directly imported into the SIDD program and used as homogenous zones. Users would then need to adjust default mapping schemes to link directly to the land use classes or develop new mapping schemes using surveys or other methods.

3 Loading Data and File Management

SIDD is divided into 3 main sections: loading data, mapping schemes and modifiers, and mapping results and base data through the preview tab. This section primarily describes the base data sets and how they are loaded into the system.

3.1 Creating and Opening Existing SIDD Projects

When the user first opens the program by clicking on the SIDD executable, the program will take up to a minute to load while the files are configured. At the first screen, all options are disabled until the user creates a new project or opens an existing project.

3.1.1 *Creating a SIDD project*

To create a project, select file, Create Project, Blank Project, name the file and save it on the hard drive. The resulting file has a .sidd extension and is a SQL lite file. It is strongly recommended that as a user works through the various steps, that projects are saved frequently and that the user selects “Save Project As” frequently to create iterative backups. This complex processes, it is common to make mistakes and want to revert to an older file. In addition, this will guard against file corruption.

The following data is stored in a project file:

- 1) Paths to input data
- 2) Mapping schemes and mapping scheme assignments
- 3) Modifiers

The following data is not stored in the project file:

- 1) Input data
- 2) GEF files

Thus, the user is advised to take the following precautions:

- 1) When distributing data, make sure to distribute the base data. If the files have been moved, users will have the opportunity to re-establish paths to the data.
- 2) After building an exposure database, the user should export the data to a GEF file. Otherwise, this data will have to be recreated next time the project is opened.

3.1.2 *Opening an existing SIDD project*

To open an existing project, select open existing from the file menu. The file may take a considerable amount of time to load, as the GIS data must be loaded into memory.

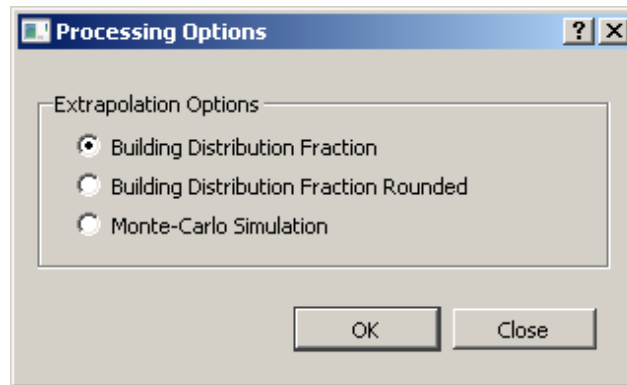


Figure 3.1 Processing Options

An important setup option is provided under the menu option: options, as displayed in Figure 3.1. These options determine how SIDD distributes building counts to zones based upon the mapping schemes.

In option one, fractional buildings are possible. In this case, if a mapping scheme has 100 equally probable building types, and there are 50 building to be distributed, each type will be assigned a half a building, but there will be 50 buildings (or 100 half-buildings) assigned to the zone.

In the second option, the 0.5 figure is rounded, in this case to 1 building for each type, resulting in 100 buildings for the zone. Note, with 101 equally-probable building types, there would be 0 buildings assigned to the zone.

The third option runs a Monte-Carlo simulation with replacement. There will be 50 zones assigned to the zone. The apportionment will dictate the likelihood of each class being selected. There will be no more than 50 classes selected, and most likely many less, with many classes having more than one building.

Although some modellers strongly prefer one method to the others, none are ideal. With large numbers of buildings in a zone, the methods should begin to be closer. The example above was designed to illustrate an extreme; however, with smaller zones and more classes, the problem is significant. Engineers and geographers desire precision when classifying buildings and representing risk spatially, but these benefits are not realized when the ratio of the number of building classes to the number of buildings being distributed is too high. Users should of the SIDD platform should ideally be familiar with these trade-offs, or at the very least be very cognizant that SIDD is a workbench designed for engineers apply these techniques, but does not simplify or alleviate the decision making process.

Spatial Inventory Data Developer

File Options Help

Data Input Mapping Scheme Modifiers Preview

Define required input data

Specify input data
Different combination of data can be used toward generating mapping schemes and exposure.

Homogeneous zones data

What type of data do you have for zones?

- ☐ No Data
- ☒ Homogeneous zones
- ☐ Homogeneous zones with building count

Select file: C:/SIDD/Argentina zones.shp ...

Select field containing zone identifier: Class

Select field containing building count:

Verify projection: WGS 84

Building footprint data

What type of data do you have for building footprints?

- ☒ No Data
- ☐ Building footprints without number of stories
- ☐ Building footprints with number of stories

Select file: ...

Select field containing number of stories:

Verify projection:

Population Grid

Population

- ☐ No Data
- ☒ Population Grid from GED

Select file: C:/SIDD/Argentina pop.shp ...

Select field containing population: pop_count

Average Number of People Per Building: 6.0

Verify projection: WGS 84

Survey & field data

What type of survey / field data do you have?

- ☒ No Data
- ☐ Sampled buildings from survey area
- ☐ Complete building stock/survey area

Select file: ...

Aggregation

How do you wish to aggregate your output data?

- ☒ Output into defined zones
- ☐ GED Compatible 30 arc-second grid

You have supplied the following types of data

- ☒ Footprint
- ☒ Zones
- ☒ Population Grid
- ☒ Survey

Output data aggregation

- ☒ Zone
- ☒ GED Grid

Following dataset required for building exposure
-No action is defined for type of data provided

Verify Input Data

Project Loaded

Figure 3.2 The Data Input Tab

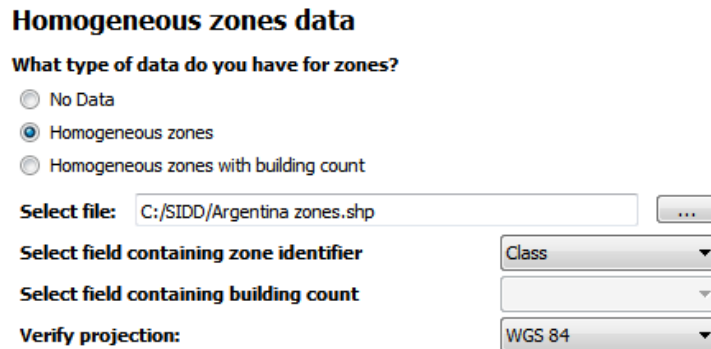
Once a project has been opened or created, the user has the opportunity to adjust or enter the paths to the base data to be analyzed. These options are available on the Data Input Tab. The Data Input Tab provides 6 main sections which should be examined in turn:

- 1) Homogenous zones data
- 2) Building footprint data
- 3) Population grid
- 4) Survey & field data
- 5) Aggregation
- 6) Verification

Each of these sections will be described in detail below.

3.2 Homogenous Zones Data

Figure 3.3 below displays the homogenous zones section of the Data Input screen, populated with sample data.



Homogeneous zones data

What type of data do you have for zones?

☐ No Data
☒ Homogeneous zones
☐ Homogeneous zones with building count

Select file: C:/SIDD/Argentina zones.shp

Select field containing zone identifier Class

Select field containing building count

Verify projection: WGS 84

Figure 3.3 Homogenous zones data

Homogenous zones must be provided in ESRI Shapefile format with a valid .prj file. The “Verify projection” dialog is provided simply to allow the user to verify that the projection has registered correctly. Homogenous zones can be postal codes, census tracts, city boundaries, land use maps, detailed zones assessed from aerial photographs, or automatically extracted zones from image processing routines. As long as the user provides the zones in a shapefile format with an identifier, SIDD will process the data. This allows the user to process a wide variety of data sets, as discussed in the introduction. At the very least, one homogenous zone must be provided to establish the study area.

Users are able to enter homogenous zones with building count data included. This is not required if footprint data is provided. This feature is useful when buildings are provided as points or in aggregated format (as with census tracts). To analyse census tracts or similar data, it is advised that users add a column to the data to characterize various development patterns to link to mapping schemes. When complete databases of building inventory are available without proper attribution, users should aggregate the counts into homogenous zones, add these counts as a field in the shapefile, and select the file in the “Select field containing building count” field.

Lastly, the user will need to enter the “zone identifier” under the option “Select field containing zone identifier.” The zone identifier will be a field with a text identifier indicating the type of homogenous zones. A descriptive identifier is suggested, such as “residential”, “CBD”, or “Pre-code commercial.” Minimizing the total number of classes is preferable, since mapping schemes will need to be assigned to each class. Users may want to count the unique number of classes though a GIS or database program before entering the data into the system.

3.3 Building Footprint Data

Figure 3.4 below displays the building footprint section of the Data Input screen.

Building footprint data

What type of data do you have for building footprints?

- ☒ No Data
☐ Building footprints without number of stories
☐ Building footprints with number of stories

Select file: ...
 Select field containing number of stories:
 Verify projection:

Figure 3.4 Building footprint data

Building footprints can be obtained in a variety of ways, including digitization in Google Earth, in a GIS, with remote sensing software, or often obtained through the local government GIS. SIDD accepts footprints with or without the number of stories. The number of stories is used internally by SIDD to create mapping schemes given field deployment data, as well as develop the final counts and square footage information. The heights, when provided, serve as a stratified sampling field in accordance with the *In-field Sampling Protocol* document, and should be supplied where feasible.

It is important to recognize that the building footprint data are 1) not attached to specific field recordings, and 2) are not exported from SIDD. SIDD creates building exposure data based on the homogenous zones, and the building footprints, when provided, are used to populate the count of buildings for those zones.

Data must be provided in an ESRI shapefile with a valid .prj file, preferably in a geographic, or latitude and longitude projection and a datum or WGS 84. The “Verify projection” dialog is provided simply to allow the user to verify that the projection has registered correctly. Users with building footprints simply supply nature of the file (with or without footprints), select the file through the file open dialog, and select the field containing number of stories when available.

3.4 Population Grid

Figure 3.5 below displays the population grid section of the Data Input screen, with sample input.

Population Grid

Population

- ☐ No Data
☒ Population Grid from GED

Select file: ...
 Select field containing population:
 Average Number of People Per Building
 Verify projection:

Figure 3.5 Population Grid

The population grid option is for situations where users do not have building footprints or building counts associated with the homogenous zones. It is appropriate for large regions of analysis, such as the update for Argentina demonstrated here. As a crude, national proxy, the population can be obtained from GEM, and the average number of people per building estimated to yield the number of buildings. If this data is available regionally, then the user is advised to aggregate this data into zones using the methods described in the homogenous zone section above.

Data must be provided in an ESRI shapefile with a valid .prj file, preferably in a geographic, or latitude and longitude projection and a datum or WGS 84. This is not available directly from GEM, and users will have to make a valid shapefile from a CSV using off-the-shelf GIS software (see Vicini *et al.* 2013 for more information). The “Verify projection” dialog is provided simply to allow the user to verify that the projection has registered correctly. This is the projection supplied by GEM. To enter in the data, simply select the “Population Grid from GED” option, select the file, and select the column containing the population. If no estimate is provided for the average number of people per building, the program defaults to a value of 1 (i.e., no actual default is provided). If building count data is available in this format, it can be loaded into the population grid. By entering a default value of one person per building, the value in the population field will be transformed into the correct number of buildings.

3.5 Survey and field data

Figure 3.6 below displays the “Survey & field data” section of the Data Input screen.

Survey & field data

What type of survey / field data do you have?

☒ No Data

☐ Sampled buildings from survey area

☐ Complete building stock/survey area

Select file:

Figure 3.6 Survey and Field Data

The survey data will be loaded directly from the field survey tools- either the windows tool or the android tool. In most circumstances, there will be more than one device used to collect the data, and it will be aggregated before being loaded into SIDD. The file, given as a .db extension, is a SQL lite database.

To load the survey data into SIDD, first select whether the survey represents “sampled buildings from the survey area” or “complete building stock/survey area.” The typical case will be sampled data, which will be used to develop the mapping schemes. If complete data is available, this data will be processed directly into a GEF, aggregating into a GEM compliant Grid.

3.6 Aggregation

Figure 3.7 below displays the aggregation section of the Data Input screen.

Aggregation

How do you wish to aggregate your output data?

- ☒ Output into defined zones
☐ GED Compatible 30 arc-second grid

Figure 3.7 Aggregation

Aggregation in this context refers to the format of the output data. The first selection, “Output into defined zones”, will output a vector-based GIS data set corresponding to the homogenous zones. There is a many-to-one relationship in the final database between objects (in this case, homogenous zones) and structure types. In order to be consistent with the OpenQuake format, the polygon is replicated for each structure type, so that each line of data in the table can be analyzed independently by the OpenQuake engine.

The “GED Compatible 30 arc-second grid” option will output the data in a 30 arc-second “grid” (approximately 1 kilometre square depending on location). The grid is represented as a series of points representing the grid centroids. As with the polygons, there is a many-to-one relationship to the points, and so there are many overlapping points at each location representing the various construction types.

Whether to use the GED data or the zones or the 30-arc second grid depends largely on the scale of the analysis. The grid method will result in the area associated with a grid point representing multiple regions in the homogenous zone file, so this method should only be used when a 30-arc second posting of results is consistent with the resolution of analysis.

Users are advised that with a vector file, the OpenQuake engine will simply analyze the location of the centre, so for large areas the 30 arc-second grids are preferable.

3.7 Verify input data

You have supplied the following types of data

- ☒ Footprint
- ☒ Zones
- ☒ Population Grid
- ☒ Survey

Output data aggregation

- ☒ Zone
- ☒ GED Grid

Following dataset required for building exposure
-No action is defined for type of data provided

Verify Input Data

Figure 3.8 Verifying data

Once data has been entered in the above dialogs, the user selects the “Verify Input Data” option. The “verify input data” option simply makes sure that the user has entered enough data to be analyzed. Users must ultimately have a source for building counts, and zone data that it is applied to. The green checks and red crosses simply verify to the user what data has been loaded in.

The table below provides a brief summary of the valid options that are will allow the user to continue to additional tabs in the SIDD program, accompanied description of the use case, as well as any caveats.

Table 3.1 Valid Entry for the Data Input Tab, with use case and caveats

Use Case	Homogenous Zones	Sampled Surveys*	Building Footprints	Population Grid from GED
Base data available to user, but collection of survey data not feasible. Requires expert user for mapping scheme development or adjustment.	X		X	
This option is only valid when homogenous zones have counts. Counts might be available from census data or provided in other aggregated form, or may be aggregated from point-level data. Requires expert user for mapping scheme development or adjustment.	X			
May be appropriate for enhanced regional or national studies where building count is not available. User must supply an estimate of persons per building. Requires expert user for mapping scheme development or adjustment.	X			X
Ideal method of analysis, but may not be feasible for widespread assessment.	X	X	X	
Local analysis may harness building counts from other sources then footprints, which will need to be associated with homogenous zones before data entry.	X	X		
This option may be employed when no footprints or counts are available. Not an ideal option, as building counts will likely not be sufficient for a local analysis- which some users might deem appropriate given surveyed results. Requires an estimate of persons per building. Use this option with extreme caution.	X	X		X

* Complete surveys are valid given a study are for homogenous zones, and output aggregated by 30 arc second grid

4 Mapping Schemes and Modifiers

This section presents the mapping scheme and modifiers tools which are used to create mapping schemes that are consistent with the GEM Basic Building Taxonomy (Brzev et al., 2012).

4.1 Mapping Scheme Overview

Figure 4.1 below presents the mapping scheme tab. The basic functions of this tab are to create mapping schemes, either from scratch or from surveyed data, and editing mapping schemes- either to assure quality or to adjust existing schemes to match a new area. The text in this section will examine each of the features in Figure 4.1.

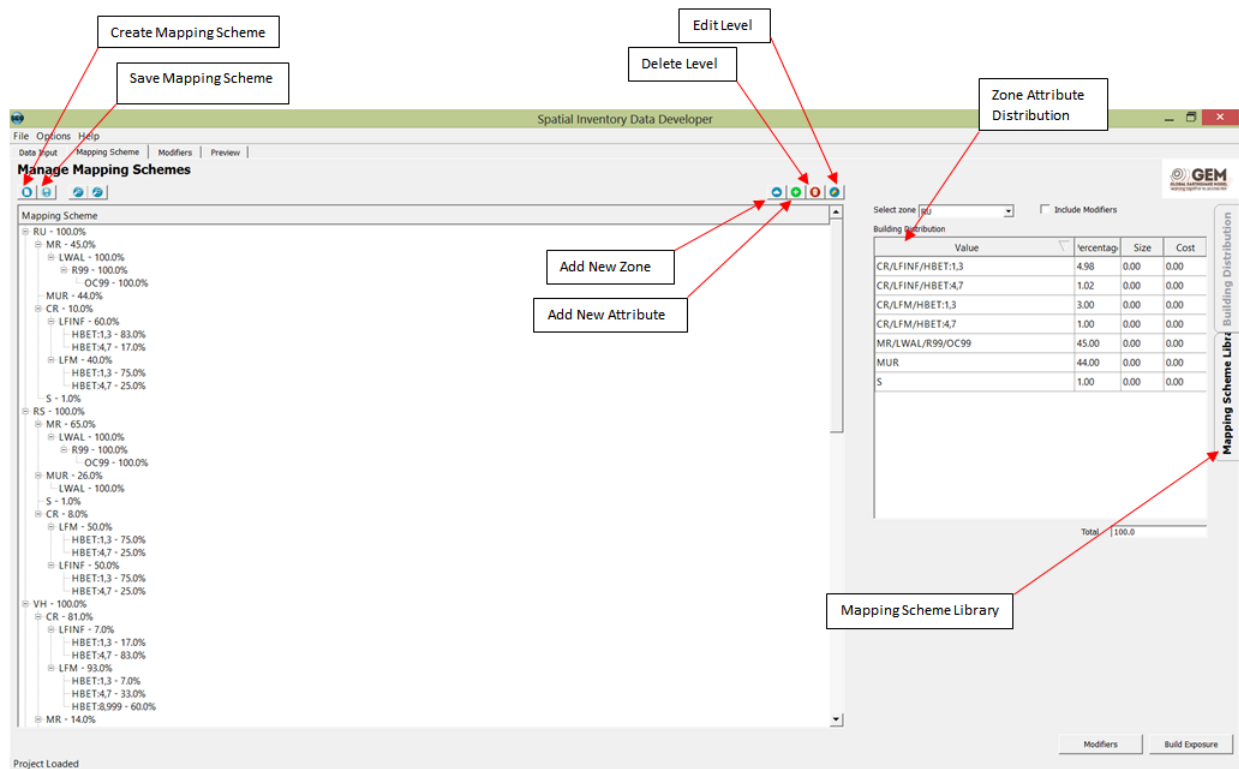



Figure 4.1 Managing Mapping Schemes

4.2 Creating a Mapping Scheme

Users have three options when creating a mapping scheme: build from an empty mapping scheme, create it from survey data or load a predefined scheme from the mapping scheme library. The sections below address each method.

4.3 Creating an Empty Mapping Scheme

Creating an empty mapping scheme allows the user to manually define and edit all attributes within each homogenous zone. Each branch within the scheme will have to be added by the user and a distribution of the attributes will need to be created. Users should have a thorough understanding of construction practices within the specified region and be comfortable assigning values and approximate distributions for each attribute.

To create an empty mapping scheme, first click on the “Mapping Scheme” tab, then click the “Create New Mapping Scheme” button  in the top left corner. A “Create Mapping Scheme” window will appear (see Figure 4.2). Choose “Create Empty Mapping Scheme” and click “OK”.

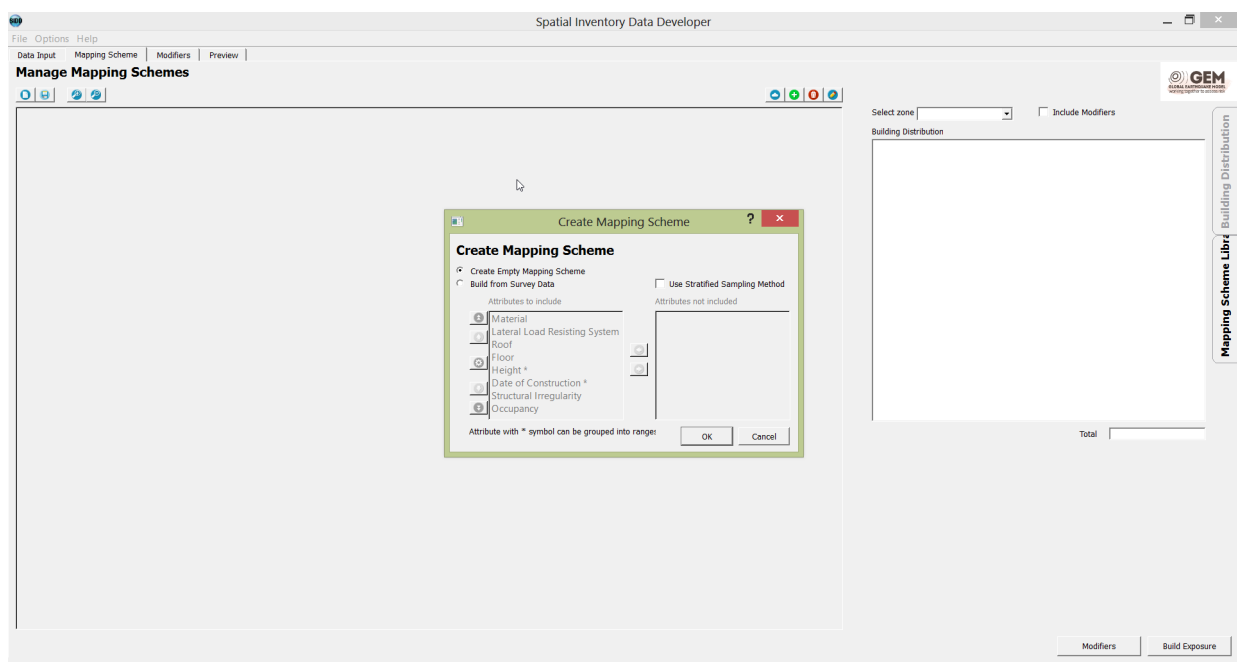


Figure 4.2 Creating a Mapping Scheme

A list of all homogenous zones previously loaded in the “data input” tab will appear in the “Mapping Scheme” window (see Figure 4.3). At this point, no attributes (structural type, occupancy, height, etc.) are associated with each zone and will have to be manually inputted by the user.

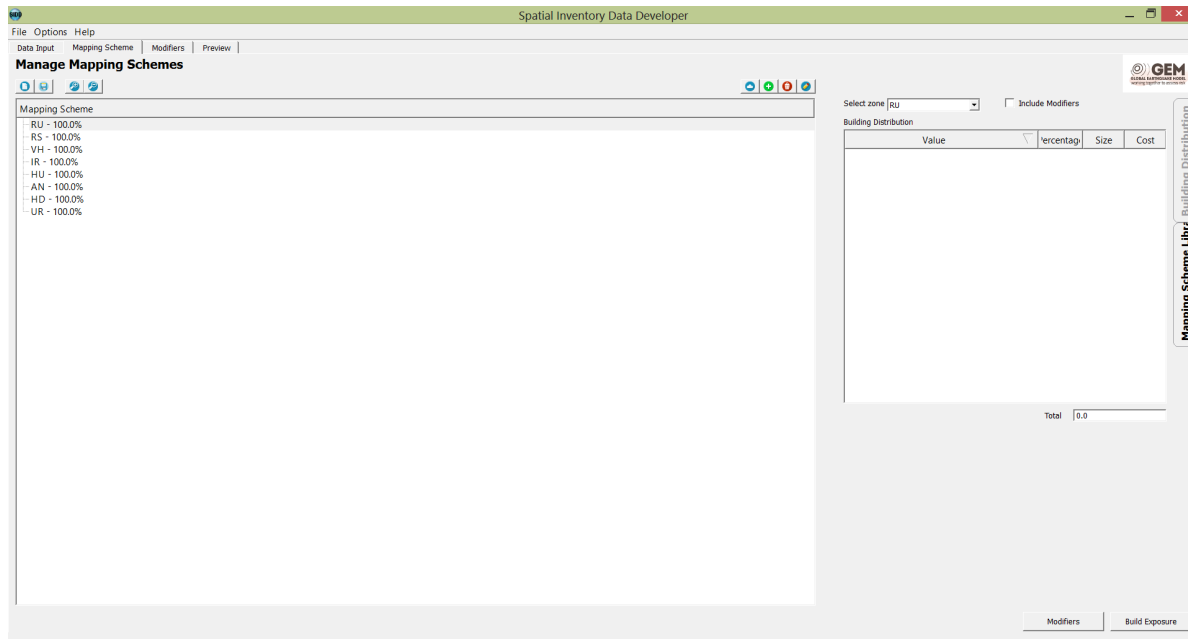




Figure 4.3 Initial Mapping Scheme for each Homogenous Zone

4.4 Editing Mapping Schemes

SIDD allows users to define and adjust the following eight different attributes for each zone:

1. Material
2. Lateral Load Resisting System
3. Roof
4. Floor
5. Height
6. Date of Construction
7. Structural Irregularity
8. Occupancy

To edit a zone, highlight the respective zone in the mapping scheme window and click the “Add/Edit Child Node” . An “Edit Mapping Scheme Branch” dialog will appear with a drop down menu listing all attributes that can be manually adjusted (see Figure 4.4). Select the attribute to be adjusted.

The user will now need to create a distribution of attributes for the given land class. Click the “Add Value” button  and a blank cell with 0.00 % will appear. Double clicking the blank value cell will bring up a drop down menu that lists all available options for the given attribute.

Edit Mapping Scheme Branch

Attribute Name

Material

Value	Percentage
Concrete, composite with steel section	0.00
Concrete, composite with steel section	
Concrete, reinforced	
Concrete, unknown reinforcement	
Concrete, unreinforced	
Earth, reinforced	
Earth, unknown reinforcement	
Earth, unreinforced	
Masonry, confined	
Masonry, reinforced	
Masonry, unknown reinforcement	

Sum of weights: 0.0 %

Apply Close

Figure 4.4 Editing the Mapping Scheme Branch, Drop Down

Select the appropriate value and manually enter the percentage associated with the value. Repeat until all values are accounted for and click “Apply” (see Figure 4.5). The “Sum of Weights” must equal 100%, or the user will not be able to proceed. Click apply and repeat the process for the rest of the zones.

Edit Mapping Scheme Branch

Attribute Name


Material

Value	Percentage
C99	25.00
E99	25.00
S	25.00
W	25.00

Sum of weights: 100.0 %

Apply Close

Figure 4.5 Editing the Mapping Scheme Branch, Sum of Weights

Each level of the mapping scheme can be further expanded by following the same procedure described above. In the “Mapping Scheme” dialog (Figure 4.6), highlight the level you wish to expand and click the “Add/Edit Child Node” . Select the attribute to be evaluated and assign both a value and percentage for that level.

Once all zones have been completed, a distribution tree will be visible in the “Manage Mapping Schemes” window. Users can expand or collapse the tree by clicking the + or – adjacent to each level. Each level informs the user both the value and distribution (in %) of each attribute.

The “Mapping Scheme” tab also allows users to identify the distribution of attributes within a specified zone. The table on the right provides all possible distributions within the selected zone, as well as the percent of exposure, average size of building (if available, and the replacement cost (if available). Users can hover over any of the cells to get a definition of the attributes.

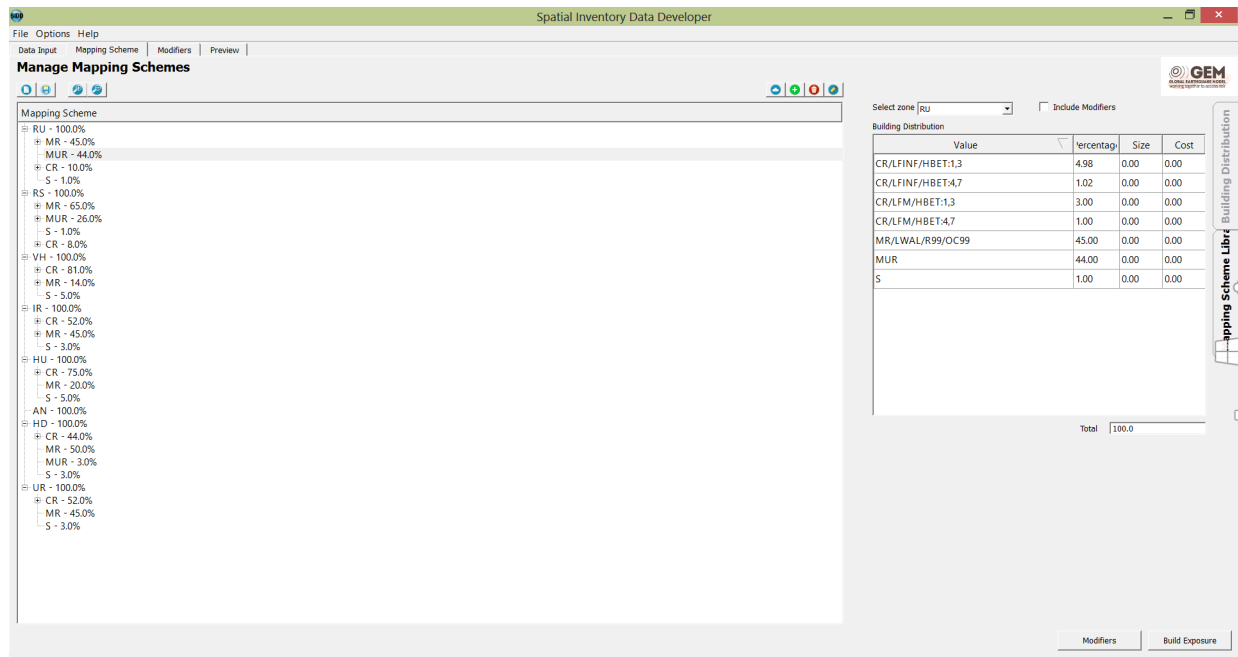


Figure 4.6 Adding Branches to the Mapping Scheme Dialog

4.5 Creating a Distribution of Heights and Date of Construction

Both the height and date of construction differ from the others attributes because it allows users to specify ranges (in stories for height and years for date of construction). Users will predefine these ranges and create a distribution as required.

To create a distribution of heights for a given zone, the user is will manually set the ranges for each height bin. First, select the level for which the height distribution will be created in the mapping scheme dialog (Figure 4.7).

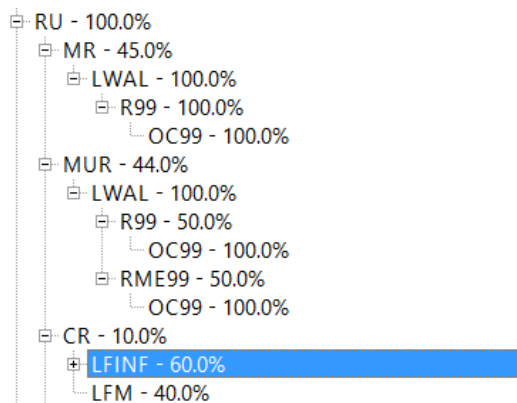


Figure 4.7 Selecting the Level for Height Distribution




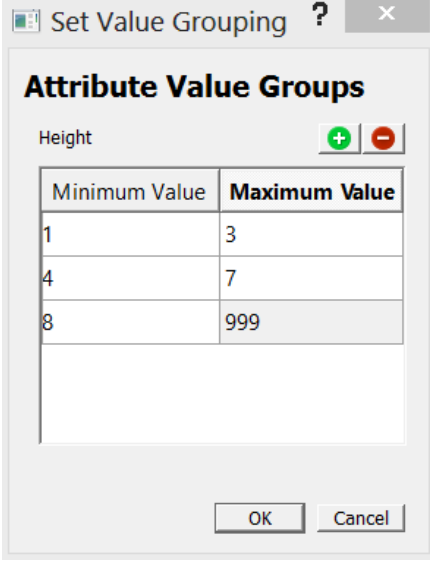
Click the “Add/Edit Child Node”  button and the “Edit Mapping Scheme Branch” window will appear (Figure 4.8). Choose “Height” from the drop down menu and click the “Edit Value Grouping” icon .

Figure 4.8 Editing the mapping scheme branch, height

An “Attribute Value Groups” window will appear. Bins containing the ranges for height can now be created by clicking the “add value group”  button.

Enter both the minimum and maximum height values. Repeat until all height distributions are created, then click “OK”.



Set Value Grouping ? x

Attribute Value Groups


Height + -

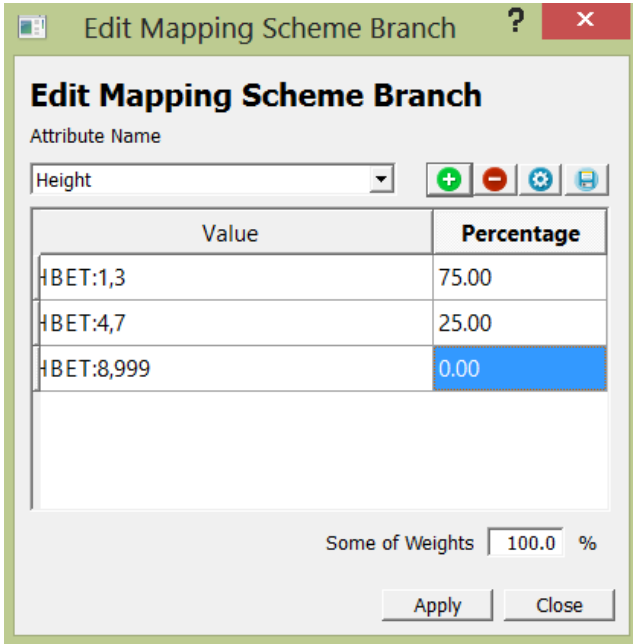
Minimum Value	Maximum Value
1	3
4	7
8	999

OK Cancel

Figure 4.9 Attribute Value Groups

Once all height distributions are created, click “OK” in the “Attribute Value Groups” window. This will take the user back to the “Edit Mapping Scheme Branch”.

Ensure “Height” is showing in the drop down menu and click the “Add Value”  button. Double click the blank cell below “Value” in the table and a drop down menu will appear. Choose one of the height distributions created and the percent value this distribution represents. Repeat for all height categories created and click “Apply”. Remember the “Sum of Weights” must equal 100% before you can proceed (Figure 4.10).



Edit Mapping Scheme Branch ? x

Attribute Name

Height + - ⚙️ 📄

Value	Percentage
HBET:1,3	75.00
HBET:4,7	25.00
HBET:8,999	0.00

Some of Weights 100.0 %

Apply Close

Figure 4.10 Editing Mapping Scheme Branch, All height Entries

The same process should be followed when creating a distribution of construction dates. Select “Date of Construction” from the drop down menu and create a range of dates instead of stories.

4.6 Editing the Distribution of Attributes within a Given Homogenous Class

Users can manually adjust the distribution of attributes within any homogenous zone classification, by using the “Edit Level” option. First click on the attribute that needs to be adjusted in the “Mapping Scheme” dialog.

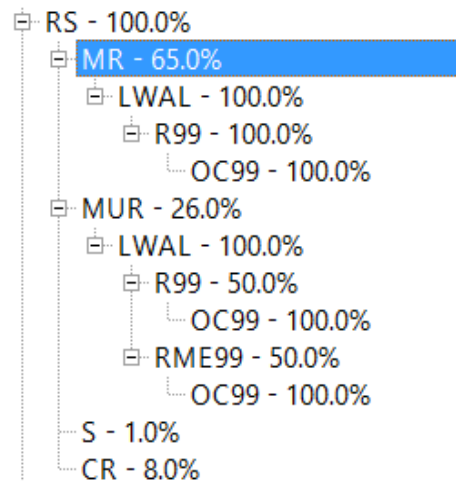


Figure 4.11 Selecting the Attribute for Editing

Click the “Edit Level”  and an “Edit Mapping Scheme Branch” window will appear.

Value	Percentage
CR	10.00
MR	45.00
MUR	44.00
S	1.00

Sum of weights 100.0 %



Apply Close

Figure 4.12 Editing mapping scheme beach, adjusting values

Both the “Value” and “Percentage” of the attribute can be manually adjusted.

4.7 Creating a Mapping Scheme Using Building Survey Data

A mapping scheme can also be created using building survey data. The building survey data has been collected by the Direct Observation Tools and has been used to create mapping schemes specific to the region.

To create a mapping scheme using survey data, first click the “Create New Mapping Scheme” button . A new window will appear and click on the “Build from Survey Data” option. The list on the left represent all attributes that will be included in the mapping scheme. To exclude any of the items, select the attribute and click the “Exclude Selected Attribute” button .

It is important to note the attribute first on the list will be the first level of the distribution tree. Users can manually adjust the location of each attribute by using the arrow buttons on the left.

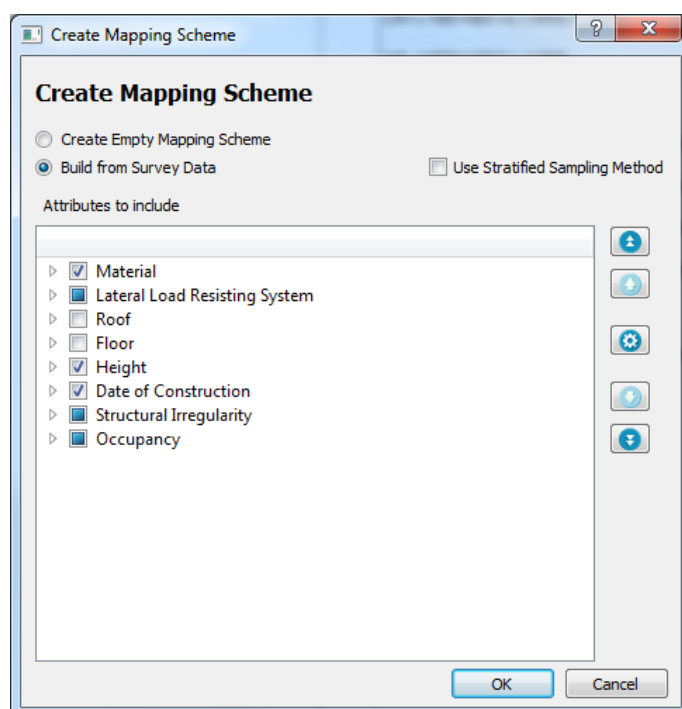


Figure 4.13 Create Mapping Scheme from a Survey

Once all required attributes have been included and the hierarchy of attributes has been established, click the “OK” button. The distribution of zones will appear in the mapping scheme window.

4.8 Accessing the Mapping Scheme Library

The mapping scheme library allows users to access default schemes created by the USGS PAGER project. These default schemes represent typical areas (Non-residential rural, non-residential urban, residential rural and residential urban) and can be used either as the finalized mapping scheme or a starting point if more detailed information is available.

The user will first need to load the land classes into the mapping scheme window using one of the two methods mentioned previously (creating an empty mapping scheme or building from surveys). Once the schemes are visible in the “Mapping Scheme” window, click on the “Mapping Scheme Library” tab on the right side of the program. Choose the country from the “Regions” box and select the data source from the “Types” box (Figure 4.14). Once these two options have been specified, a list of available mapping schemes will appear in the box directly below.

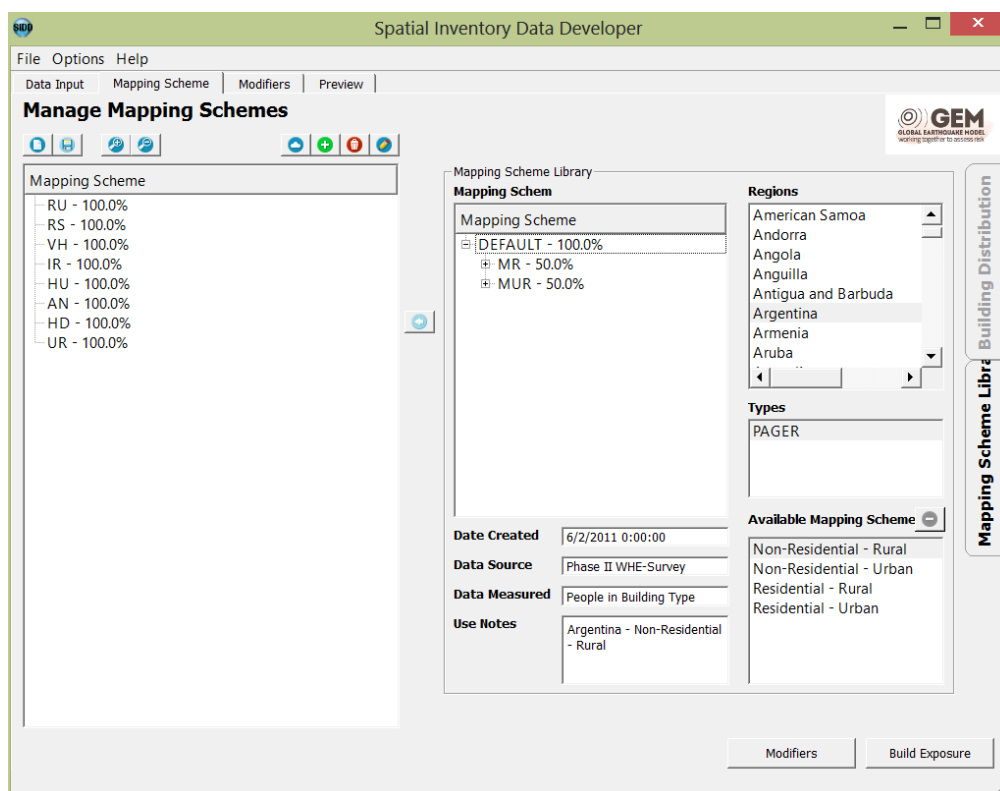



Figure 4.14 Managing Mapping Scheme, the Mapping Scheme Library

The user must now identify which available mapping scheme is the most representative of the homogenous class that needs to be mapped. In most cases, this will not be a one-to-one match, therefore adjustments will need to be made within the attributes distribution.

Select the appropriate mapping scheme and a distribution of attributes will appear in the adjacent window. In addition, information regarding the date created, data source, data measured and other notes is available. Select the land class you wish the PAGER data to be mapped to and click the “Append Selected Mapping

Scheme from Library to Zone” button . The PAGER mapping scheme should now be transferred to the specified zone. It is important to note that the PAGER estimates are based on the percentage of people dwelling in a given type of building, rather than the percentage of buildings.

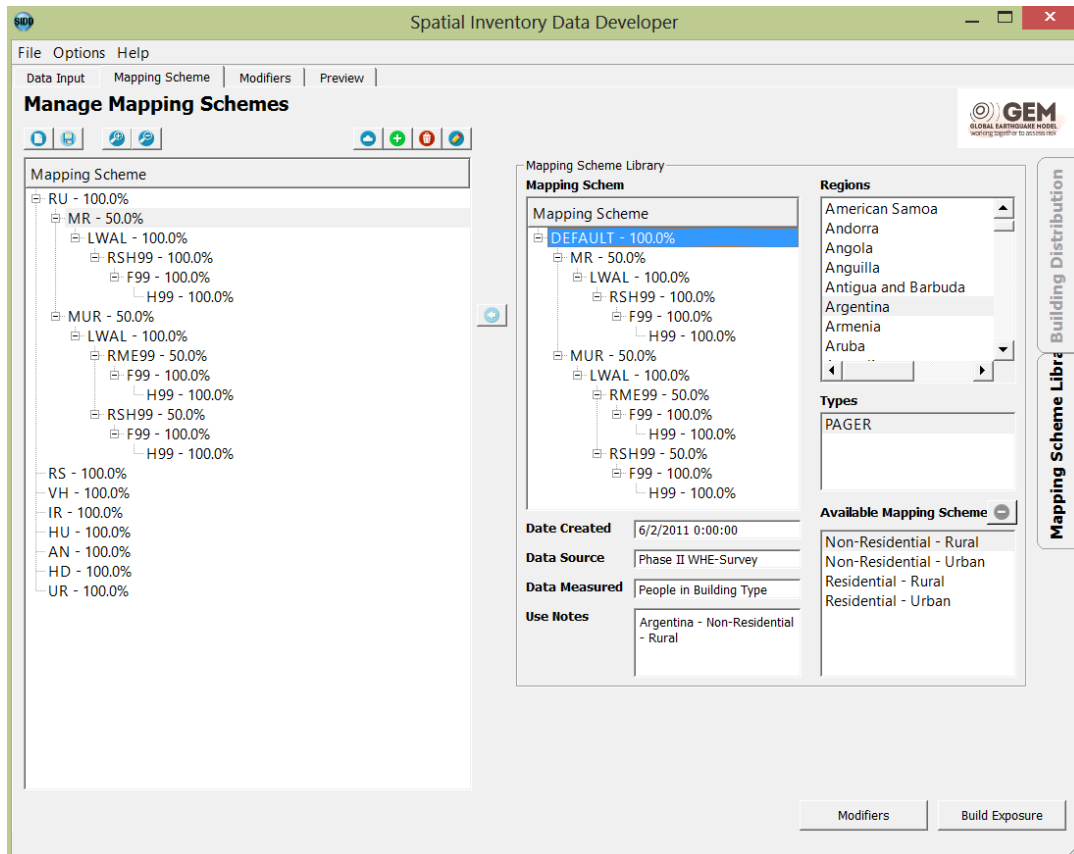


Figure 4.15 Managing Mapping Schemes, Assigning the Saved Mapping Scheme to a Branch

The process should be repeated for any other homogenous zones that need to be mapped and adjusted (if required) following the instructions described in the “Editing the Distribution of Attributes within a Given Homogenous Class” section.

It is important to note that preloaded mapping schemes can be mapped to any level of the distribution tree. This can be useful when trying to create mapping schemes for a single attribute (i.e. occupancies). The user must first load the empty zones and create a distribution of occupancies. Find the appropriate mapping scheme in the library and transfer the scheme to the occupancy level. The mapping scheme has now been applied to the occupancy, as opposed to the homogenous zone.

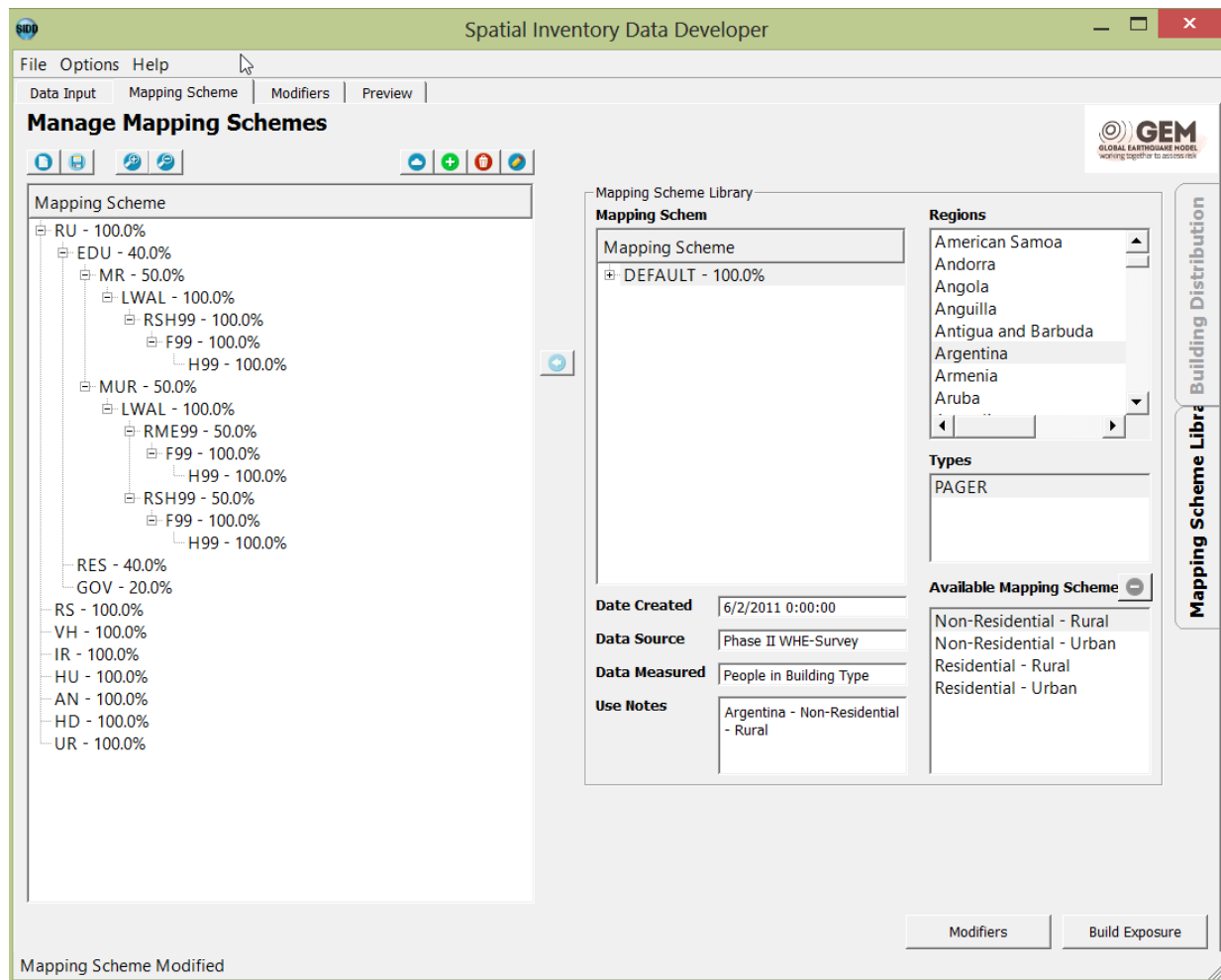


Figure 4.16 Managing Mapping Scheme, Final Distributions

The direct observation tools are designed to be flexible and allow users to collect as much data as possible without requiring certain fields to be filled in. This data can then be imported into SIDD, which can result in blank data entries, as displayed in Figure 4.16. There are three options for cleaning the data: 1) reclassify the value through editing, 2) delete the branch and all the percentages will adjust automatically, or 3) clean the data externally and re-import the more accurate database.

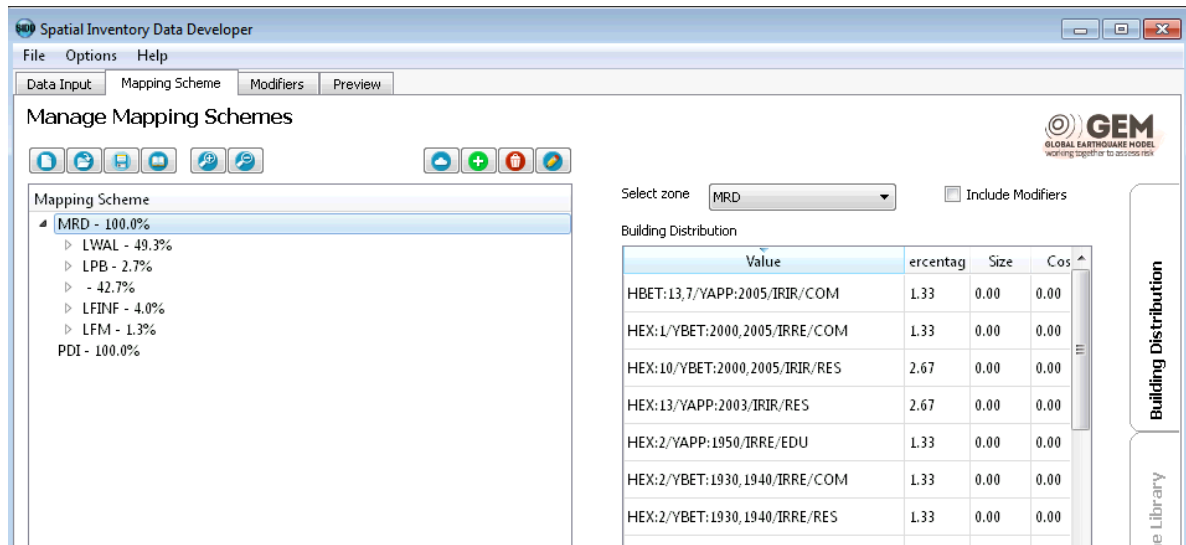


Figure 4.17 Faulty Data in the Mapping Scheme Pane

4.9 Modifiers

Modifiers provide a detailed description of the various attributes, however are not included in the mapping scheme tree. In order to view a list of all modifiers, the user will need to switch to the “Modifiers” tab (See Figure 4.18).

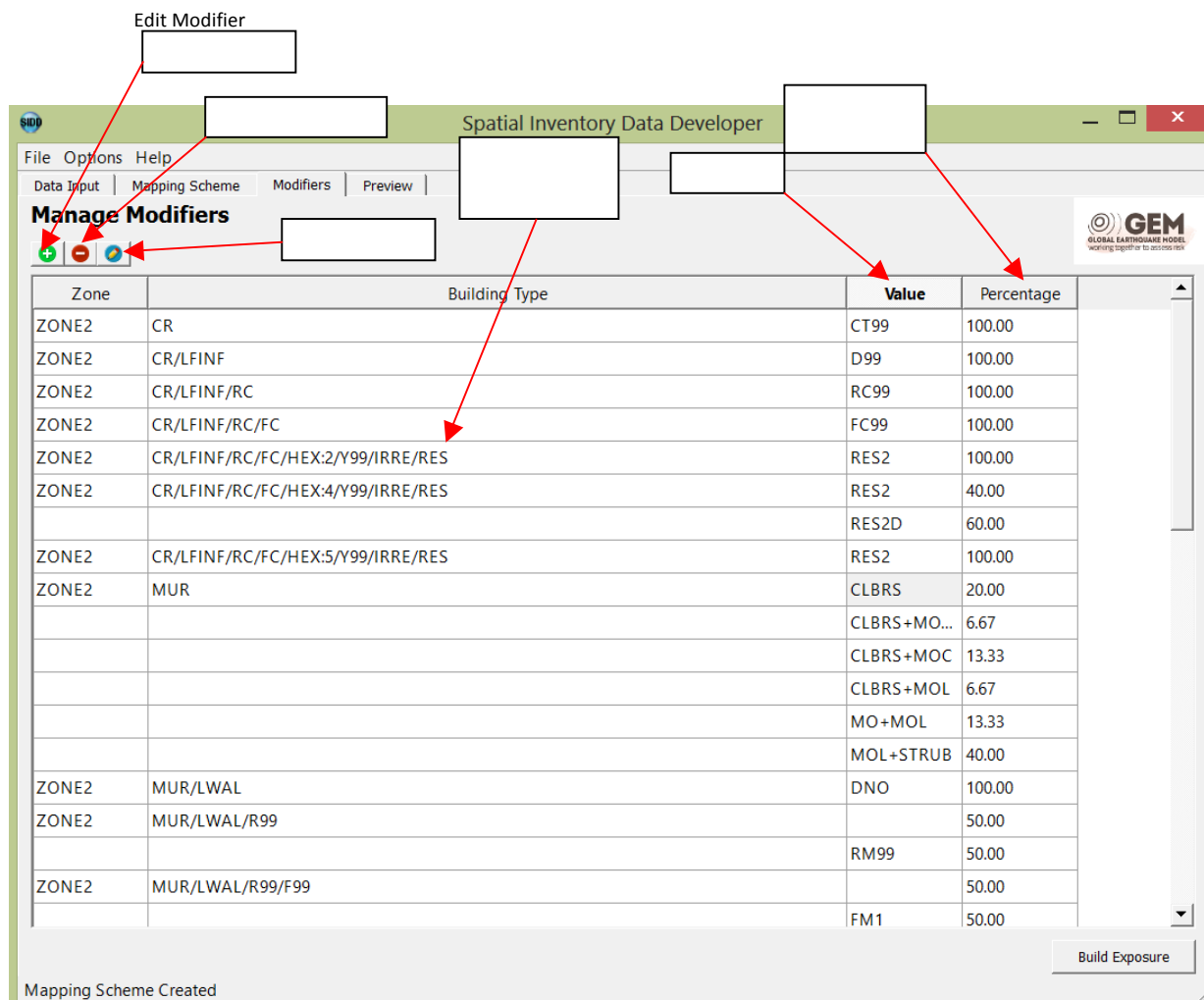




Figure 4.18 Managing Mapping Scheme, Final Distributions

To add a modifier, first click the “Add Modifier” button . This will open an “Edit Building Distribution Modifier” dialog (Figure 4.19), with the current mapping scheme shown in the left window. Click the attribute you wish to add a modifier to (it is important to note that all attributes except “height” and “date of construction” can have modifiers added). Once the attribute has been selected, click the  next to “Modifier Detail” (Figure 4.20). A blank value cell and 0.00 percentage will appear in the table. Double clicking the value cell will bring up an “Edit Modifier Value” window.

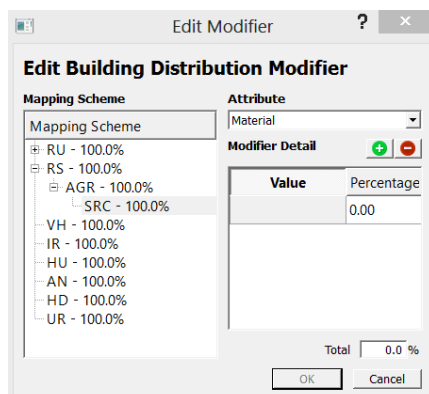



Figure 4.19 Editing building distribution modifier

The “Edit Modifier Value” window has a drop down menu that lists all applicable modifiers for the specified attribute. Once the appropriate modifier is selected, a description will appear adjacent to the drop down menu and a “Modifier Value” will be created. Users can specify multiple modifiers by pressing the “Add Value” button  in the “Edit Modifier Value” window. Once the task is complete, click “OK” and the modifier will have been added to the previous window. Enter the percentage for the specified modifier and click “OK”.

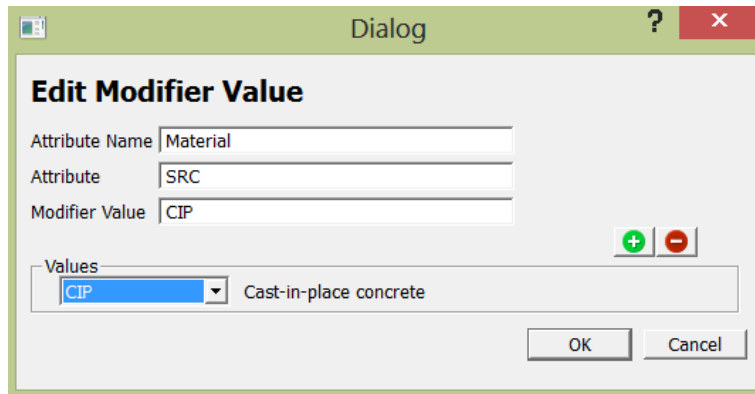



Figure 4.20 Edit Modifier Value

4.10 Saving Mapping Scheme

Custom mapping schemes can be saved and loaded at another time. Clicking the “Save Mapping Scheme” button  in the mapping scheme tab will open a “Save Mapping Scheme” dialog (Figure 4.10.1). Users should specify the country in the “Region” drop down menu (an option of “Undefined-Region” is also available). Specify a name in the “Name” box and add a data source, measure and any notes if required and click “Save”.

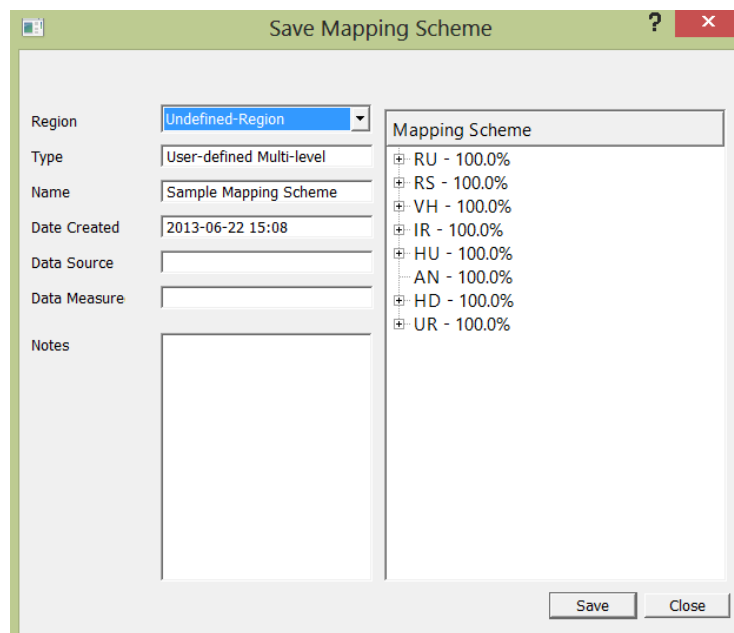


Figure 4.21 Saving the mapping schemes

4.11 Modifying Average Dwelling Size and Replacement Cost

To modify the average dwelling size or replacement cost, make sure the Building Distribution tab is selected on the right hand panel. SIDD displays the size and cost in a table for each zone type, and for each combination of factors. To edit the size and cost, double click on any of the size and cost values and a separate dialog comes up to allow editing, as shown in Figure 4.22. The user can modify these values directly, as well as navigate to any other size and replacement cost assumptions in the mapping scheme pane on the left. If users modify the values with a high-level branch selected, the price and size assumptions are distributed to all the corresponding structural types.

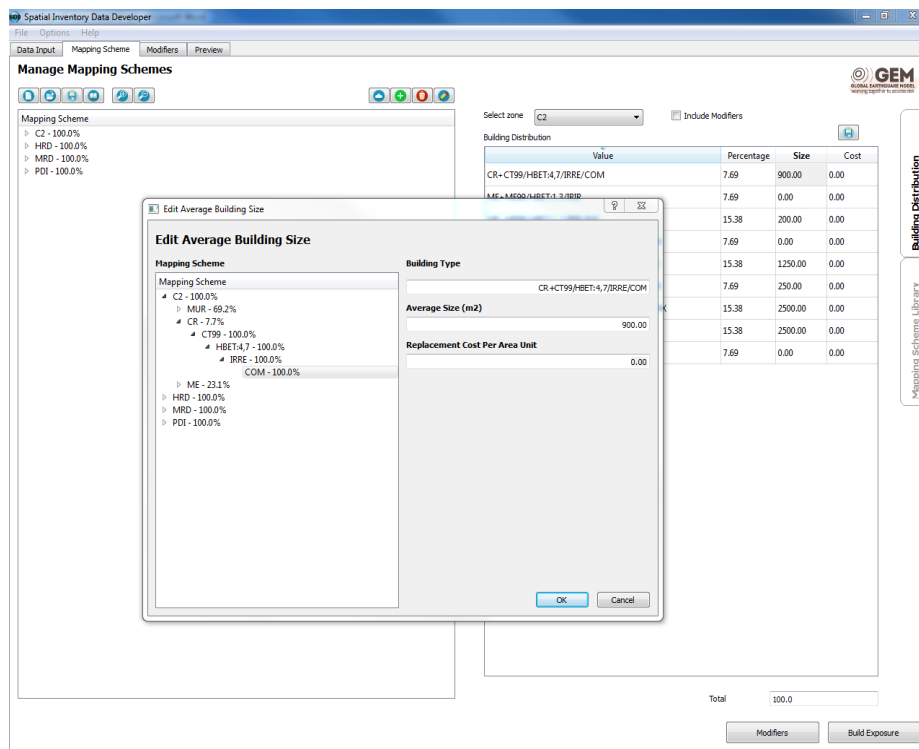


Figure 4.22 Edit Average Building Size and Cost

If the Include Modifiers checkbox is selected, all building percentage distributions are calculated for user review. Editing the size and cost feature is not available when reviewing the impact of the modifiers.

5 GIS Mapping

SIDD incorporates Quantum GIS as a platform, and thus is able to offer powerful mapping tools without licensing considerations. The preview tab offers the user the ability to map and interrogate both the base data and the results.

5.1 GIS Use

This chapter introduces some of the critical mapping features, as well as describes the data quality and export features. Users who are generally familiar with GIS programs should be able to navigate the additional GIS features given this brief introduction. Figure 5.1 presents the preview tab as it is first encountered after an analysis. This chapter will examine some of the visualization options to improve the appearance.

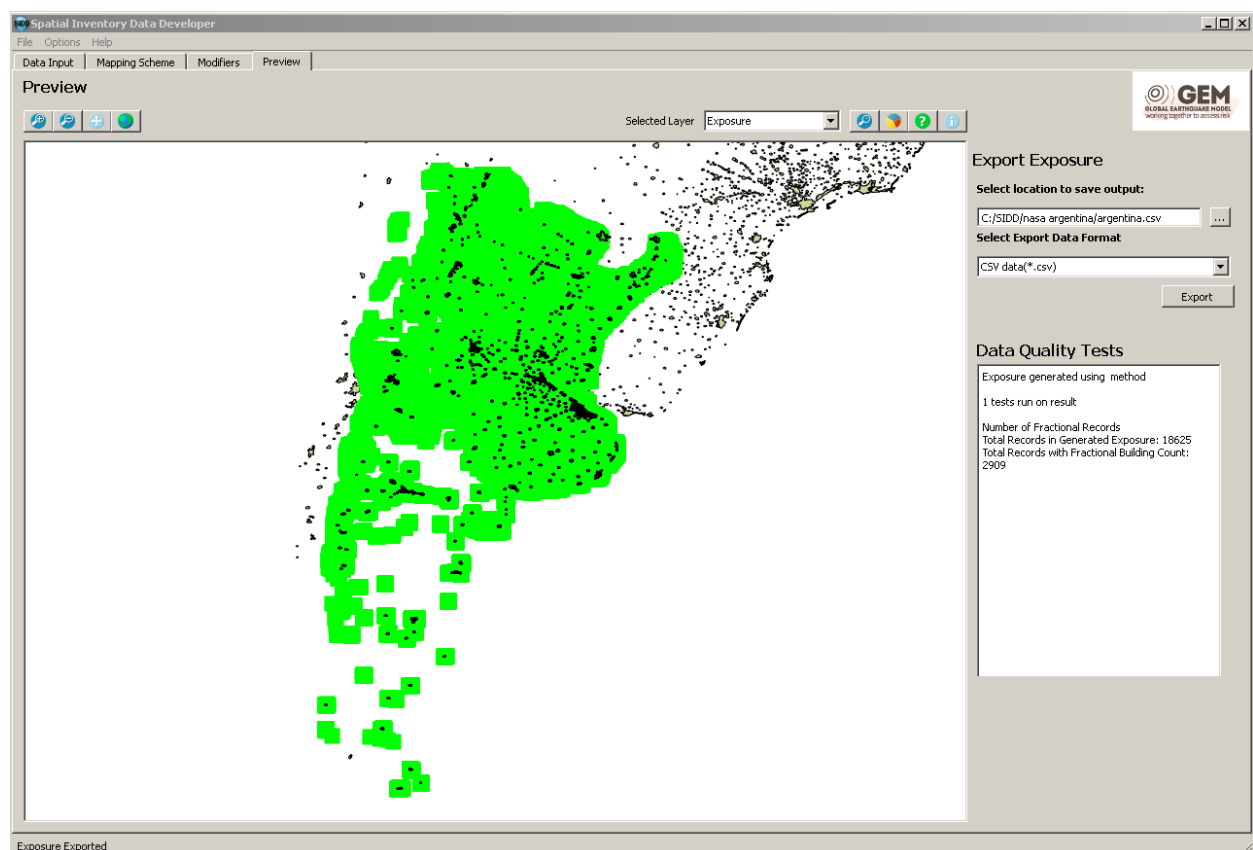


Figure 5.1 Sample output in the preview tab.

A critical feature of the preview tab is the export exposure feature. Users are reminded that the results are not stored within the SIDD project file, and thus it is recommended that user export all results before exiting the program. When a user exports a .csv file, a text file is constructed that includes an ID, longitude, latitude, the GEM taxonomy string, the homogenous zone, the number of buildings, area, and replacement cost. Table

5.1 provides a sample from Argentina. If a shapefile is exported, it is accompanied with the same attribute data, but every row links to a geographic object, either a zone or a point, depending on the aggregation method selected on the import tab. It is highly recommended that users examine the data closely in a variety of environments before under taking risk analysis.

Table 5.1 Sample CSV export results

GID	LON	LAT	TAXONOMY	ZONE	NUM_BLDGS	AREA	REPCOST
-568334233	-72.2725	-50.3408333275	MR+MUN99+R5	HD	20	0	0
-568334233	-72.2725	-50.3408333275	MUR+CLBR5	HD	0.6	0	0
-568334233	-72.2725	-50.3408333275	CR+CIP/LFM+DUC/HBET:1,3	HD	1.9712	0	0
-568334233	-72.2725	-50.3408333275	CR+CIP/LFINF+DNO/HBET:1,3	HD	9.9968	0	0
-568334233	-72.2725	-50.3408333275	S	HD	1.2	0	0
-568334233	-72.2725	-50.3408333275	CR+CIP/LFINF+DNO/HBET:4,7	HD	4.0832	0	0
-568334233	-72.2725	-50.3408333275	MUR+MON+STRUB	HD	0.6	0	0
-568334233	-72.2725	-50.3408333275	CR+CIP/LFM+DUC/HBET:8,999	HD	0.3872	0	0
-568334233	-72.2725	-50.3408333275	CR+CIP/LFM+DUC/HBET:4,7	HD	1.1616	0	0
-560862238	-71.325	-42.9166666609	MR+MUN99+R5	HD	43.5	0	0
-560862238	-71.325	-42.9166666609	MUR+CLBR5	HD	1.305	0	0
-560862238	-71.325	-42.9166666609	CR+CIP/LFM+DUC/HBET:1,3	HD	4.28736	0	0
-560862238	-71.325	-42.9166666609	CR+CIP/LFINF+DNO/HBET:1,3	HD	21.74304	0	0
-560862238	-71.325	-42.9166666609	S	HD	2.61	0	0
-560862238	-71.325	-42.9166666609	CR+CIP/LFINF+DNO/HBET:4,7	HD	8.88096	0	0
-560862238	-71.325	-42.9166666609	MUR+MON+STRUB	HD	1.305	0	0
-560862238	-71.325	-42.9166666609	CR+CIP/LFM+DUC/HBET:8,999	HD	0.84216	0	0
-560862238	-71.325	-42.9166666609	CR+CIP/LFM+DUC/HBET:4,7	HD	2.52648	0	0
-560730954	-71.3083333333	-41.1499999942	MR+MUN99+R5	HD	132.5	0	0
-560730954	-71.3083333333	-41.1499999942	MUR+CLBR5	HD	3.975	0	0
-560730954	-71.3083333333	-41.1499999942	CR+CIP/LFM+DUC/HBET:1,3	HD	13.0592	0	0
-560730954	-71.3083333333	-41.1499999942	CR+CIP/LFINF+DNO/HBET:1,3	HD	66.2288	0	0
-560730954	-71.3083333333	-41.1499999942	S	HD	7.95	0	0
-560730954	-71.3083333333	-41.1499999942	CR+CIP/LFINF+DNO/HBET:4,7	HD	27.0512	0	0
-560730954	-71.3083333333	-41.1499999942	MUR+MON+STRUB	HD	3.975	0	0
-560730954	-71.3083333333	-41.1499999942	CR+CIP/LFM+DUC/HBET:8,999	HD	2.5652	0	0
-560730954	-71.3083333333	-41.1499999942	CR+CIP/LFM+DUC/HBET:4,7	HD	7.6956	0	0
-560862025	-71.3208779113	-41.1367823317	MR+MUN99+R5	HD	16.5	0	0
-560862025	-71.3208779113	-41.1367823317	MUR+CLBR5	HD	0.495	0	0
-560862025	-71.3208779113	-41.1367823317	CR+CIP/LFM+DUC/HBET:1,3	HD	1.62624	0	0
-560862025	-71.3208779113	-41.1367823317	CR+CIP/LFINF+DNO/HBET:1,3	HD	8.24736	0	0

SIDD runs 2 data quality test and presents the results back to the user for consideration. These tests give the user an indication of issues that might arise from zones that are too small or mapping schemes that are too complex, as well as the processing options chosen under the options menu. The first test compares the count of buildings input to the count of buildings output. These will always match except under the second processing option where rounding has occurred (Building Distribution Fraction Rounded). This comparison will give the user an indication of the net effect of the rounding. The second test indicates to the user the number of “fractional records” or the number of exposures where there has been less than a single building assigned for a given zone/structure type combination. The second problem can be alleviated by reducing the number of structure type classifications in the mapping schemes, particularly the smaller values.

5.2 Toolbar

Figure 5.2 displays the preview toolbar. The toolbar is divided into two sections. The tools on the left are for navigation, and the tools on the right set options for the layer selected from the “selected layer” drop-down menu. Table 5.2 presents an introduction to the main features of the toolbar. Some visualization features are discussed in more detail later in the chapter. The user can obtain a description of each tool through tool-tips by hovering over each icon.

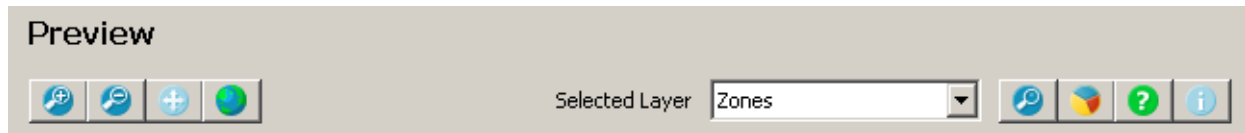


Figure 5.2 The preview toolbar

Table 5.2 Feature of the preview toolbar





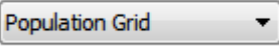




Icon/menu	Description	Comments
	Zoom in	
	Zoom out	
	Pan	
	Zoom to full extent	
Selected Layer 	Select layer	Determines layer for subsequent icons
	Zoom to selected layer	
	Layer rendering options...	Additional visualization options discussed in sections below
	Zoom to feature...	Additional options discussed below, see
	Query feature in layer	See Figure 5.4

Figure 5.3 displays the zoom to feature option invoked by the question mark icon. Users can enter an attribute field and value, and the map will zoom to all features with that description. This feature is valuable for checking specific homogenous zones or taxonomy strings.

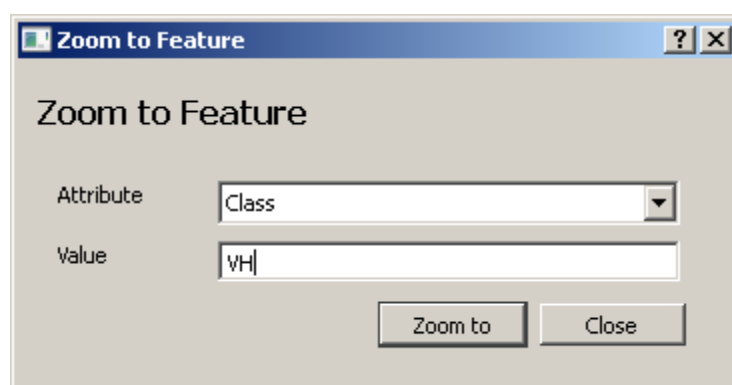


Figure 5.3 Zoom to Feature

Through the Quantum GIS tool SIDD integrates some powerful mapping tools, and users are advised to use them to verify the inputs and outputs of SIDD very carefully before proceeding with a loss assessment. Figures 5.4 and 5.5 below present the results of the feature query tool. The ideal way to use this tool is in conjunction with the thematic options. With thematics on key fields, users can query inputs and outputs to look for anomalies. The following sections examine these mapping options.

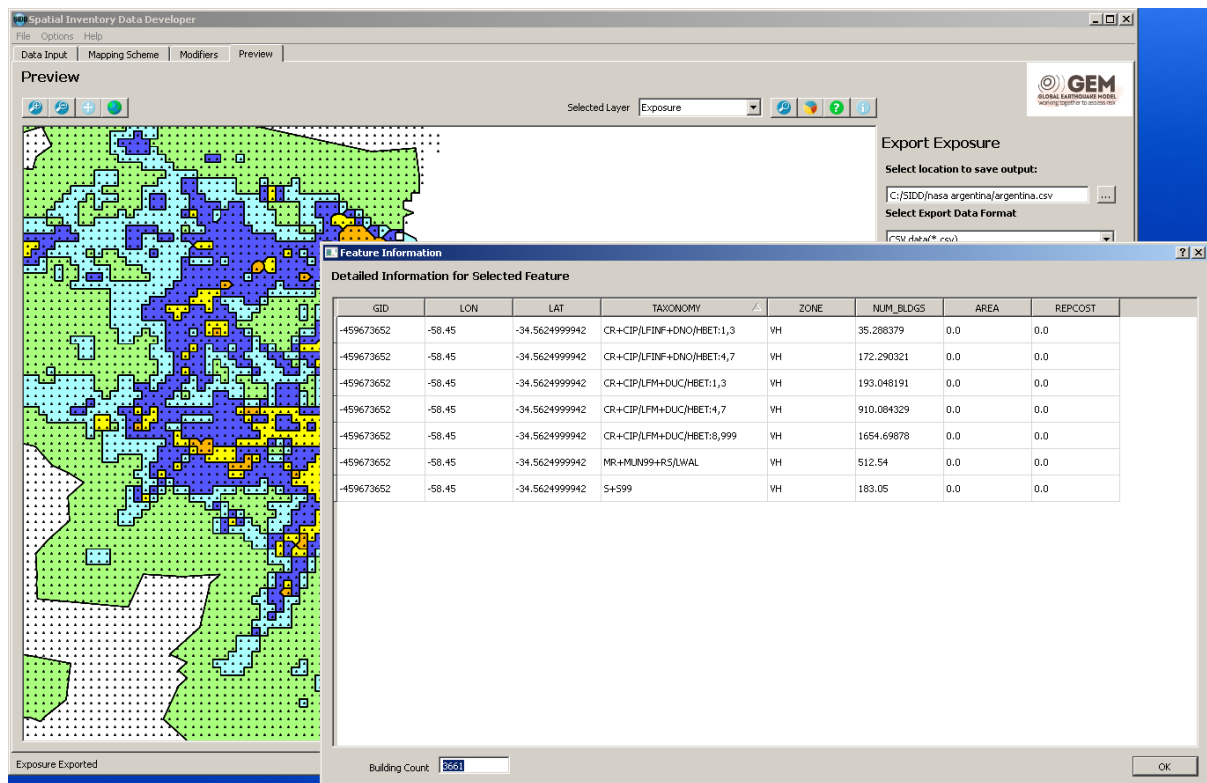


Figure 5.4 Query feature for results

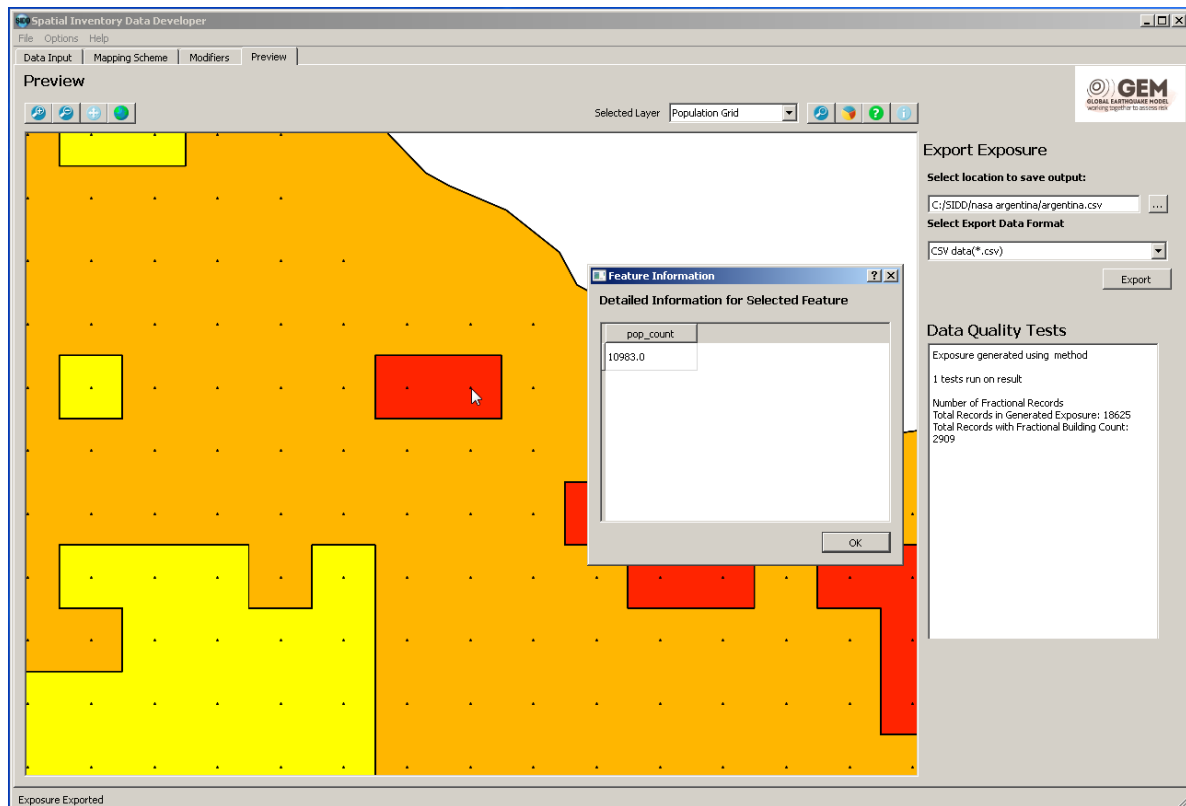



Figure 5.5 Query feature for input data

5.3 Thematic mapping

This section describes the process of making thematics using the Quantum GIS tools. Figure 5.6 presents the Layer Rendering Settings that are displayed after the user selects the layer rendering options ().

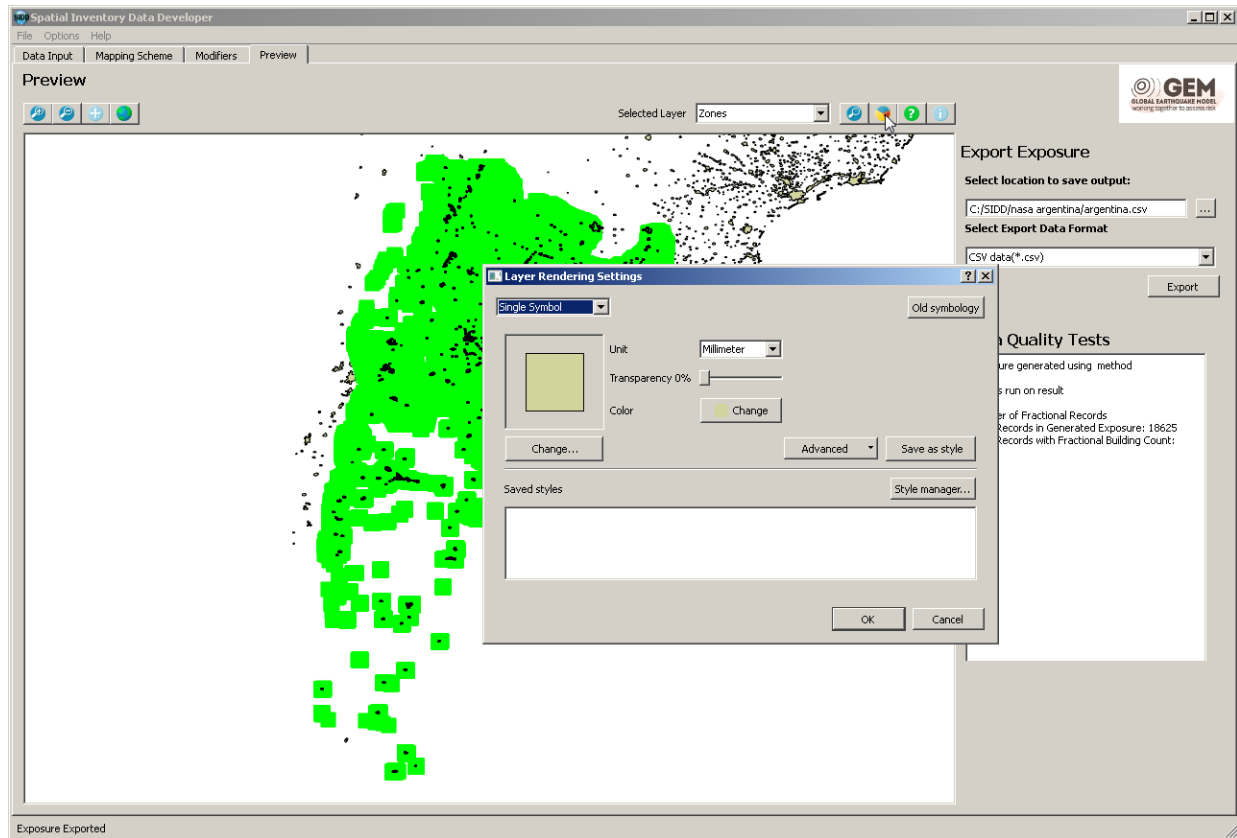


Figure 5.6 Layer Rendering Settings

The first step to setting a thematic is to change the drop down option from single symbol to either categorized or graduated. Most of the features apply to more than one feature, but users may opt to explore the options in more detail using the Quantum GIS manual available online. In this example, categorized is selected.

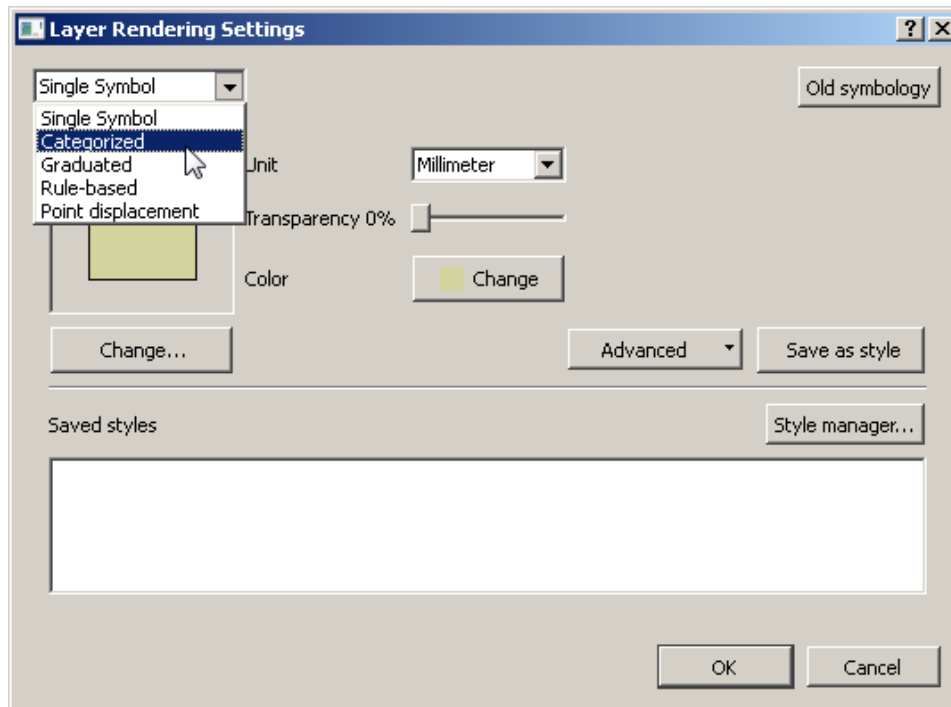


Figure 5.7 Layer Rendering Setting, Selecting the thematic option

Once a thematic option is selected, a colour ramp option appears. The user can select a saved ramp, if available, or make a new one. The user must select the type of ramp (Figure 5.8). Gradient (Figure 5.9) provides a smooth dissolve between two colours. Random provides one colour per class, with a specific hue or scheme, and ColorBrewer (Figure 5.10) provides a compromise between these two options that might have a progression between colours.

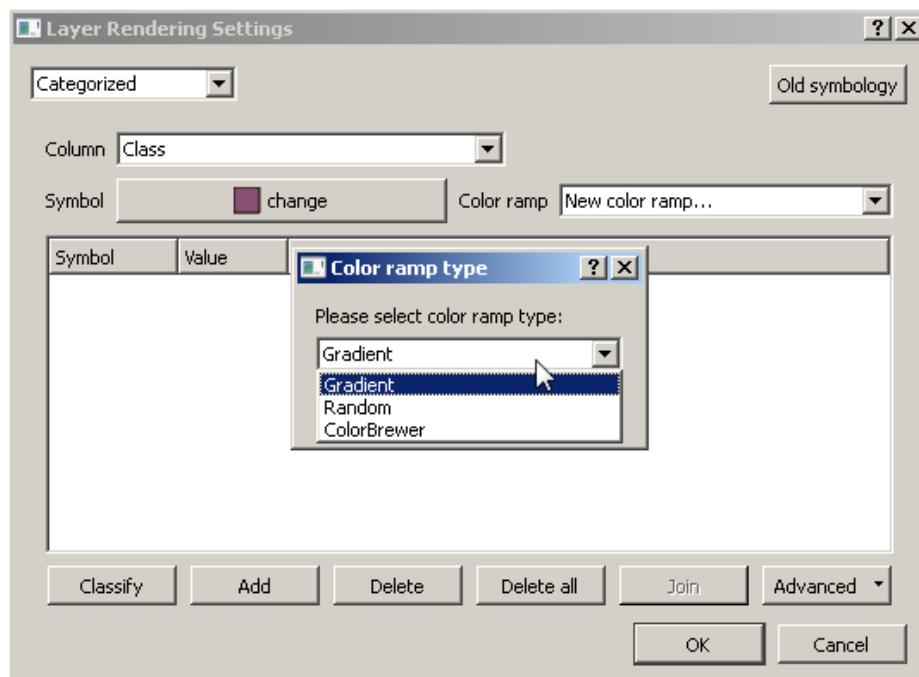


Figure 5.8 Layer Rendering Setting, making a colour ramp.

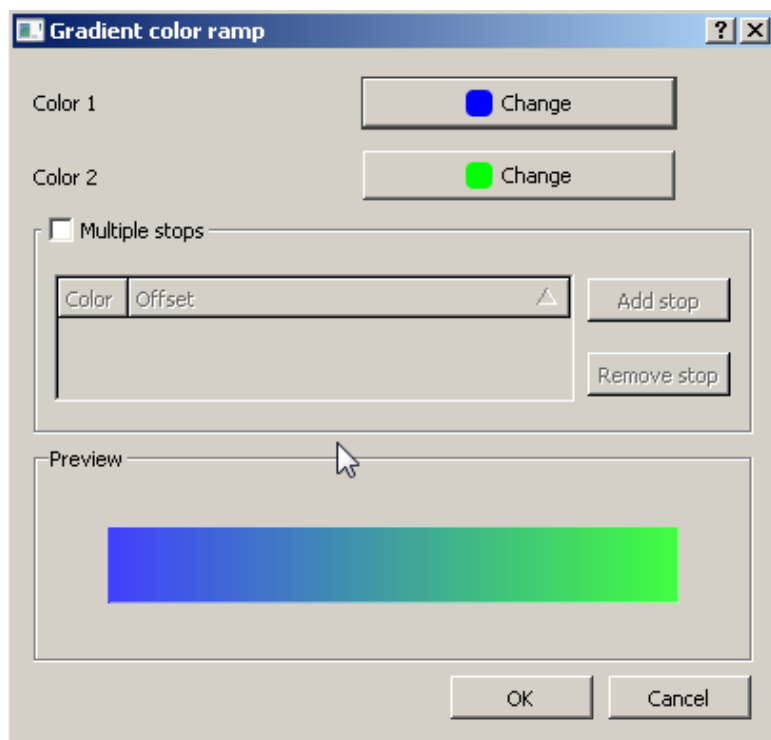


Figure 5.9 Gradient colour Map

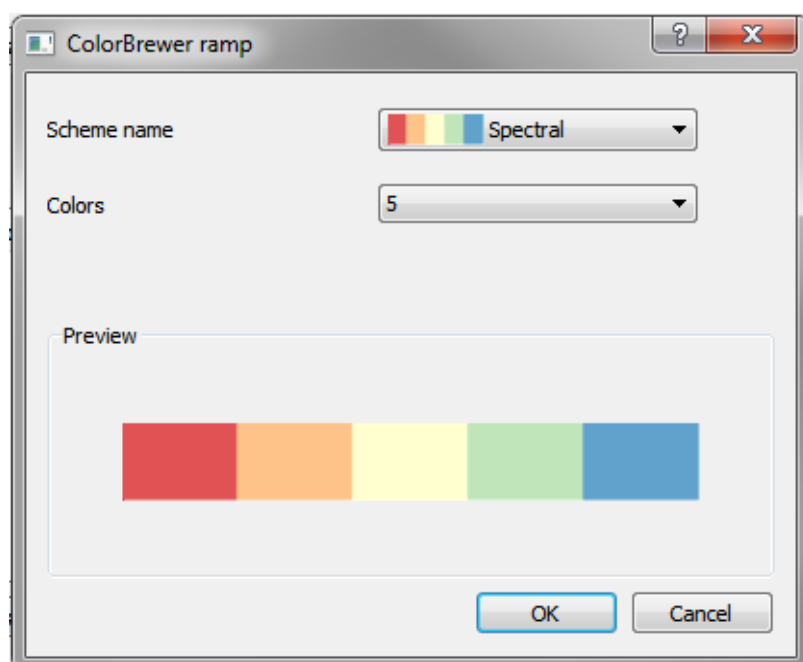


Figure 5.10 ColorBrewer ramp

In each of these dialogs, the colour options are changed by double clicking on a colour, and the colour is presented.

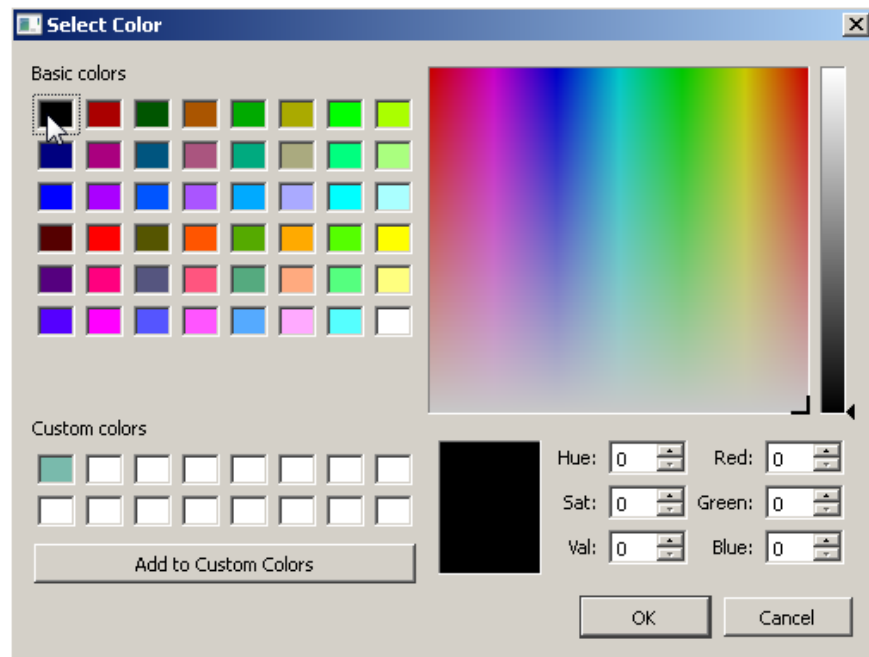


Figure 5.11 Select Colour

Once the user has selected the type of ramp, they select “OK” until they are back to the Layer Rendering Settings menu. They create the thematic by selecting the classify button, bring up the values of each value in the field. Users are invited to explore the advanced menu for additional settings such as symbol rotation.

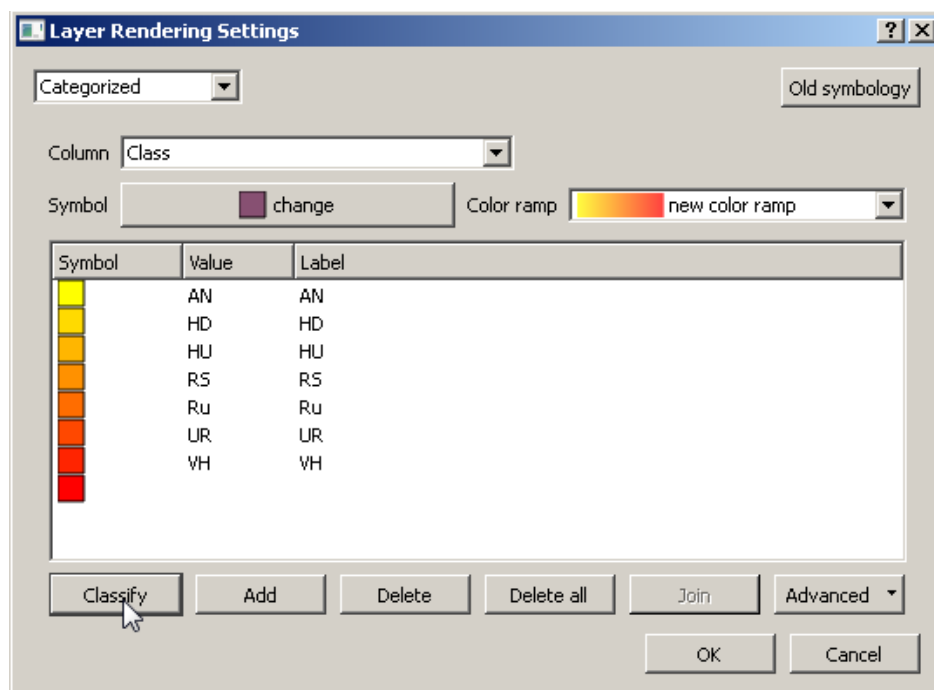


Figure 5.12 Layer Rendering Setting, Adding the class values

The colour of individual colours may not lend themselves to a particular ramp, as illustrated in the scheme in Figure 5.12. In general, ramps may be more applicable to numeric values, such as number of buildings. To change the colour of each symbol, the user double clicks on the colour in the symbol palette, and the dialog in Figure 5.13 is presented.

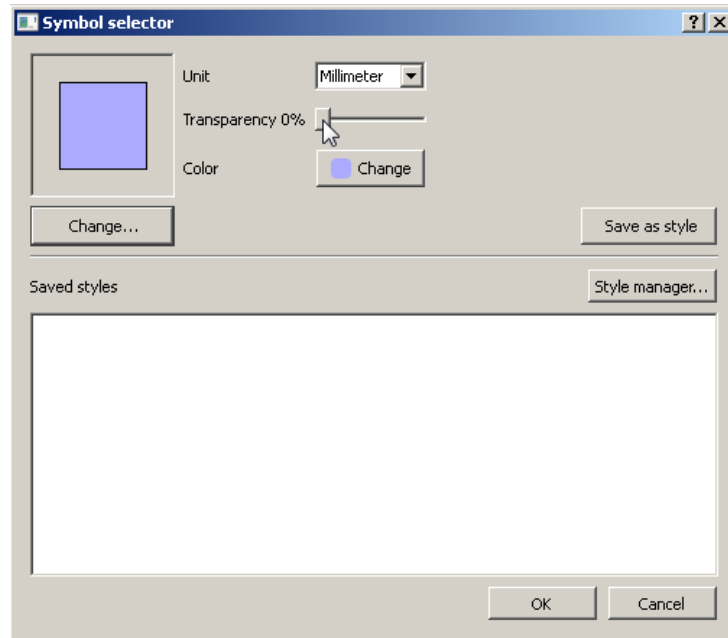


Figure 5.13 Symbol Selector

The symbol selector allows the user to adjust transparency, as well as change the symbol colour or other attributes, such as the border colour (Figure 5.14). The style manager allows users to save the schemes once developed and retrieve them from the saved styles menu (Figure 5.15).

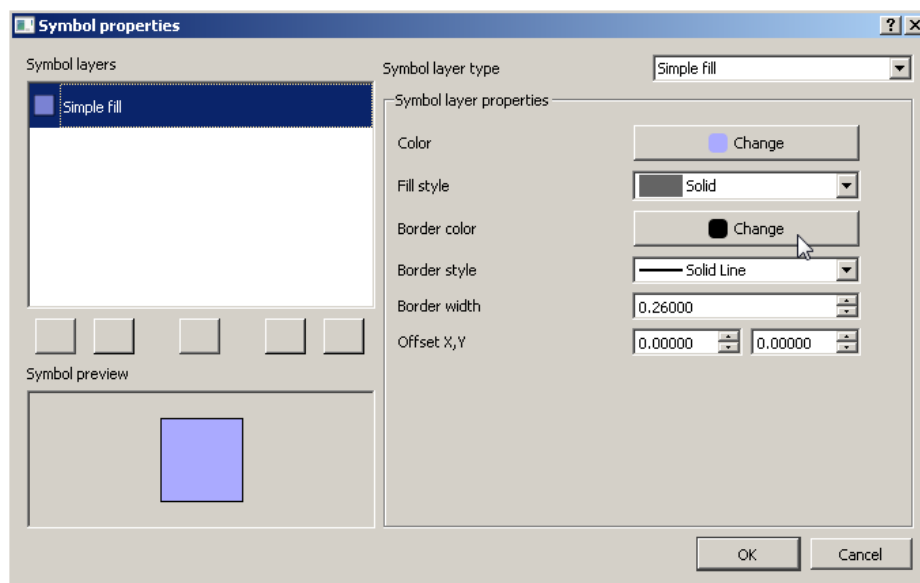


Figure 5.14 Symbol Properties

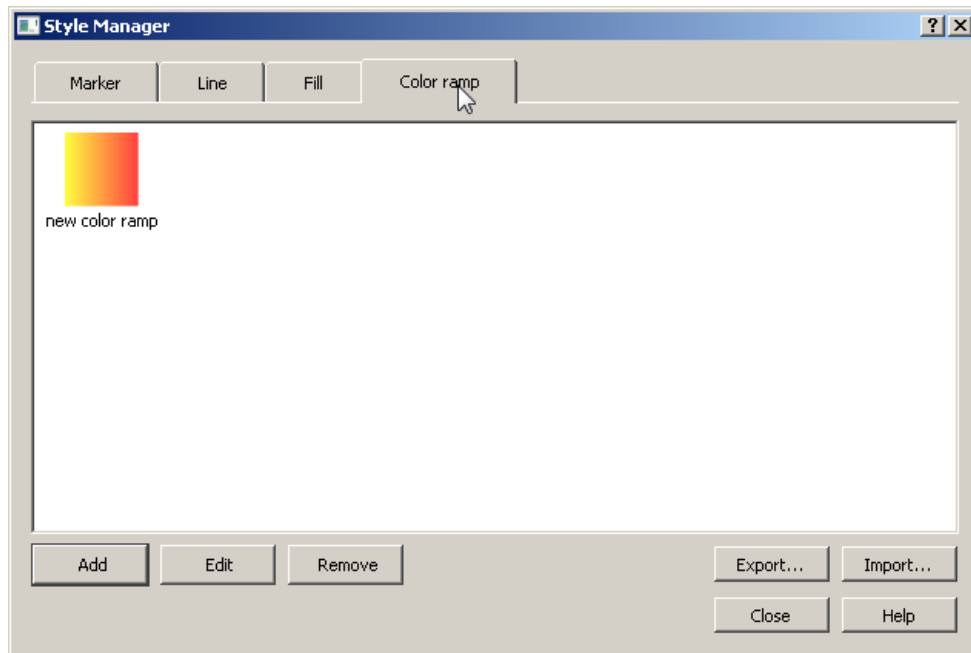


Figure 5.15 Style Manger

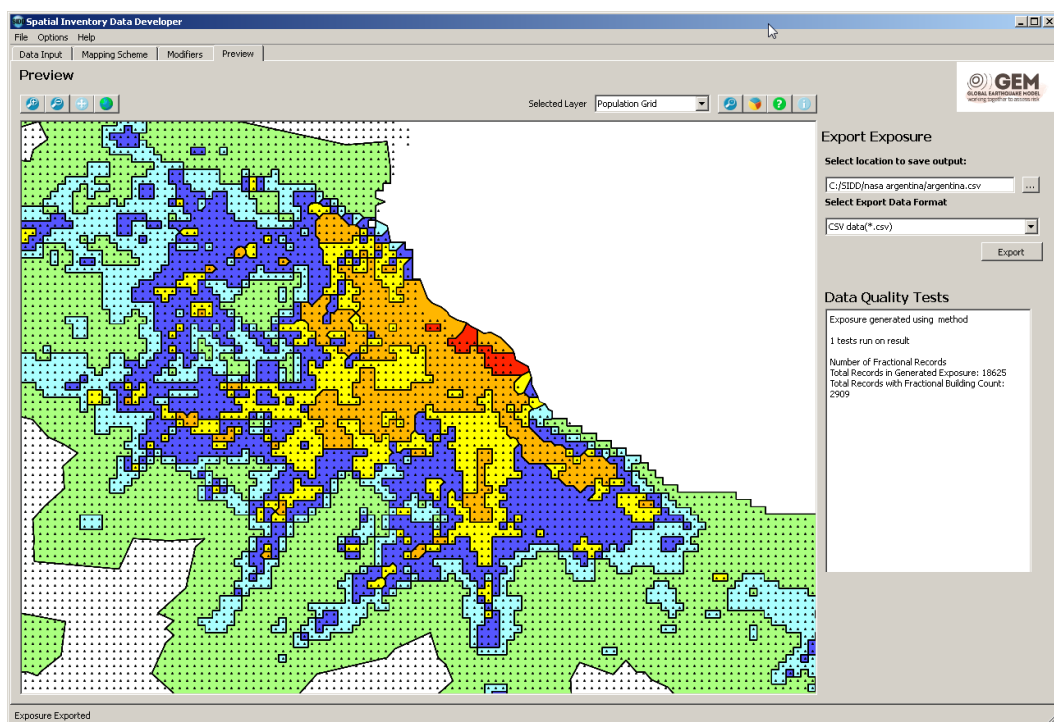



Figure 5.16 Adjust map on the preview tab

Using the tools described above, Figure 5.16 presents a more user-friendly map, allowing users to see both the land use patterns for the area of interest and the points representing the 30 arc-second grid.

5.4 Adjusting Point Symbols

To adjust a point file, first select the point file in the selected drop down layer, then select layer rendering options (). Figure 5.17 displays the resulting dialog, allowing the user to change the color, size, and save the style to the saved styles menu. To change the symbol itself, select the “change..” option, and select from the symbols in Figure 5.18.

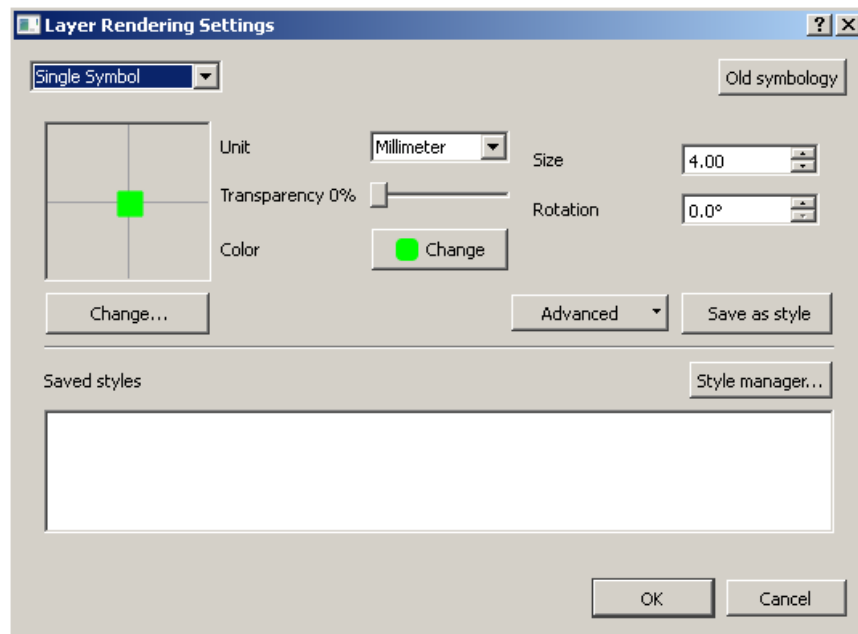


Figure 5.17 Layer rendering setting, points

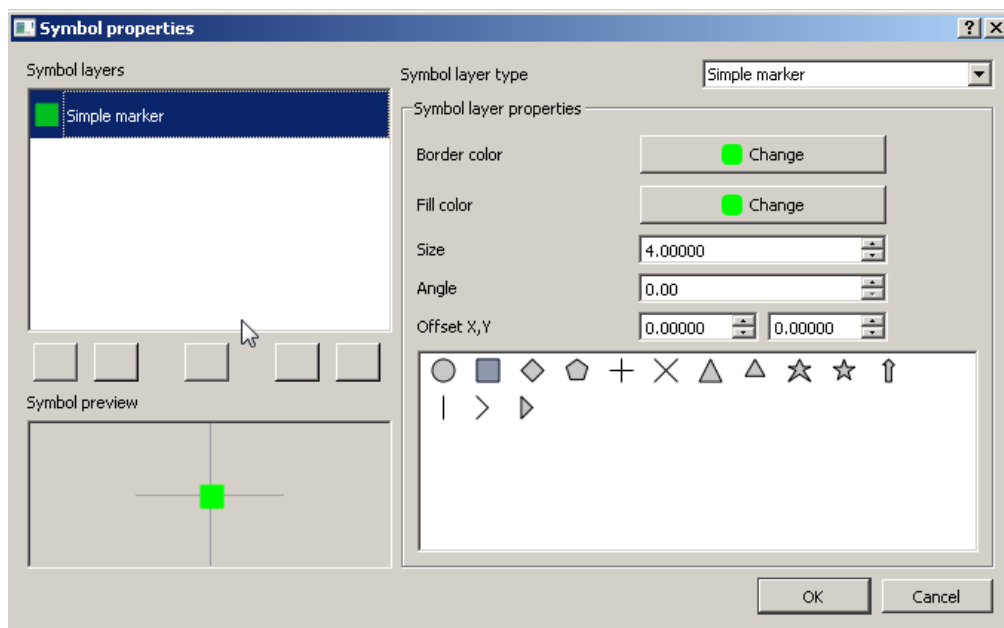


Figure 5.18 Layer rendering setting, change point symbol

REFERENCES

Document References

Brzev, S., Scawthorn, C., Charleson, A.W. and Jaiswal, K. (2012), GEM Basic Building Taxonomy Beta Version 1.0, March 2012, <http://www.nexus.globalquakemodel.org/gem-building-taxonomy/posts/gem-basic-building-taxonomy-beta-v1.0>

Hu, Z., Huyck, C., Eguchi, M. and Bevington, J. (2014), User guide: Tool for spatial inventory data development, GEM Technical Report 2014-05 V1.0.0, GEM Foundation, Pavia, Italy.

ImageCat (2011), Indonesia Building Exposure Development Report, Internal report to Australia-Indonesia Facility for Disaster Reduction (AIFDR) and the Indonesian Government, ImageCat, Inc., USA. Available from URL: <http://www.nexus.globalquakemodel.org/gem-idct/posts/building-exposure-models-in-indonesia>

Jaiswal K.S. and D.J. Wald (2008), Creating a global building inventory for earthquake loss assessment and risk management, U.S. Geological Survey Open-File Report, OF 2008-1160, 103 pp., USGS, USA.

THE GLOBAL EARTHQUAKE MODEL

The mission of the Global Earthquake Model (GEM) collaborative effort is to increase earthquake resilience worldwide.

To deliver on its mission and increase public understanding and awareness of seismic risk, the GEM Foundation, a non-profit public-private partnership, drives the GEM effort by involving and engaging with a very diverse community to:

- Share data, models, and knowledge through the OpenQuake platform
- Apply GEM tools and software to inform decision-making for risk mitigation and management
- Expand the science and understanding of earthquakes

GEM Foundation

Via Ferrata 1
27100 Pavia, Italy
Phone: +39 0382 5169865
Fax: +39 0382 529131
info@globalquakemodel.org
www.globalquakemodel.org

Copyright © 2014 GEM Foundation,
Hu, Z., C. Huyck, Eguchi, M., Bevington, J.



Creative Commons Attribution 3.0 Unported License Except where otherwise noted, this work is licensed under a **Creative Commons Attribution 3.0 Unported License**

JANUARY 2014

 **GEM**
GLOBAL EARTHQUAKE MODEL
working together to assess risk