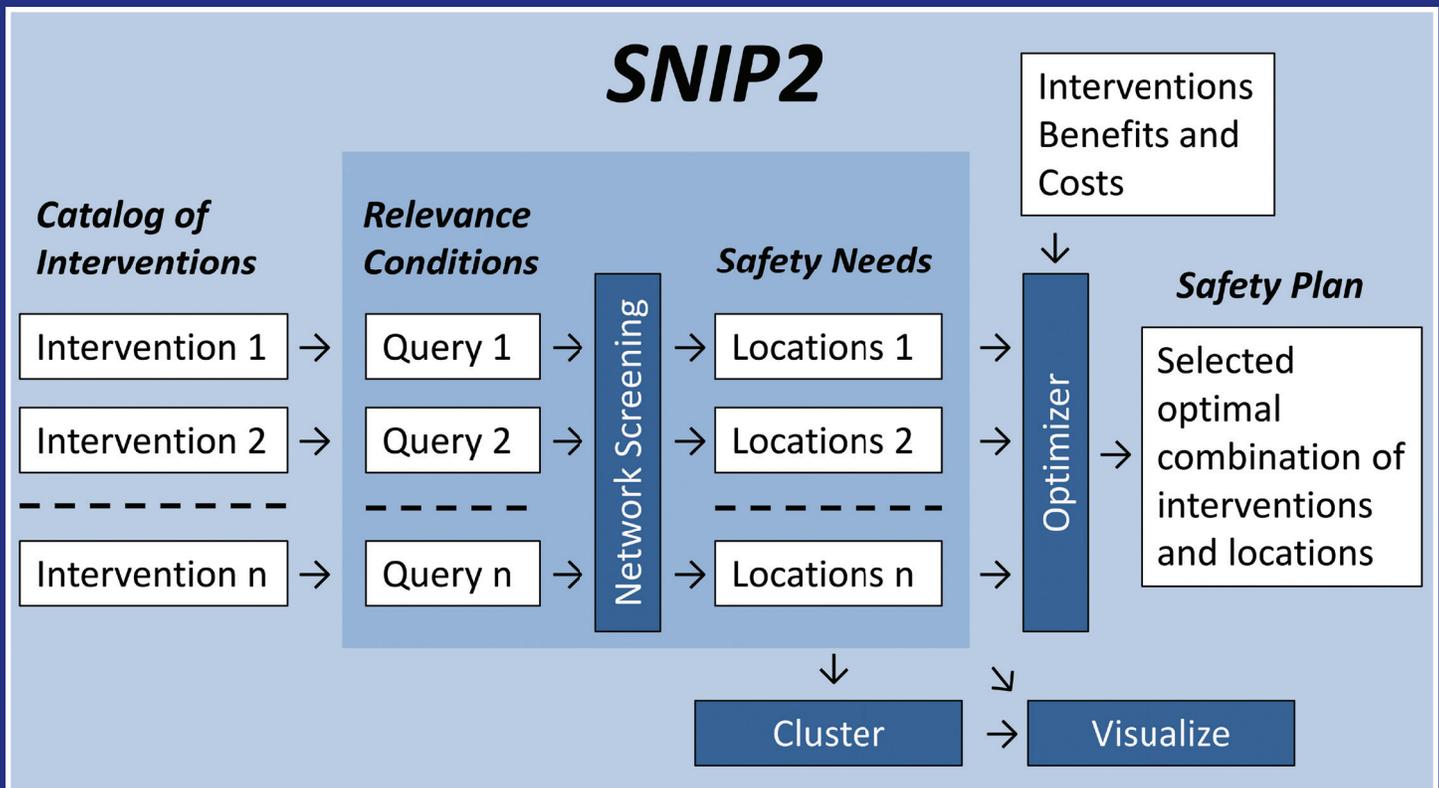


JOINT TRANSPORTATION RESEARCH PROGRAM

INDIANA DEPARTMENT OF TRANSPORTATION
AND PURDUE UNIVERSITY



A Systematic Approach to Identifying Traffic Safety Needs and Intervention Programs for Indiana *Volume II—SNIP2 User Manual*



Andrew P. Tarko, Mingyang Li, Mario Romero, Jose Thomaz

RECOMMENDED CITATION

Tarko, A. P., Li, M., Romero, M., & Thomaz, J. (2014). *A systematic approach to identifying traffic safety needs and intervention programs for Indiana: Volume II—SNIP2 user manual* (Joint Transportation Research Program Publication No. FHWA/IN/JTRP-2014/04). West Lafayette, IN: Purdue University. <http://dx.doi.org/10.5703/1288284315498>

AUTHORS

Andrew P. Tarko, PhD

Professor of Civil Engineering
Lyles School of Civil Engineering
Purdue University
(765) 494-5027
tarko@purdue.edu
Corresponding Author

Mingyang Li

Graduate Research Assistant
Lyles School of Civil Engineering
Purdue University

Mario Romero

Graduate Research Assistant
Lyles School of Civil Engineering
Purdue University

Jose Thomaz

Graduate Research Assistant
Lyles School of Civil Engineering
Purdue University

JOINT TRANSPORTATION RESEARCH PROGRAM

The Joint Transportation Research Program serves as a vehicle for INDOT collaboration with higher education institutions and industry in Indiana to facilitate innovation that results in continuous improvement in the planning, design, construction, operation, management and economic efficiency of the Indiana transportation infrastructure. https://engineering.purdue.edu/JTRP/index_html

Published reports of the Joint Transportation Research Program are available at: <http://docs.lib.purdue.edu/jtrp/>

NOTICE

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views and policies of the Indiana Department of Transportation or the Federal Highway Administration. The report does not constitute a standard, specification or regulation.

COPYRIGHT

Copyright 2014 by Purdue University. All rights reserved.
Print ISBN: 978-1-62260-319-0
ePUB ISBN: 978-1-62260-320-6

1. Report No. FHWA/IN/JTRP-2014/04	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle A Systematic Approach to Identifying Traffic Safety Needs and Intervention Programs for Indiana: Volume II—SNIP2 User Manual		5. Report Date April 2014	6. Performing Organization Code
7. Author(s) Andrew P. Tarko, Mingyang Li, Mario Romero, Jose Thomaz		8. Performing Organization Report No. FHWA/IN/JTRP-2014/04	
9. Performing Organization Name and Address Joint Transportation Research Program Purdue University 550 Stadium Mall Drive West Lafayette, IN 47907-2051		10. Work Unit No. 11. Contract or Grant No. SPR-3616	
12. Sponsoring Agency Name and Address Indiana Department of Transportation State Office Building 100 North Senate Avenue Indianapolis, IN 46204		13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code	
15. Supplementary Notes Prepared in cooperation with the Indiana Department of Transportation and Federal Highway Administration.			
16. Abstract This report presents the results of JTRP Project "A Systematic Approach of Identifying Safety Intervention Programs for Indiana (SNIP2)," which aimed to develop SNIP2 to support identification of roads that have excessive crashes of the types defined by the user. In addition, this tool is capable of selecting the best combination of high-crash roads and relevant safety interventions that maximizes the safety benefits and keeps the total cost within the budget and other user-defined constraints. Unlike other studies considering the implementation time of safety projects, the optimization objective of SNIP2 is to identify an optimal combination of countermeasures renewable within a long time horizon. This simplification is accomplished by representing the projects through their annualized costs and benefits. It allows consideration of many projects for large road networks and it makes the SNIP2 suitable for identification of safety focus areas in strategic safety plans. The SNIP optimizer – a heuristic approximation of a large-size mixed integer knapsack problem based on a greedy search was extensively tested and evaluated. It was found producing optimal or near-optimal solutions in a sufficiently short time. Another research result is a comprehensive catalog of countermeasures for Indiana – a list of countermeasure names, road and crash conditions for the countermeasure relevance, corresponding crash modification factors, and countermeasure costs. The SNIP2 is computer software developed with close collaboration with the INDOT future users. It includes an updated crash and state road database. A user's manual describes on the necessary details of the software and various aspects of its use. Two example studies are also included in the manual to illustrate its use and to better presents the SNIP2 features.			
17. Key Words road safety, road safety management, safety planning, heuristic optimization, computer tool		18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161.	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 76	22. Price

PREFACE

A SYSTEMATIC APPROACH TO IDENTIFYING TRAFFIC SAFETY NEEDS AND INTERVENTION PROGRAMS FOR INDIANA: VOLUME II—SNIP2 USER MANUAL

The Safety Needs Identification Package (SNIP2) is a tool developed by the Center for Road Safety at the request of the Indiana Department of Transportation, and its development was supported through the Joint Transportation Research Program of Purdue University and the Indiana Department of Transportation. SNIP2 is an extended next version of SNIP, a tool for identifying roads and areas that experience an excessive number of crashes and may require special attention. SNIP2 has the additional ability to select countermeasures to address identified safety needs in a cost-effective way and within a user-defined budget.

This user manual describes SNIP2 and presents examples that illustrate its features. Additional information about the foundation of SNIP2 can be found in *A Systematic Approach of Identifying Traffic Safety Needs and Intervention Programs for Indiana: Volume I—Research Report* (Tarko, Mingyang, Romero, & Thomaz, 2014). The data management procedures that must be performed annually or every other year to ensure that the data used by the SNIP2 are up to date are described in Volume I, Appendix B.

SNIP2 runs in the MS Windows XP/Vista/7/8 environment and requires the MS .NET Framework 4.0, MS SQL Server (minimum 2008 version), and Google Earth or ArcGIS Explorer. The MS .NET Framework is installed automatically during the first installation of SNIP2 if the PC on which it is being downloaded does not have this component. An Internet connection must be turned on when installing the MS .NET Framework and when using Google Earth.

CONTENTS

1. SNIP2 OVERVIEW	1
1.1 Data Renewal Process	1
1.2 Road Network Screening	1
1.3 Road Clustering	1
1.4 Results Visualization	2
1.5 Safety Program Optimizer	2
2. INSTALLATION.	3
3. RUNNING SNIP2.	3
3.1 SNIP2-SQL Connection	3
3.2 SNIP2 Interface.	4
3.3 Folders and Files.	5
3.4 Typical Operations.	6
4. SCREENING	6
4.1 Study Folder	6
4.2 Years, Scope, Element, and Exposure	7
4.3 Element Selection Criteria	7
4.4 Crash Selection Criteria	7
4.5 Saving and Retrieving Queries	10
4.6 Screening Execution and Results	10
5. CLUSTERING	13
5.1 Settings.	13
5.2 Selecting Files	13
5.3 Executing Clustering	16
5.4 Results	16
6. VISUALIZATION	16
6.1 Settings.	16
6.2 Selecting Files	18
6.3 Creating KML File	18
6.4 ArcGIS Tools	20
7. OPTIMIZATION.	24
7.1 Catalog of Countermeasures.	24
7.2 Running Optimization	26
8. EXAMPLE: WIDENING SHOULDERS	33
8.1 Screening	33
8.2 Clustering	37
8.3 Visualization Using KML files	38
9. EXAMPLE: SAFETY PLAN OPTIMIZATION	43
9.1 Screening	43
9.2 Optimization	44
REFERENCES	59
APPENDICES	
Appendix A. Screening Concepts	60
Appendix B. Road Clustering	65
Appendix C. Results Presentation.	67
Appendix D. Optimization.	70

LIST OF TABLES

Table	Page
Table A.1 Exposure measures for different road elements	62
Table A.2 Confidence F and index I for the three screening criteria	64
Table B.1 Levels of statistical evidence	65

LIST OF FIGURES

Figure	Page
Figure 1.1 SNIP Data Renewal Process	1
Figure 1.2 SNIP2 architecture	2
Figure 1.3 Screening module architecture	2
Figure 3.1 Successful connection to SQL server	4
Figure 3.2 Missing SQL server string connection file dialog	4
Figure 3.3 SNIP2 main interface	5
Figure 3.4 Command and tab bars in the SNIP2 interface	5
Figure 3.5 SNIP2 folder structure	6
Figure 4.1 Screening buttons	7
Figure 4.2 Create or open study folder	7
Figure 4.3 Year/scope/element/exposure tab	8
Figure 4.4 Element selection criteria	9
Figure 4.5 Reset element criteria button	9
Figure 4.6 Crash selection criteria tab	10
Figure 4.7 Buttons for saving, retrieving, or resetting queries	11
Figure 4.8 Folder with <i>qry</i> files	11
Figure 4.9 All collapsed branches, except for branches with selected options	11
Figure 4.10 Results pane	12
Figure 4.11 Output table	12
Figure 5.1 Settings	14
Figure 5.2 Selecting data	15
Figure 5.3 Clustering summary	16
Figure 6.1 SNIP2 visualization tab	17
Figure 6.2 Changing color code	18
Figure 6.3 Selecting files for visualization	19
Figure 6.4 KML file displayed in Google Earth	20
Figure 6.5 Joining shape file with output	21
Figure 6.6 Choosing setting for symbology	22
Figure 6.7 Classification settings in symbology	23
Figure 6.8 Final results of symbology in map	24
Figure 7.1 Catalog <i>ctl</i> files	25
Figure 7.2 Catalog of countermeasures	25
Figure 7.3 Running the catalog of countermeasures	26
Figure 7.4 Catalog error message	27
Figure 7.5 Catalog run completion message	27
Figure 7.6 Input data	28
Figure 7.7 Opening the file with the results of the processing catalog	29
Figure 7.8 Saving the optimization problem input	30
Figure 7.9 Constrains tab	31

Figure 7.10 Optimization summary: general results	32
Figure 7.11 Optimization summary: total benefit	33
Figure 8.1 Beginning the screening process	34
Figure 8.2 Segment type = rural two-lane and shoulder width ≤ 6 ft and with no curb	35
Figure 8.3 Choosing the relevance crashes	35
Figure 8.4 Selection tree for the reference and target crashes	35
Figure 8.5 Final results table	36
Figure 8.6 Exporting results to current study query folder	36
Figure 8.7 Saving query definition to current study query folder	37
Figure 8.8 Selecting files for clustering	37
Figure 8.9 Clustering selected files	38
Figure 8.10 Clustering summary	39
Figure 8.11 Visualization tool settings	40
Figure 8.12 High-crash rural two-lane segments with narrow shoulders in Indiana	41
Figure 8.13 Brown County: High-crash rural two-lane segments with narrow shoulders	41
Figure 8.14 High-crash rural two-lane segments with narrow shoulders on SR 135 near Gnaw Bone, Indiana	42
Figure 8.15 Clusters of high-crash rural two-lane segments with narrow shoulders on SR 135 near Gnaw Bone, Indiana	42
Figure 9.1 Segment criteria for adding rumble strips	44
Figure 9.2 Saving the query results and definitions in the current study query folder	45
Figure 9.3 Choosing intersections will assign exposure to “traffic volume”	46
Figure 9.4 Intersections without signals and right angle crashes	46
Figure 9.5 Urban and rural intersections processed separately	46
Figure 9.6 Resulting five queries for this example	47
Figure 9.7 Assembled catalog of countermeasures	47
Figure 9.8 Selecting associated queries	48
Figure 9.9 Saving example catalog of countermeasure	48
Figure 9.10 Selecting the catalog of countermeasures results file	49
Figure 9.11 Optimization settings data	50
Figure 9.12 Safety evaluation program results	51
Figure 9.13 Optimization output data file	52
Figure 9.14 Safety plan cost predication settings	53
Figure 9.15 Safety plan cost prediction results	54
Figure 9.16 Safety plan optimization results: total cost	55
Figure 9.17 Safety plan optimization constraints	56
Figure 9.18 Safety plan optimization results with extra constraints: total cost	57
Figure 9.19 Optimization log file	57
Figure 9.20 Saving the optimization problem input	58
Figure A.1 Scope, elements, and crashes and their selection/screening criteria	61
Figure A.2 Relationship between the index of Frequency I and the significance of Level F	62
Figure A.3 Dependence between crash counts	63
Figure A.4 Case 1: Simulated (blue) versus calculated (red) distributions of m estimates	63

Figure A.5 Case 2: Simulated (blue) versus calculated (green) distributions of m estimates	63
Figure B.1 Clustering algorithm flowchart	66
Figure C.1 Window in ArcGIS	67
Figure C.2 Labeling feature (left: no labeling; right: labeling)	68
Figure C.3 Selection by attribute	69
Figure D.1 Heuristic algorithm flowchart	71

1. SNIP2 OVERVIEW

There are two major components:

1. The Data Renewal Process (DRP) which updates the SNIP2 database (Figure 1.1).
2. SNIP2, which includes the user interface and the up-to-date database produced with the DRP (Figure 1.2).

The DRP is performed on a regular basis, typically once a year, by a dedicated team in charge of maintaining SNIP2 in an up-to-date version.

The SNIP2 tool is a computer application that supports the following four operations (see Figure 1.2):

1. Identification of high-crash road elements (segments and intersection) that exhibit excessive numbers or proportions of crashes of a type defined by the user.
2. Clustering of the identified high-crash road elements into larger sections that exhibit similar safety needs.
3. Visualization of the individual road elements and road clusters on digital maps.
4. Selection of the most cost-effective combinations of road elements and safety countermeasures according to a user-defined budget and other constraints.

1.1 Data Renewal Process

The Data Renewal Process (DRP) includes updating the existing data by acquiring new data at their sources, reformatting them to meet the standards of the Purdue University Center for Road Safety (CRS) database (called also “the master database”), integrating these data into tables that meet the master database specifications, and replacing the existing data. These new formatted and integrated data are then post-processed to prepare them for use by SNIP2. The data maintenance is facilitated by a suite of procedures developed by CRS or available in ArcGIS. The data updating may be performed annually or when a major change of data at any of the data sources occurs to reflect these changes in the SNIP2 database.

The DRP facilitates the updating of the GIS and non-GIS data in a convenient and efficient process. The data

management procedures include ArcGIS geo-processing and VBA implementation and Model Builder codes that are not packaged as a single module, but rather are used separately as needed to maintain the flexibility of the data management process. The DRP acquires data from the sources, reformats and preprocesses it, and links it together. More details can be found in the SNIP2 research report (Tarko, Mingyang, Romero, & Thomaz, 2014). The final stage of data preparation uses SAS and SQL scripts to merge crash data with their respective assigned road elements by re-codifying, rearranging, and renaming variables. The resulting tables are then uploaded in their final format to the SNIP2 database in SQL Server, where they can be accessed by the SNIP2 interface.

1.2 Road Network Screening

The Road Network Screening module facilitates building queries and performs screening tasks that identify crashes and road elements that meet the query criteria. For example, the user may need a list of rural road segments with narrow shoulders that are experiencing a considerable number of severe single-vehicle crashes in order to identify locations where widening shoulders might be justified. The definitions of queries and the results of their execution are saved in a study folder for later project continuation. The user also has an option of saving the queries to libraries to be used in other future studies. The Road Network Screening processor executes the screening task by accessing the SNIP2 database and searches for crashes and roads according to a query currently in use, the results of which are saved in the *Study Queries* folder. These results then can be accessed by these other SNIP2 processors: clustering, visualization, and optimizer (Figure 1.3). The screening process is described in Appendix A.

1.3 Road Clustering

Road segments and intersections that exhibit an excessive number of crashes may be concentrated along longer road sections. Clustering these elements can

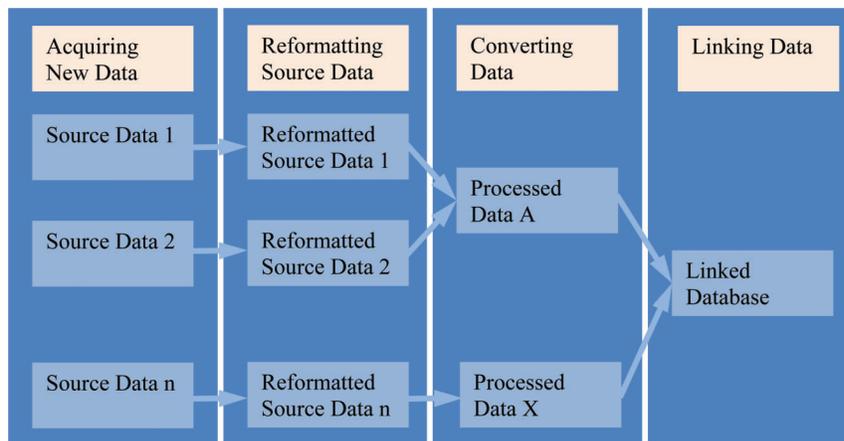


Figure 1.1 SNIP Data Renewal Process.

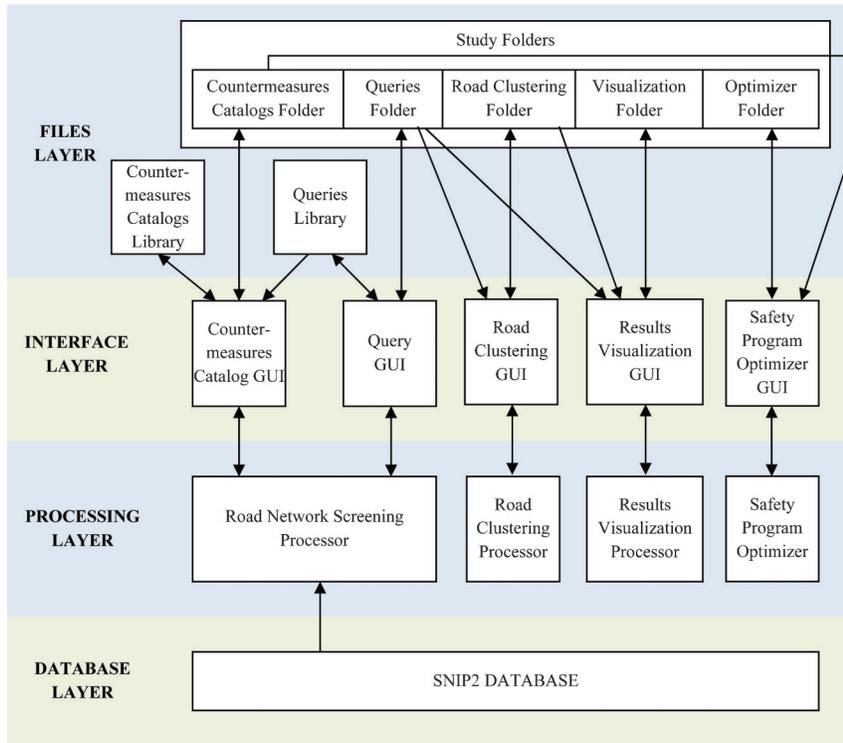


Figure 1.2 SNIP2 architecture.

reveal large-scale safety issues that otherwise might be overlooked if the screening analysis is focused on individual spots. For example, clustering segments with excessive numbers of rear-end crashes may reveal a spill-over safety effect that originated at a signalized intersection with a capacity shortage or where traffic signals are poorly coordinated.

Clustering state road segments and intersections along state routes can help INDOT identify parts of corridors that require certain road improvements from a safety standpoint. These clusters might be found useful in scoping such projects. The clustering procedure is presented in Appendix B.

1.4 Results Visualization

The Road Network Screening module saves the results of a query in a tabular format convenient for

clustering and for additional processing as needed. The final results also may be displayed on GIS maps to visualize the spatial distribution of the identified roads. Such visualization is beneficial in presenting the results to decision-makers and to identify spatial patterns not detectable otherwise. Since the identified road components are geo-coded with their respective latitude and longitude, they can be visualized with the display features offered by Google Earth and ArcGIS. This procedure is presented in the SNIP2 research report (Tarko, Mingyang, Romero, & Thomaz, 2014) and also in Appendix C.

1.5 Safety Program Optimizer

The Safety Program Optimizer facilitates development of a Catalog of Countermeasures which is then utilized to find a set of road elements and relevant

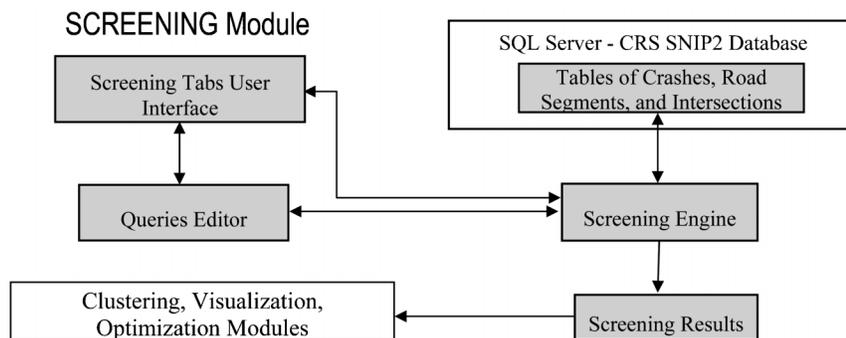


Figure 1.3 Screening module architecture.

safety countermeasures for these roads that maximize the safety benefit within a pre-selected budget level. Thus, the Countermeasures Catalog is the primary input to the optimization process. The catalog includes:

- Countermeasures considered in the optimization.
- Names and directories of query files that include definitions of road and safety conditions (coded in SQL) that justify the countermeasures.
- Crash Modification Factors (CMFs) to evaluate the safety benefit of the countermeasure applied to a relevant road element.
- Costs of countermeasures to estimate the countermeasure's implementation cost.

The queries in the Catalog of Countermeasures are executed before the optimization. Then, the optimization process is executed and its results are saved in the Results Folder. The results are in the form of a list of road elements with applied countermeasures and corresponding economic benefits and costs. Summarized economic indicators are also available.

The Road Network Screening, Road Clustering, and Results Visualization modules are described in more detail in Tarko, Mingyang, Romero, and Thomaz (2014) and in Appendix D. The remainder of this manual focuses on the features of the User Interface and their efficient use.

2. INSTALLATION

SNIP2 is compatible with Windows XP/Vista/7/8. In order to run SNIP2, the following components first must be installed:

1. Microsoft SQL Server (version 2008 or greater). A free version (SQL Server Express) is available at <http://www.microsoft.com/en-us/download/details.aspx?id=29062>
2. MS .NET 4.0 Framework or later. If not present during the installation, SNIP2 will attempt to install this component if the PC is connected to the Internet.
3. Google Earth or ArcGIS Explorer. If Google Earth is not installed, it can be downloaded and installed at no charge at <http://www.google.com/earth/index.html> Alternatively, the latest version of ArcGIS Explorer can be downloaded at <http://www.esri.com/software/arcgis/explorer>.

Once the above components are installed on the personal computer (PC), the next step in the SNIP2 installation is the creation of a folder *My Documents/SNIP2*. The user should then unzip the provided file into a folder of his/her choice. The SNIP2 installation package comes with two main components: the SNIP2 installation program and the SNIP2 database. Copy the SNIP2 database to the *My Documents/SNIP2* folder that was just created.

The administrator of the PC where SQL Server is installed should then attach the SNIP2 database, which was copied into the *My Documents/SNIP2* folder, using the *Attach* function in the SQL Server database menu. The future user of the program also must have or (or must be given) administrative rights over the database, as SNIP2 needs to be able to both read and

write to the database. After the SNIP2 database has been attached to MS SQL Server, the following commands should be executed to give SQL Server permission to run the user-defined functions:

```
sp_configure 'show advanced options', 1;
GO
RECONFIGURE;
GO
sp_configure 'clr enabled', 1;
GO
RECONFIGURE;
GO
```

Once SQL Server is configured, an ODBC connection needs to be established, assigning a connection name to the SNIP2 database. The proper parameters for this connection need to be saved to a file named *snipiconnection.txt* which should reside inside the *My Documents/SNIP2* folder. A sample file is included in the installation package and can be custom-edited. The content of the text file is a simple string, repeated below, where the PC name and SNIP2_connection_name should be replaced with the proper values of the user's PC:

```
String to be edited by user and saved to file My Documents/SNIP2/snipiconnection.txt:
Server =PC_name\SQLEXPRESS; Database =SNIP2_
connection_name; integrated security = true"
```

To install the actual SNIP2 interface, the user should return to the folder where the content of the zip file was extracted and run the installation process by clicking on the *setup.exe* file. The *readme.txt* file explains the installation steps.

Once the program is installed and the user has made sure it is working, the user may delete the unzipped files in the folder with the *setup.exe* file to save disk space. The user should save the zipped/compressed file in case it is needed to reinstall the program. **Do not delete** the *My Documents/SNIP2* folder nor the files in the folder. Also, **do not delete** the files in the folder to which the program was installed.

3. RUNNING SNIP2

The program can be executed from the Windows start menu where the SNIP2 program should be listed if already installed.

3.1 SNIP2-SQL Connection

Every time SNIP2 is run, it will attempt to establish a connection to the SNIP2 database in the SQL Server. It will look for the *SNIP2connection.txt* file inside the *SNIP2* folder. This file will contain the string that establishes the parameters needed to connect to the user's SQL server. If the file is found, the program will read the string and establish the proper connection. A message announcing the successful connection (Figure 3.1) will be momentarily flashed on screen before the full interface loads.

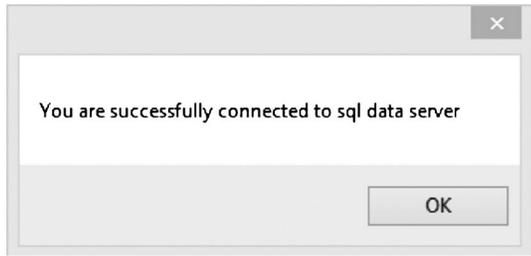


Figure 3.1 Successful connection to SQL server.

In case the *SNIP2connection.txt* file is not found, a file dialog will open asking the user to locate where a file with the proper connection string can be found. A copy of any selected file will be copied to the *SNIP2* folder and renamed *SNIP2connection.txt*. The connection string then will be read and connection established. If for some reason the user desires to connect to a different SQL Server, he or she can use the *Settings/Read Configuration File* menu option and the same file dialog window will appear, allowing for a new *SNIP2connection.txt* to be selected. (See Figure 3.2.)

Assuming that the connection to SQL Server is successful, SNIP2 then initializes itself and removes any temporary tables remaining open from a previous run. The main interface should appear within several seconds (Figure 3.3).

3.2 SNIP2 Interface

The SNIP2 interface includes a command bar and tabs arranged below the command bar (Figure 3.4). The command bar facilitates operations on study files and folders and changes to the general SNIP2 settings. The tabs arranged in the second row facilitate defining queries and catalogs, execution of screening, road

clustering, graphical visualization, and optimization operations. The third row of queries appears as needed after selecting one of the tabs in the second row.

3.2.1 Screening Tab

The *Screening* tab opens another row of six tabs to facilitate the screening operations (Figure 3.4). These screening tabs allow the user to specify the screening conditions (i.e., types of road elements and types of crashes) to be considered in searching for high-crash roads. These screening conditions, once defined, can be executed right away or saved as queries for future use by means of the appropriate buttons present inside the screening tab pane. The user also has an option to reset the conditions of a query. Chapter Screening explains how to use the screening tabs.

3.2.2 Clustering Tab

The second tab of the interface is the *Clustering* tab. It allows clustering the road elements identified in the screening phase. The clustering combines road elements experiencing similar safety problems into larger parts of the road network based on their safety performance. The clusters may be suitable for determining the scope of certain safety studies. Chapter Clustering provides more details about the clustering tool.

3.2.3 Optimization Tab

The third tab opens the optimization pane. This module allows the user to develop new or select an existing catalog of countermeasures with the collection of countermeasures to be processed. In addition, the user specifies one of the three types of analysis to be performed:

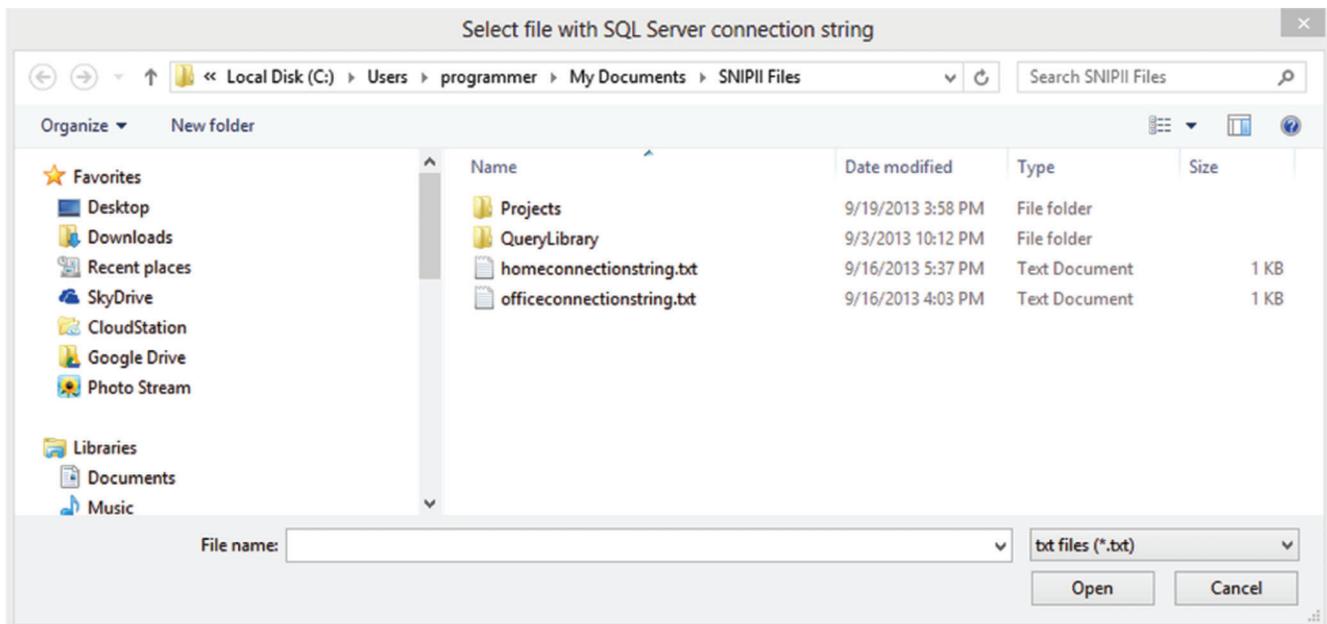


Figure 3.2 Missing SQL server string connection file dialog.

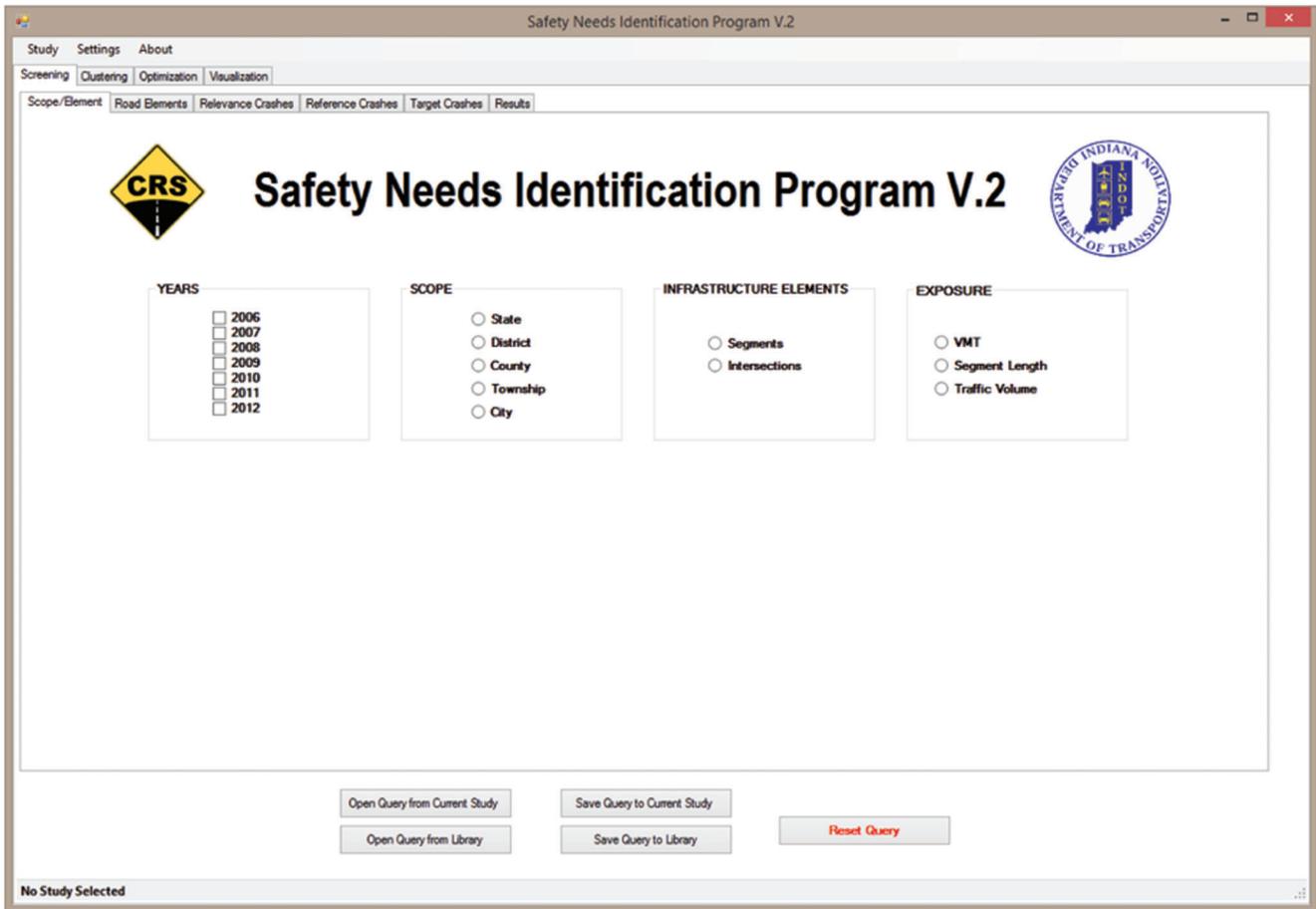


Figure 3.3 SNIP2 main interface.

1. Safety plan optimization
2. Safety plan cost prediction
3. Safety plan evaluation

Depending on the type of analysis, the user may need to specify the total budget and other constraints. The details of these three analyses are explained in chapter Optimization.

3.2.4 Visualization Tab

The fourth and last tab, *Visualization*, allows specifying the visualization settings, creating Keyhole Markup Language (KML) files for results visualization, and execution of visualization employing Google Earth or other earth browser such as ArcGIS Explorer Desktop. The KML files can also be displayed outside

of SNIP2 by other suitable tools such as ArcGIS. Chapter Visualization explains the visualization tool.

3.3 Folders and Files

When SNIP2 is run for the first time, it will generate three subfolders inside the SNIP2 main folder. The *Study* subfolder is the suggested place for the user to create and store new study files. The other two subfolders are *QueryLibrary* and *CatalogLibrary*. The *QueryLibrary* folder is where the user can store query definitions that he or she deems useful for future studies, thereby saving the user's time defining queries again. The queries stored in this library may be imported, customized, and saved later as needed.

In a similar fashion, the *CatalogLibrary* folder should be used to store user-created catalogs of countermeasures

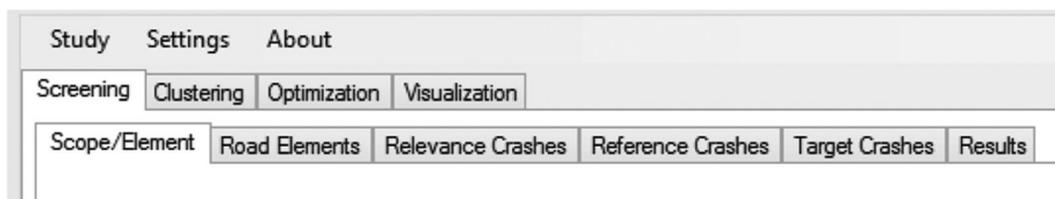


Figure 3.4 Command and tab bars in the SNIP2 interface.

that may be re-used or used as building blocks for other catalogs in the future. The basic SNIP2 folder structure for studies and queries is shown in Figure 3.5. It should be noted that although this folder structure is suggested and automatically created by SNIP2, the user is given the freedom to create study folders anywhere in the PC. Regardless of where they may be created by the user, the study folders will always contain five subfolders to store the files created by the different modules (see Figure 3.5, study 4) Once an existing study folder is selected for use, SNIP2 will also create and save its files properly.

3.4 Typical Operations

The SNIP2 program allows performing a number of operations by selecting the appropriate tabs. Some of these tasks need to be performed in a specific order, while other tasks can be executed separately. The tasks that a user may typically perform include:

1. Creating new *study folders* or selecting existing ones. These folders and their subfolders are used to store query definitions, cluster results, visualization files, catalogs of countermeasures, and optimization analyses. Every study folder created automatically includes subfolders for each of these categories.
2. Screening road segments or intersections for crashes of user-defined types. This screening is executed with queries defined by the user. These queries can be saved and executed individually or as a batch of queries organized in a *catalog of countermeasures*.
3. Ranking safety needs by certain types of segments and intersections. This screening is based on safety performance measures that reflect the number of the *relevance* and *reference crashes* on these segments and at intersections. The resulting ranked list is saved in a Comma Separated Values (csv) file placed in the queries subfolder of the current study folder.
4. Clustering segments and intersections in groups that exhibit excessive numbers of crashes of certain type while the surrounding roads do not. Such a grouping of high-crash roads may help scope safety improvement studies. The resulting clustered roads are saved in csv files in the clustering subfolder of the current study folder.

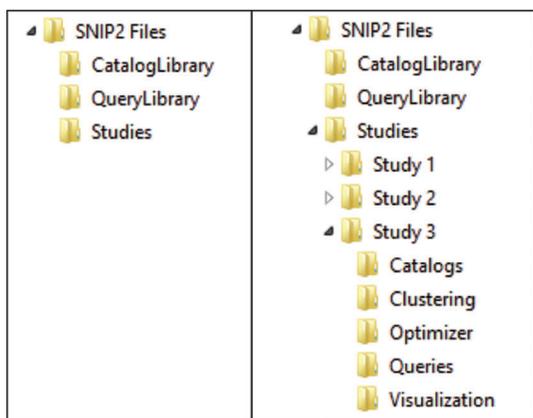


Figure 3.5 SNIP2 folder structure.

5. Displaying the results (high-crash roads) in Google Earth or ArcGIS Explorer. Since the program creates *KML visualization files*, the user may visualize results from previously run analyses as long as the files are saved in the visualization subfolder, respective query, or clustering subfolders of the current study folder.
6. Creating new or editing existing *catalogs of countermeasures*. A catalog of countermeasures is a set of predefined queries supplemented with additional inputs. The catalog of countermeasures is read by the *optimizer module*. The user can add a query to a catalog by reading it from the queries folder of the current study, from the queries folders of other studies, or from the *query library folder*.
7. Running the optimizer to obtain the best combination of relevant safety countermeasures for the identified safety needs. These results are obtained from a catalog of countermeasures by maximizing the safety benefits under limited resources and are saved to the optimizer sub-folder.

4. SCREENING

The screening process includes a number of steps and the six tabs contained inside the main screening tab facilitate entering all the inputs required by the screening task, executing the screening, and inspecting the results. The screening process also requires creating, reading, and saving files, which are facilitated by the file operation buttons arranged at the bottom of the interface (Figure 4.1). The user has options to open and save queries in the current study or the query library.

The user has a number of reset options available. The red button at the bottom of the pane, visible in all six screening sub-tabs resets the query completely, erasing all user choices. Alternatively, the buttons present in the upper region of the tabs for extracting road elements, relevance, reference, or target crashes can reset the choices made only in those individual tabs.

The screening tabs open the input windows of a convenient query editor that can help a user unfamiliar with SQL build queries by selecting crash properties and road properties from the possible. SNIP2 converts this selection to SQL expressions and executes them in the MS SQL server. The direct screening results include the road elements and crashes that meet the query conditions, which then undergo a number of statistical computations; and the final results are presented to the user in a tabular form. The results then can be used with the clustering tab or accessed with the visualization tab. Queries, if saved, can later be organized into catalogs of countermeasures to be run in a batch mode for the optimizer module.

4.1 Study Folder

Although it is not necessary for the user to create or open an existing study folder in order to perform a screening session, it is advisable to do so in order to later be able to save the query and the results. A study can be selected or created at any time before or during

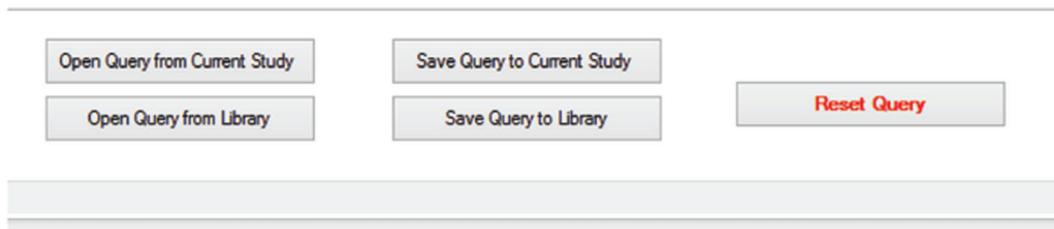


Figure 4.1 Screening buttons.

the creation of a query by selecting the proper menu option (Figure 4.2).

If a new study folder is created, then it becomes the study folder in use. Every created study folder has five subfolders dedicated to different tasks. If the user decides to save queries or create new catalogs of countermeasures, SNIP2 will save them in the current study and in the proper subfolder. The user has the option to choose a different folder.

4.2 Years, Scope, Element, and Exposure

The first screening tab (Figure 4.3) takes the user to the screening pane where the selection of the years, scope, elements, and exposure of the study are executed in four fields.

1. *Years.* The years with crashes available in the database are presented. The user defined the study period by selecting years. It is advisable that a period of consecutive years be selected.
2. *Scope.* Five scope options are available: state, district, county, township, and city. Clicking any option other than state opens a window displaying a list of smaller geographical units: districts, counties, townships, or cities from which the user can select. Multiple geographical units within a given scope can be selected.
3. *Infrastructure element.* SNIP2 offers two different types of road elements that can be explored: segments and intersections. When an element type is chosen, only the appropriate set of element selection criteria is available in the *Road Elements* tab.
4. *Exposure.* Depending on the element type chosen, different measures of exposure are available for selection. *VMT* and *Length* are choices for segments, and *Traffic Volume* is automatically chosen for intersections.

4.3 Element Selection Criteria

The Road Elements tab opens a page that allows the user to select specific road types and geometry conditions for consideration in the study as selection

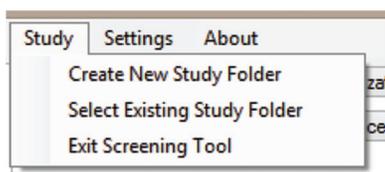


Figure 4.2 Create or open study folder.

criteria. This page serves two types of roads: segments and intersections. The type of road selected in the *Scope* field will be active while the other option will be disabled and greyed out. Figure 4.4 illustrates the case where segments were previously chosen by the user. The available selection criteria for intersections are greyed out.

The selection criteria are presented as a tree. The selection group can be expanded to show choices that belong to that group or can be collapsed to hide the choices. The user can select multiple values by marking the boxes in front of the desired multiple choices (Figure 4.5).

After the selections are properly marked, the user should click on the *Extract Element Data* button. The mouse pointer is replaced with a turning wheel while a popup window warns the user to wait until the data extraction is completed. Once the query processor is finished, the mouse pointer returns and a text box at the bottom of the page displays the SQL commands that summarize the selection criteria just executed by SQL Server.

After the extraction is completed, the *View Extracted Elements* button prompts the user to view a table containing the road elements that satisfy the current selection criteria. A button to reset the element choices made by the user is also available (Figure 4.5). Similar selection reset buttons are present on each of the different crash selection tabs.

4.4 Crash Selection Criteria

4.4.1 Relevance Crashes

The *Relevance Crashes* tab allows the user to define the types of crashes used to estimate the road safety performance. If the screening is designed to identify roads for a given selected countermeasure (e.g., widening road shoulders), then the crashes caused by the road deficiency to be eliminated with the countermeasure should be specified as relevance crashes (e.g., run-off-road crashes).

There are three groups of crash criteria, which are organized in selection trees: environmental criteria, vehicle criteria, and person criteria (Figure 4.6). Since the interface has certain limitations, it is important that the user understand how the SNIP2 interprets the choices made with the selection trees. The following three rules apply.

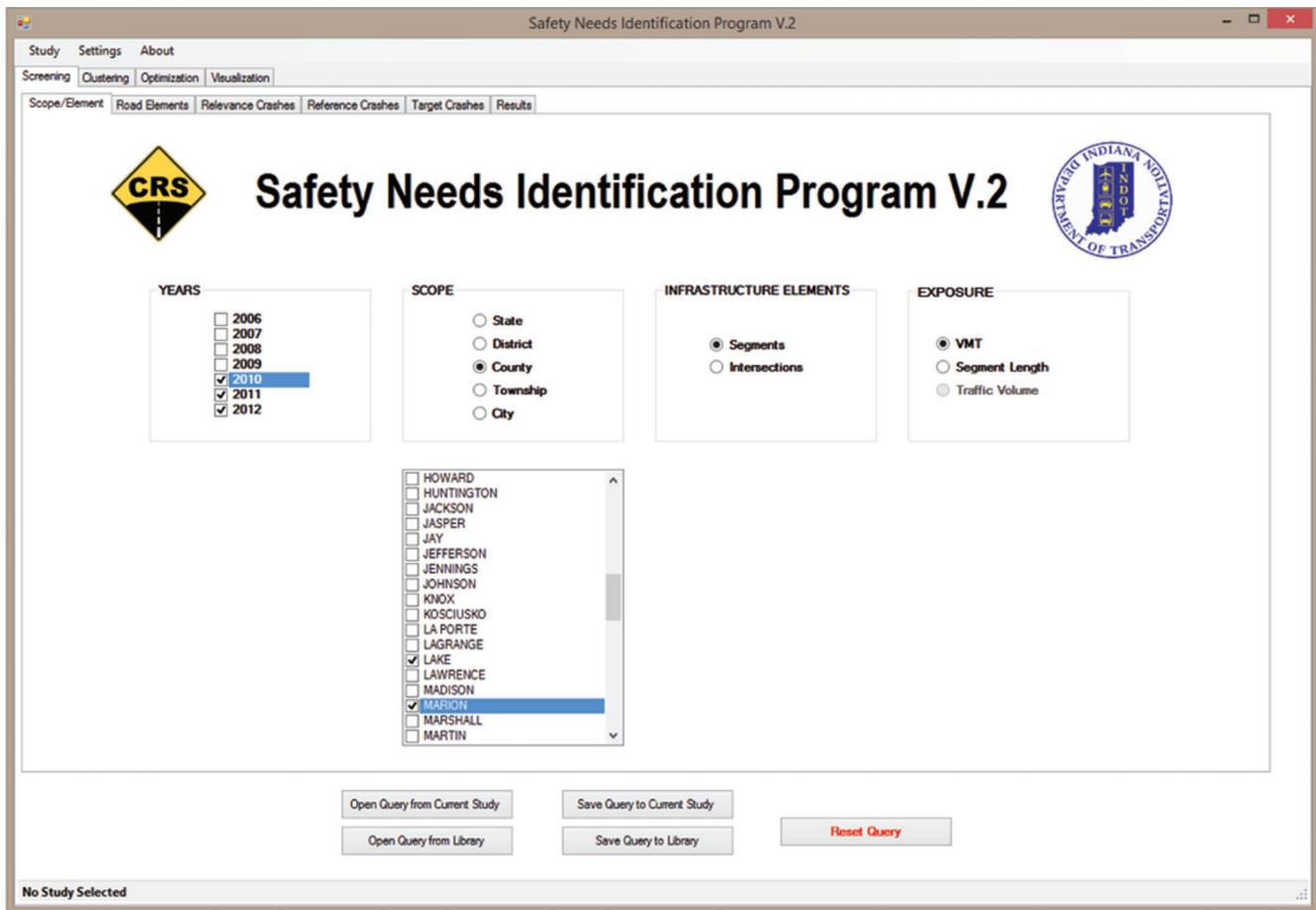


Figure 4.3 Year/scope/element/exposure tab.

Rule 1. Multiple options selected in the same group of options are equivalent to setting two conditions connected with the logical operator “OR.” For instance, selecting the *Daylight* and *Dawn/Dusk* options for the *Light Conditions* group is equivalent to looking for crashes that occurred during the daylight OR the dawn/dusk conditions:

Light Conditions = Daylight OR Light Conditions = Dawn/Dusk.

Rule 2. Multiple values selected in different groups are equivalent to setting two conditions connected with the logical operator “AND.” For instance, selecting the *Daylight* option in the *Light Conditions* group and selecting the *Friday* option in the *Day of Week* group is equivalent to looking for crashes that happened on Fridays AND during the daylight conditions:

Light Conditions = Daylight AND Day of Week = Friday.

Rule 3. If the user defines the environmental, vehicle, and person criteria, then the crashes that meet all these criteria are selected with the logical operator “AND.” For example, defining the environmental criterion shown as an example in Rule 1 and selecting the *Van* option in the *Vehicle Type* group in the *Vehicle Criteria* tree is

equivalent to searching for crashes involved a van AND occurred in the daylight OR dawn/dusk conditions: (Light Conditions = Daylight OR Light Conditions = Dawn/Dusk) AND Vehicle Type = Van.

After the criteria for selection is defined, the user should click the *Extract Relevance Crash Data* button. The mouse pointer becomes a turning wheel while a popup window advises the user to wait until the data extraction is completed. Once the extraction is completed, a text box at the bottom of the page displays the resulted SQL query submitted to SQL Server. The button *View Extracted Crashes* allows the user to view a table containing the crash data which satisfy the criteria for the type of crashes specified by the user.

4.4.2 Reference Crashes

Reference crashes constitute a larger class of crashes that include relevance crashes. For example, injury crashes at an intersection may serve as reference crashes to right-angle injury crashes at that intersection. In another example, all crashes on a road segment may be used as reference crashes for fatal crashes on the same segment. In some studies, the user may want to know the proportion of relevance crashes in the reference

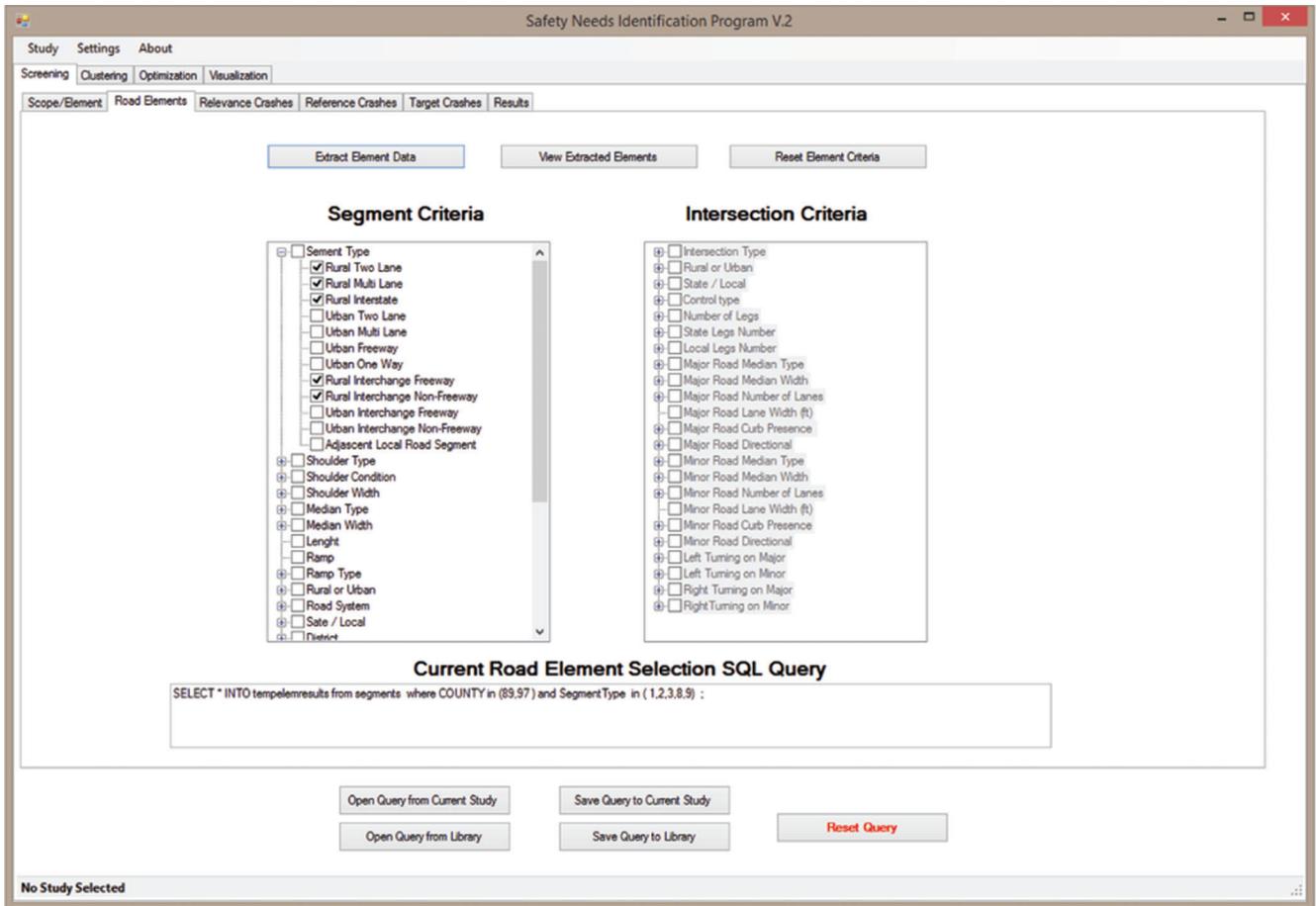


Figure 4.4 Element selection criteria.

crashes. SNIP2 offers this option and also evaluates if the crash proportion on a road element is significantly higher than the average proportion in the selected scope of road elements.

Prompted by the order of the screening tabs, the selection conditions for reference crashes may be defined after doing so for relevance crashes. After the relevance crashes are extracted, the page called by the *Reference Crashes* tab presents the three selection trees collapsed, except for groups where at least one selection was made in the *Relevance Crashes* tab. Presenting the selection tree in this way helps the user to expand the relevance crash class to make it suitable for becoming a reference crash class. The user can expand the crash class by selecting more options in the groups where at

least one selection was made for defining relevance crashes. It must be emphasized that removing ALL already selected options is equivalent to removing this group from consideration. It can be accomplished by unmarking each option or clicking twice on the box in front of the group.

Once the criteria for selecting reference crashes are determined, the remaining steps are the same as already described for relevance crashes.

4.4.3 Target Crashes

The *Target Crashes* tab allows defining the target crashes, which are the crashes that would be affected by a safety countermeasure. Their number is multiplied by

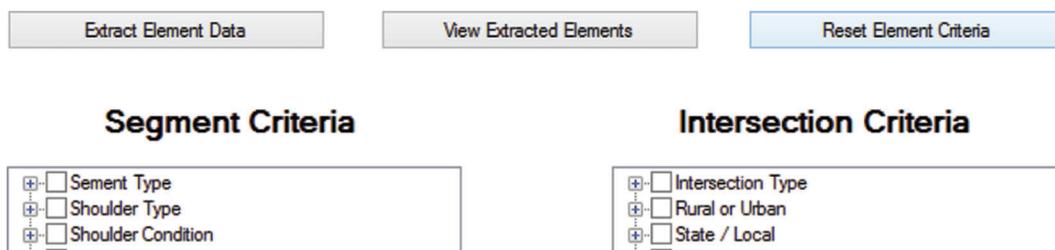


Figure 4.5 Reset element criteria button.

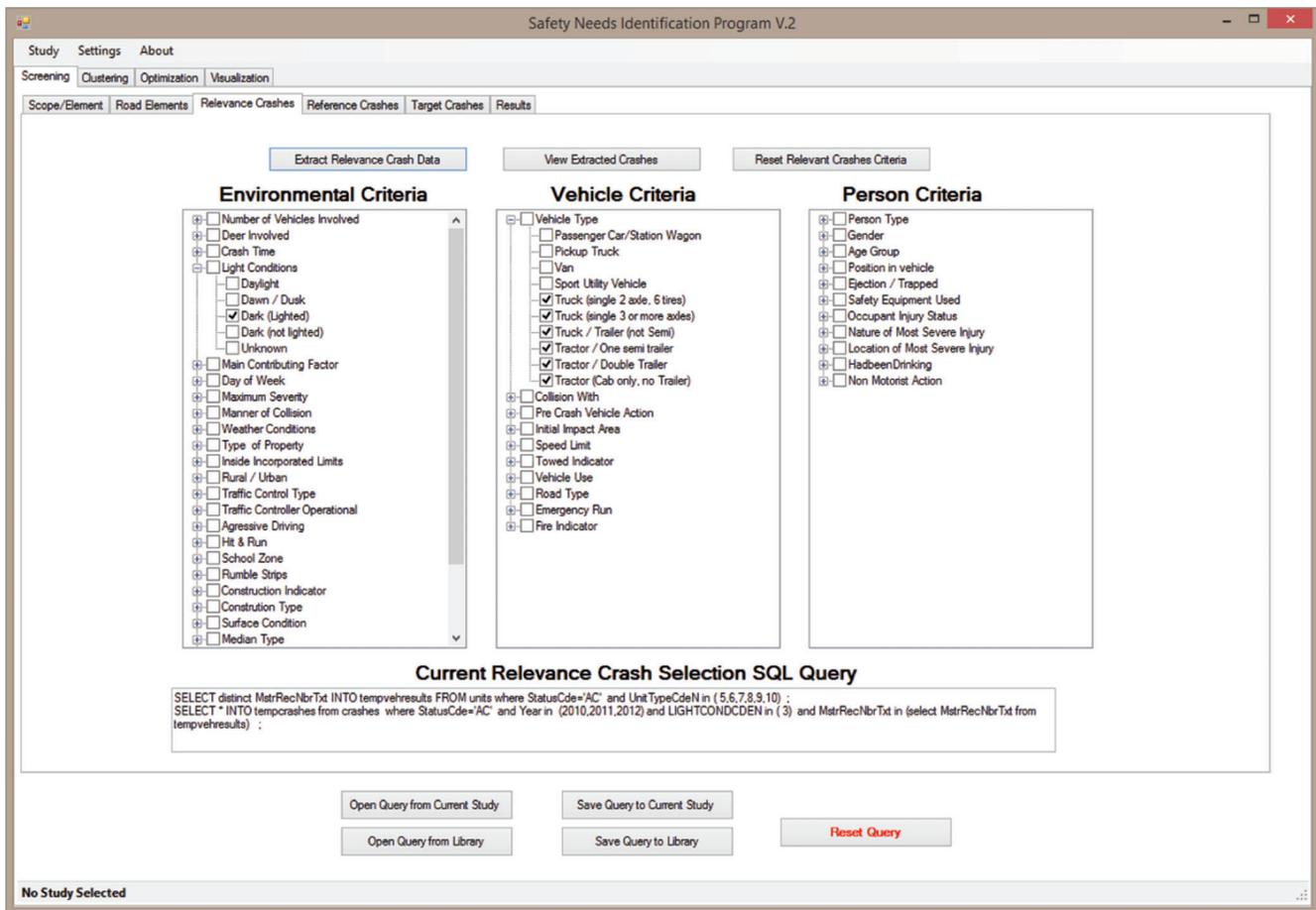


Figure 4.6 Crash selection criteria tab.

the corresponding CMF to estimate the reduced number of crashes after the countermeasure is implemented. The difference between the original and reduced crashes is the source of the safety benefit estimated in the optimization module.

The operations in this tab are identical to the *Relevance Crashes* tab.

4.5 Saving and Retrieving Queries

The user can define a new selection of conditions for roads and crashes or can open a saved query by clicking on the button that corresponds to the desired query file operation (Figure 4.7.) In SNIP2, query files have the *qry* extension (Figure 4.8). An opened query will be presented to the user via the marked options in the three selection trees. The selection trees will have all branches collapsed except for the query choices selected (Figure 4.9). The user may then continue editing the opened query or, if all criteria have been already defined, may decide to execute it.

At any time while using the screening tabs, the user can save the currently defined selections to a query file. To save a query, the user clicks on the button that saves the query to the current study or the *QueryLibrary* folder. It is

important to note that the saved query does not include any results, only the query coded in the text format.

The user can also reset the query (both road and crash conditions) at any time by clicking the red *Reset Query* button.

4.6 Screening Execution and Results

Clicking the *Start Screening Calculations* button executes the screening operation. The safety performance measures are calculated as well as the indicators for their statistical confidence. The output pane (Figure 4.10 and Figure 4.11) includes detailed information about all the elements that satisfy the element selection criteria in the scope. The latitude and longitude coordinates are also provided. This information is useful in visualizing the results with ArcGIS or Google Earth. The results also include the crash statistics as well as the safety performance values and the corresponding confidence levels.

By clicking the *Export Results* button, the user saves the results to csv file that can be imported for use by the SNIP2 clustering and visualization modules. The default folder to which to save the results is the *Queries* subfolder of the current study.

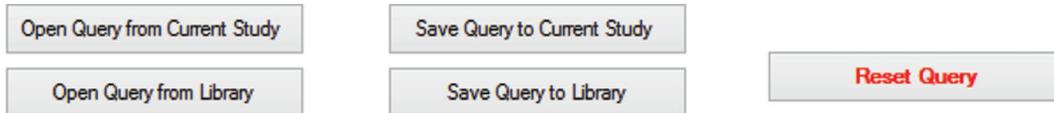


Figure 4.7 Buttons for saving, retrieving, or resetting queries.

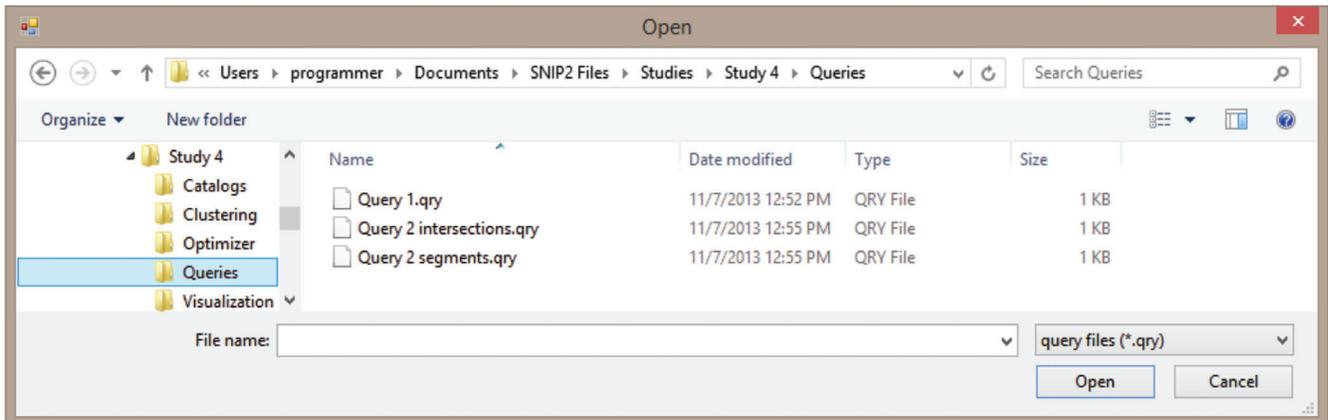


Figure 4.8 Folder with *qry* files.

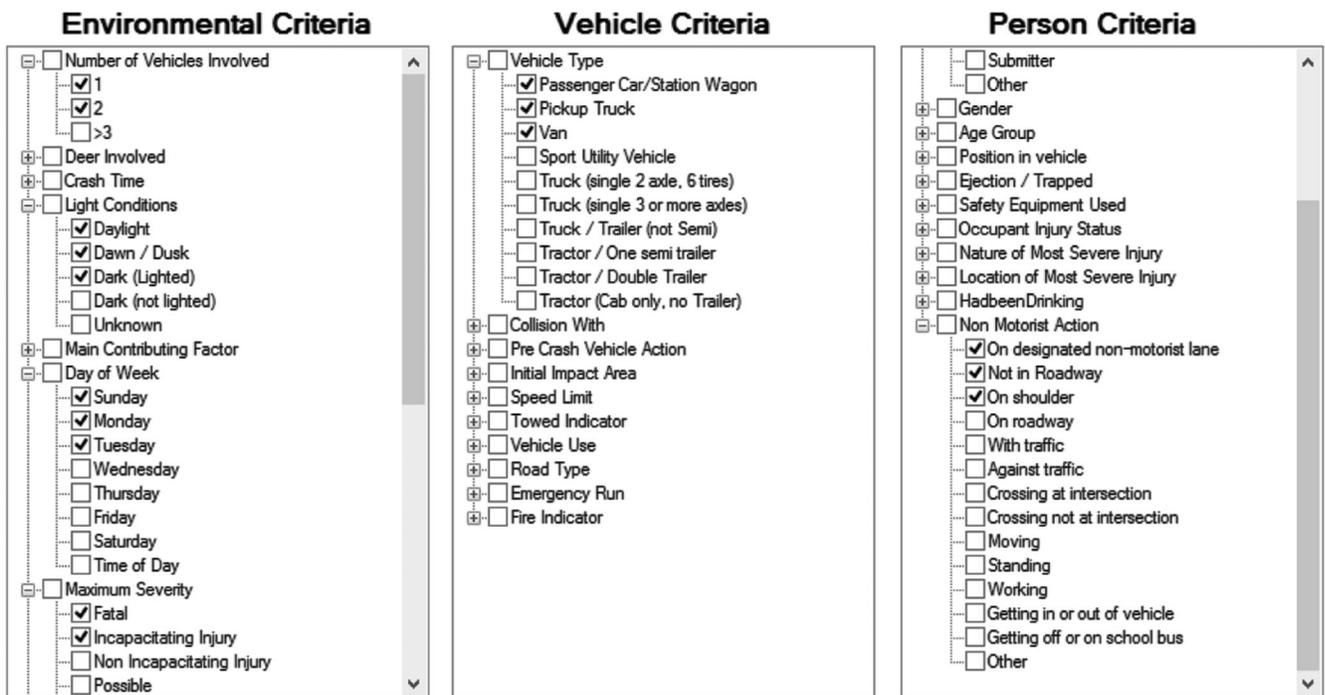


Figure 4.9 All collapsed branches, except for branches with selected options.

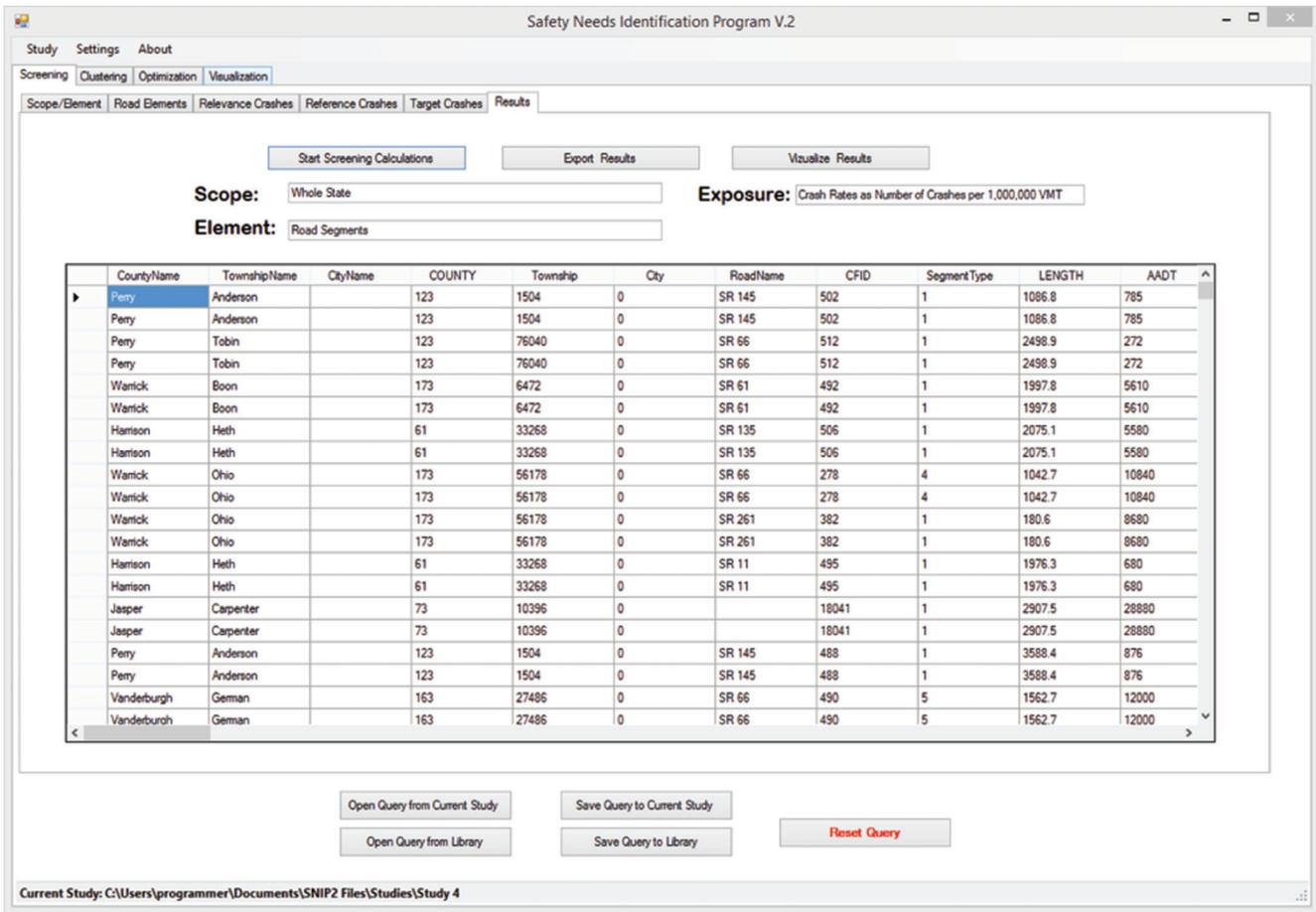


Figure 4.10 Results pane.

CountyName	TownshipName	CityName	COUNTY	Township	City	RoadName	CFID	Segment Type	LENGTH	AADT
Lake	Center		089	11476	0	I-65	33557	3	2746.3	40820
Lake	Center		089	11476	0	I-65	33558	3	2746.3	40820
Lake	Center	Crown Point	089	11476	16138	I-65	300379	3	5732.7	40820
Lake	Eagle Creek		089	19306		I-65	300378	3	5453.9	33000
Lake	Eagle Creek		089	19306	0	I-65	31360	3	2687.7	33000
Lake	Eagle Creek		089	19306	0	I-65	31361	3	2687.7	33000
Lake	Eagle Creek		089	19306	0	I-65	31356	3	2687.7	33000
Lake	Eagle Creek		089	19306	0	I-65	31364	3	2687.7	33000
Lake	Eagle Creek		089	19306	0	I-65	33560	3	2746.3	40820
Lake	Eagle Creek		089	19306	0	I-65	31362	3	2687.7	33000
Lake	Eagle Creek		089	19306	0	I-65	31359	3	2687.7	33000
Lake	Eagle Creek		089	19306	0	I-65	33565	3	2746.3	40820
Lake	Eagle Creek		089	19306	0	I-65	33566	3	2746.3	40820
Lake	Eagle Creek		089	19306	0	I-65	33567	3	2746.3	40820
Lake	Eagle Creek		089	19306	0	SR 2	21279	1	2515.8	6860
Lake	Eagle Creek		089	19306	0	SR 2	21286	1	2340	6860
Lake	Eagle Creek		089	19306	0	SR 2	21287	1	3020.8	9070
Lake	Eagle Creek		089	19306	0	SR 2	21289	1	2147.6	9070
Lake	Eagle Creek		089	19306	0	SR 2	21290	1	2386.2	9070
Lake	Eagle Creek		089	19306	0	SR 2	32900	1	250	6860

Figure 4.11 Output table.

5. CLUSTERING

The screening component identifies which road elements experience an excessive number of crashes. Clustering these elements into longer road sections may reveal useful spatial patterns that otherwise may go unnoticed and may be helpful to INDOT engineers in scoping corridor improvement studies and other safety-oriented programs.

It is important to note that elements with safety needs should be clustered based on the safety performance measures in order to obtain road clusters that are relevant from the safety management point of view.

There are three basic safety measures that can be used to identify road elements with an excessive number of crashes of a certain category: crash frequency, crash rate, and crash proportion. These measures were introduced earlier in this report and presented in Appendix A. Crashes are subject to a strong random fluctuation over time and two safety performance indices: Confidence F and Adjusted Index I_A , are proposed to estimate the level of statistical confidence that the detected excessive number of crashes indicates a systematic issue rather than just the effect of random fluctuation.

Confidence F is the probability of a safety level equal to or better than the one observed during the period of analysis if the expected safety level in the long run is average for the type of location and under the given exposure. The higher confidence F is, the stronger is the evidence that the location experiences a real safety problem. Values of $F=0.90$ and higher are typically used.

Adjusted index I_A is the difference between the safety observed during the period of analysis and the safety expected given the location type and exposure, divided by the standard deviation of the difference estimate on an equivalent normal distribution. It is a simplified measure of Confidence F . Values of $I_A=1.5$ and higher provide sufficient evidence that the location is experiencing a real safety problem.

An additional and important criterion of considering a road element as a high-crash location is the minimum number of crashes. This criterion addresses the limitation of the F statistic for roads with very low crash frequency such as low-volume county roads. In extreme cases, a segment with zero crashes can have a value of F higher than the threshold for high-crash locations. This result is correct from the statistical point of view because zero may be a highly likely outcome for low-volume roads. On the other hand, selecting such a road as a high-crash location is incorrect from the safety management point of view.

One of the important operations of clustering road elements is evaluation of the safety level in the current clusters in order to be able to claim that the obtained clusters experience excessive numbers of crashes. A practical method of updating safety evaluation in clusters is aggregation of safety measures.

Three user-selected threshold values control building clusters: the minimum number of crashes, the threshold confidence F_1 (or I_1) and the threshold confidence F_2 (or index I_2). The clustering module builds a cluster starting

with the road element which has the highest $F > F_1$ (or $I > I_1$) and the number of crashes at least equal to the minimum value. The algorithm allows adding a road element if to the current cluster if: (1) the element is adjacent to the currently built cluster, (2) it has the confidence F (or index I) greater than the threshold F_2 (or threshold I_2), and the confidence F for the cluster after adding the new element is still higher than the threshold F_1 (or threshold I_1). When no additional element can be added to the cluster, the clustering tool stops building the current cluster and searches for a next road element suitable to build a new cluster. The clustering ends when no suitable road elements can be found.

The user can restrict the clusters building only along the same routes to follow the common practice in scoping road studies. Other restrictions may be added to the algorithm as needed. A list of clusters and their elements is obtained based on the screening results, the network topology, and the parameters set by the user.

The clustering component requirements are presented below as well as a description of the interface of the *Clustering* tab, which includes the data importation process, the user settings, the clustering calculations, and the results.

A description of each necessary step to run the process is shown below, which includes the user settings, the open data files, the clustering calculations, and the results.

5.1 Settings

The clustering process uses the three user-selected threshold values: minimum number of crashes, threshold confidence F_1 (or index I_1) and threshold confidence F_2 (or index I_2). It seems that the minimum number of crashes set at 2 serves its purpose. The user may use a larger value if it better reflects the local policy. The recommended ranges for the other settings are: (0.9–0.97) for F_1 , (0.5–0.9) for F_2 with the recommendation that $F_1 > F_2$, or (1.25–2) for I_1 , (0–1.25) for I_2 with the recommendation that $I_1 > I_2$.

Thus, before running the clustering tool, there are basic settings needed (Figure 5.1): selecting the variable type to control the process either adjusted index I or confidence F , setting thresholds F_1 or I_1 and I_2 or F_2 , selecting the clustering variable criteria and the exposure options for segments and intersections, and allowing (or restricting) the clusters building along the same routes to follow the common practice in scoping road studies by selecting (deselecting) the *Allow Clustering Crossing Roads* check box. Allowing this option may lead to clusters that “turn” at an intersection to another state-administered route.

5.2 Selecting Files

Input to the clustering tool is included in the screening tool results. One or more files with the results must be selected. Clicking on the *Select Files* button shown in Figure 5.2 opens the Windows file selection window where the appropriate files can be selected. It is important to note that it is necessary to open all the

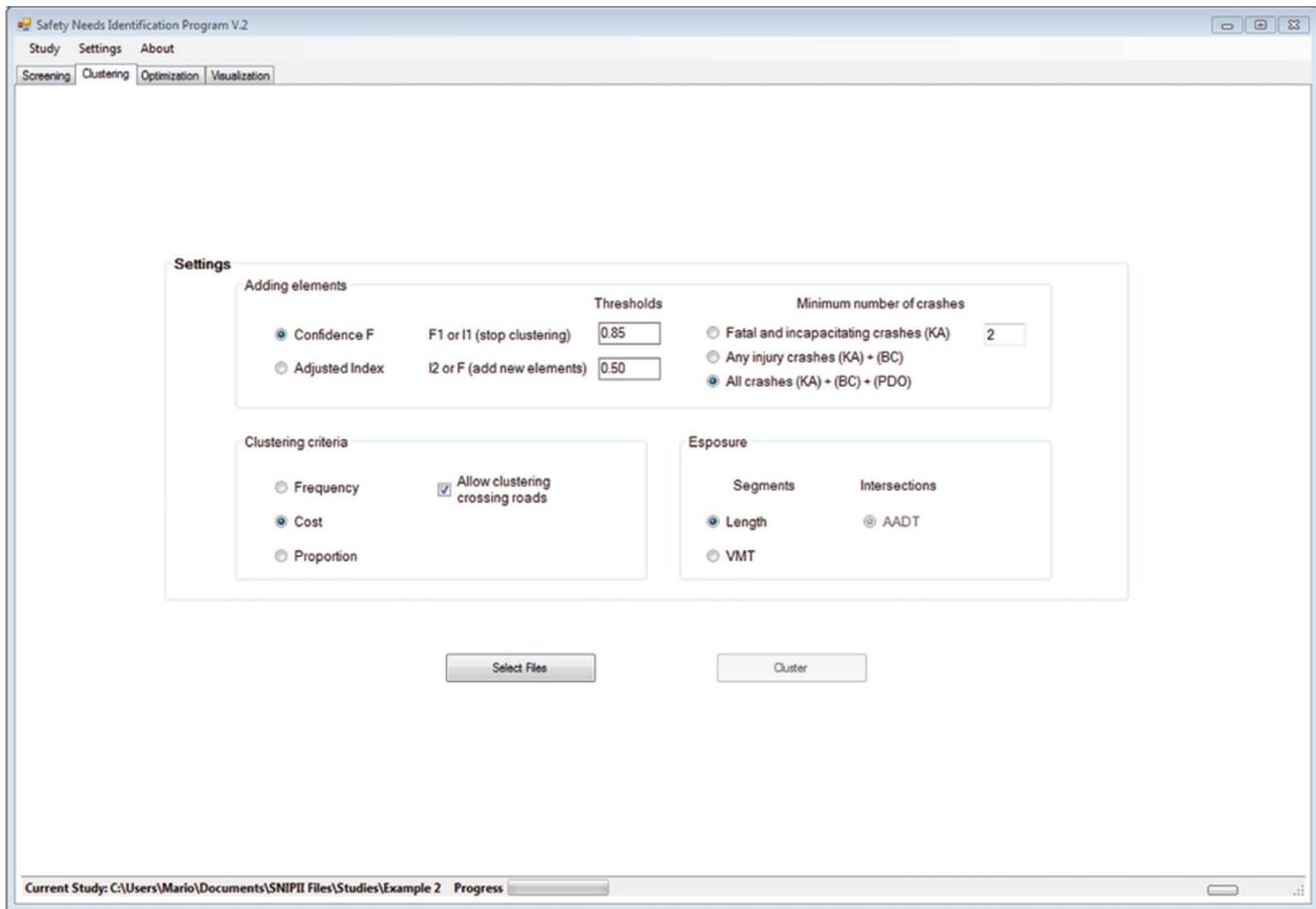


Figure 5.1 Settings.

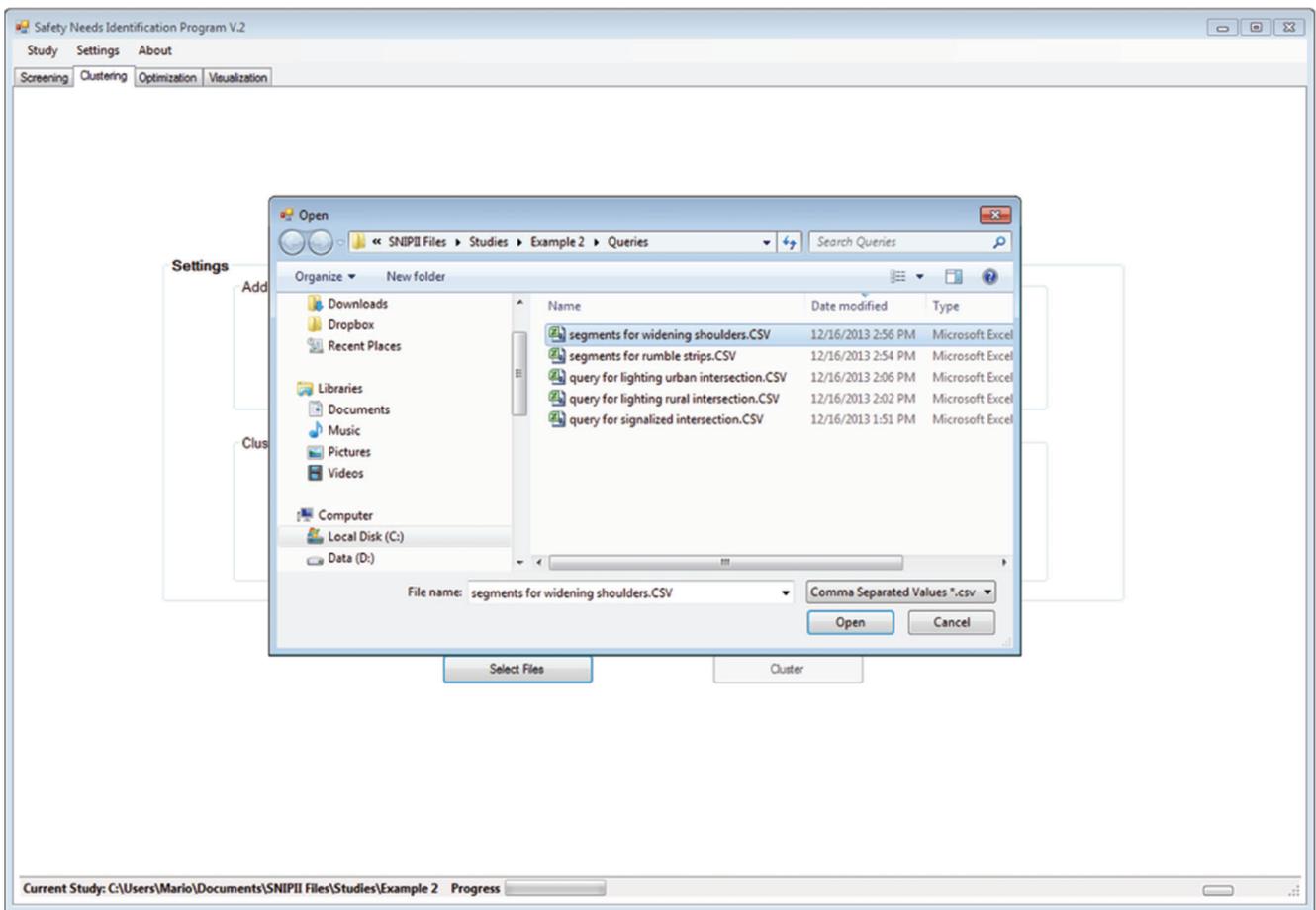


Figure 5.2 Selecting data.

files at once, which is accomplished by highlighting the selected files using the control key and clicking the *Open* button.

5.3 Executing Clustering

Once the settings and the data opening have been completed, the *Clustering* button is enabled. By clicking on it, the clustering process begins. It is important to note that this process may take several minutes. After it ends, the user is asked to provide a file name and a folder to which the results are saved.

At the end of the clustering process, a window opens with a summary report that shows the total number of clusters found, the total segments included, and the total intersection points included, as shown in Figure 5.3.

5.4 Results

The clustering process creates one csv file that contains the segment and intersection information for the clusters.

The file contains the Element ID and the Cluster ID, which is the ID of the cluster to which this element belongs. This file is used in the visualization process as explained in chapter Visualization.

6. VISUALIZATION

The user has two options for visualization of the results:

1. The visualization tool embedded in the SNIP2 package that creates files to be visualized using any earth browser such as ArcGIS Explorer Desktop or Google Earth.
2. Linking the data to an external shape file using ArcGIS.

An Internet connection is needed to run the visualization option via either Google Earth or ArcGIS Explorer. Both methods are explained in this section.

The results presentation is the final step of SNIP2. In this phase, the user can visualize the results obtained in the screening or clustering steps. This SNIP2 component, shown in Figure 6.1, allows creating KML files for both the screening and the clustering components. KML is the file format used to display geographic data in an earth browser such as Google Earth, Marble, ArcGIS Explorer Desktop, etc.

6.1 Settings

Before running the clustering tool, there are four basic settings needed: color code settings first, then line width, transparency, and variables. (See Figure 6.2.)

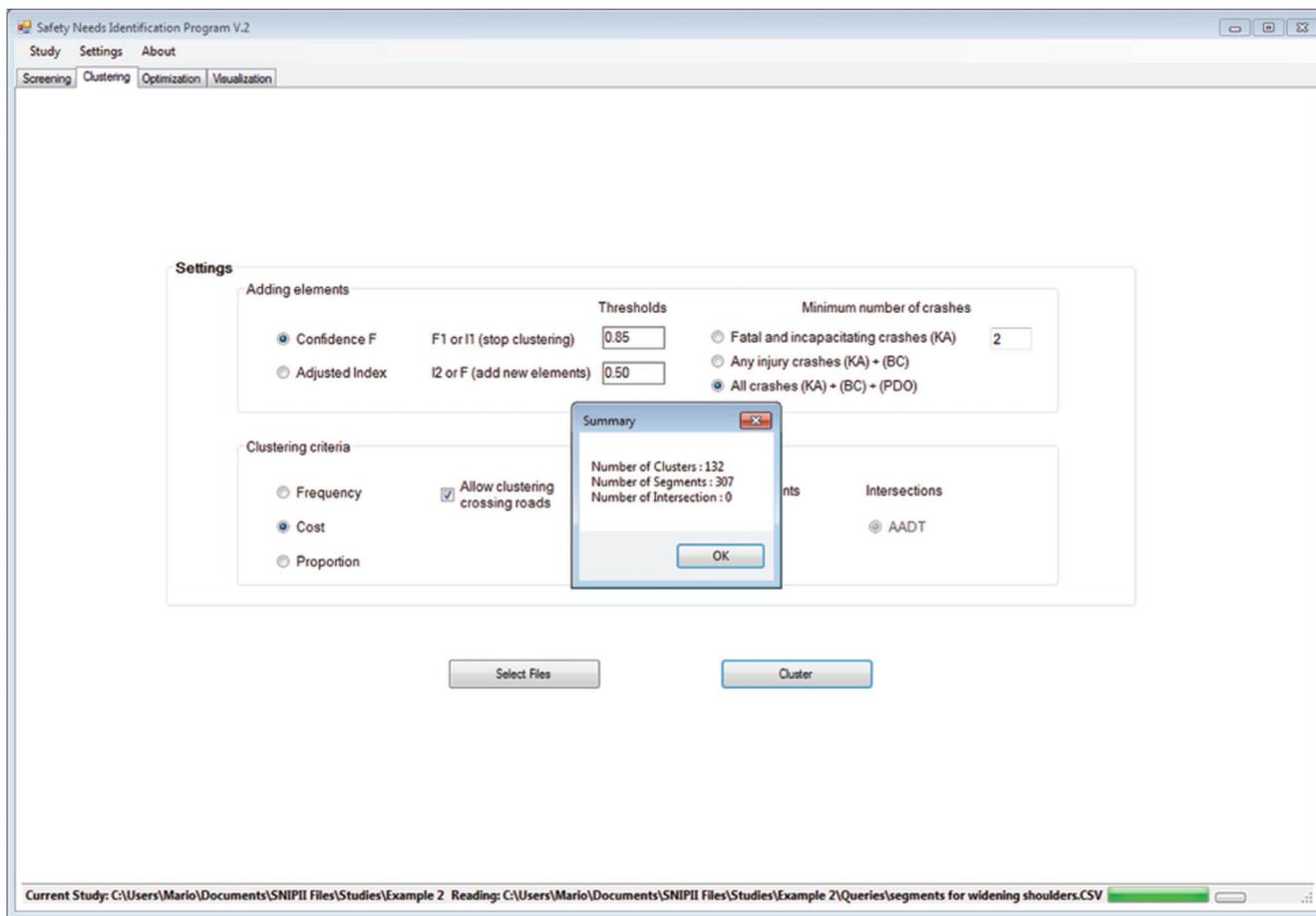


Figure 5.3 Clustering summary.

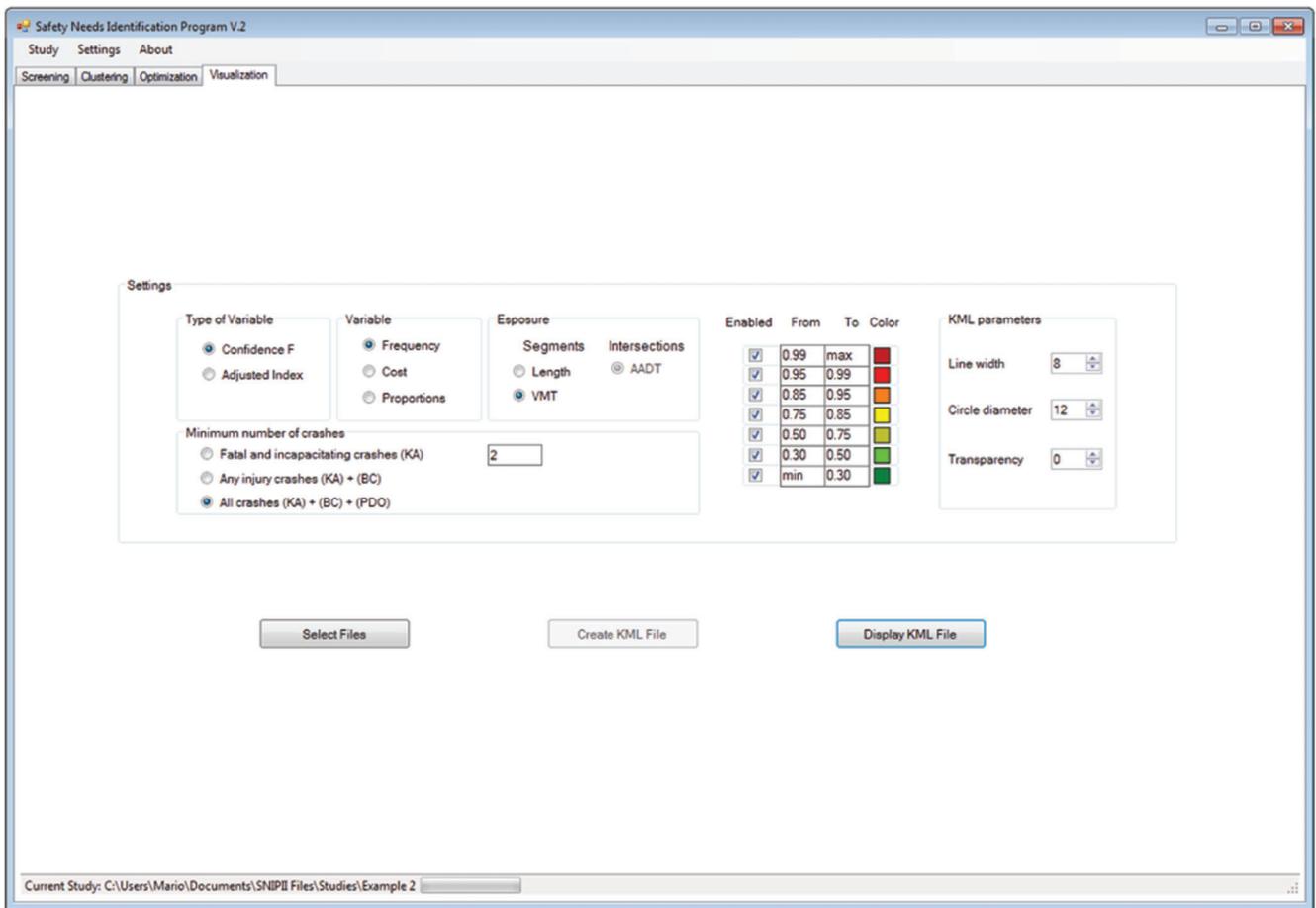


Figure 6.1 SNIP2 visualization tab.

The display settings include:

- Width of the displayed lane
- Circle diameter for displaying intersections
- Level of transparency
- Criterion for adding elements to a folder in the KML file (I_2 or F_2 and minimum number of crashes)
- Safety performance measure for segments and intersections
- Enabling/disabling ranges of values
- Minimum and maximum values of the ranges
- Color displayed for the range

The user should be aware that changing between I and F displays the default ranges of the values, thus any user-entered values will be wiped out.

6.2 Selecting Files

Clicking on the *Select Files* button shown in Figure 6.3 opens the Windows file selection window where the user may select the files with the results for display. Unlike in the case of clustering, where the user has to open all the files at once, this time the files for visualization can be converted to the KML format one by one or all at once. In both cases, multiple KML files are created. The user can display all the created KML files in Google Earth or other earth browsers simultaneously.

6.3 Creating KML File

Clicking on the *Create KML File* button creates and saves KML files in the visualization folder of the study, with the same name as the input files that were processed but which add an indication of the selected safety performance measure.

KML is the file format used to display geographic data in a GIS browser, such as Google Earth, Marble, or ArcGIS Explorer Desktop. To display the results, Google Earth or ArcGIS Explorer Desktop must be installed. If neither is currently installed, Google Earth can be downloaded at no charge at <http://www.google.com/earth/index.html>. You can download ArcGIS Explorer Desktop at no charge at <http://www.esri.com/software/arcgis/explorer/download>.

Upon opening (double-clicking), Google Earth is called automatically to open the SNIP2-created KML file. Google Earth displays the full extent of the KML file with all the elements shown according to the user's settings. For zooming in, just double-click on the zone you want to zoom. The zoom level also can be controlled with the mouse wheel. It is important to note that if the KML file reaches the maximum number of elements, Google Earth can display in a single view, and then a subset of elements is not shown. If this

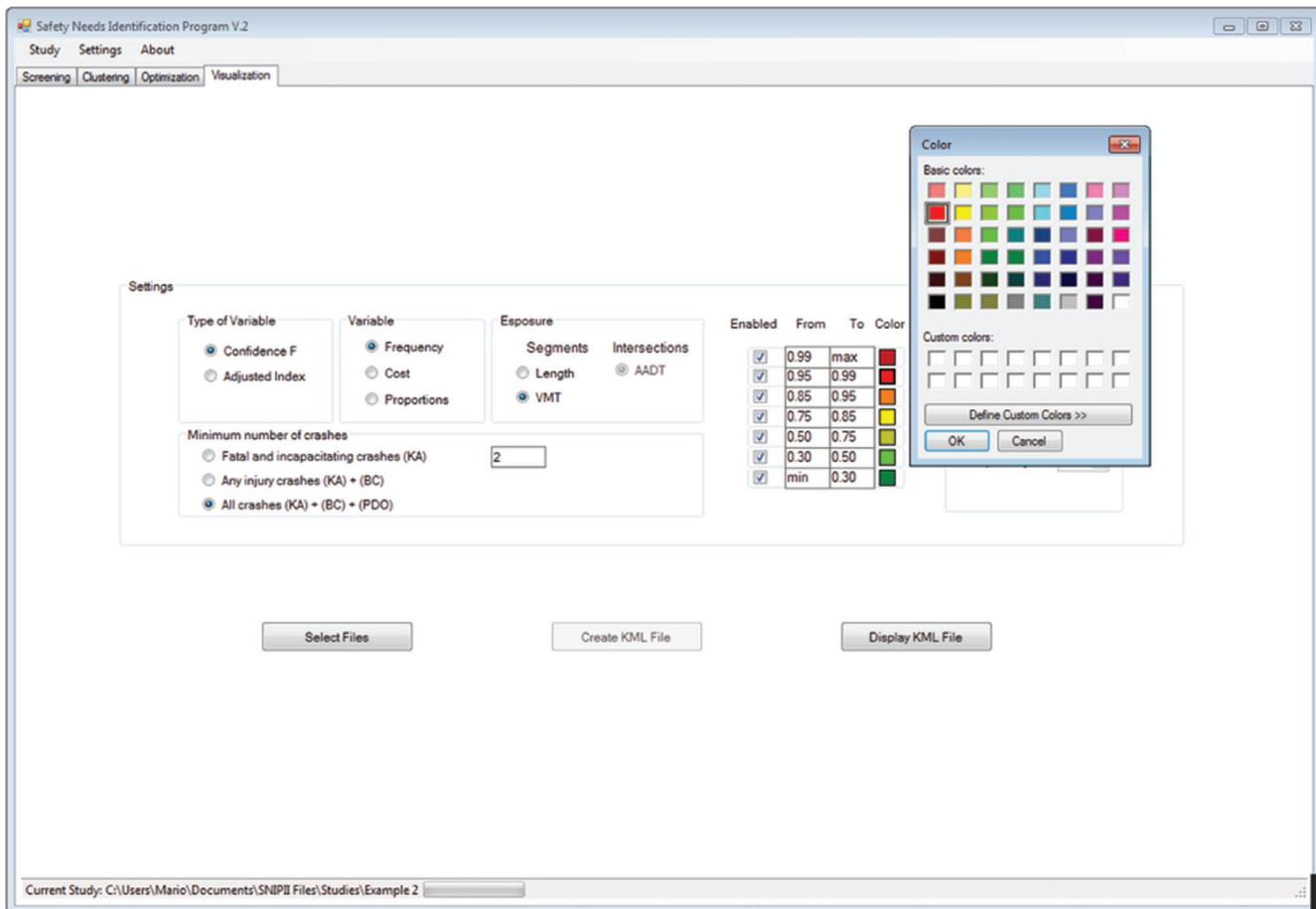


Figure 6.2 Changing color code.

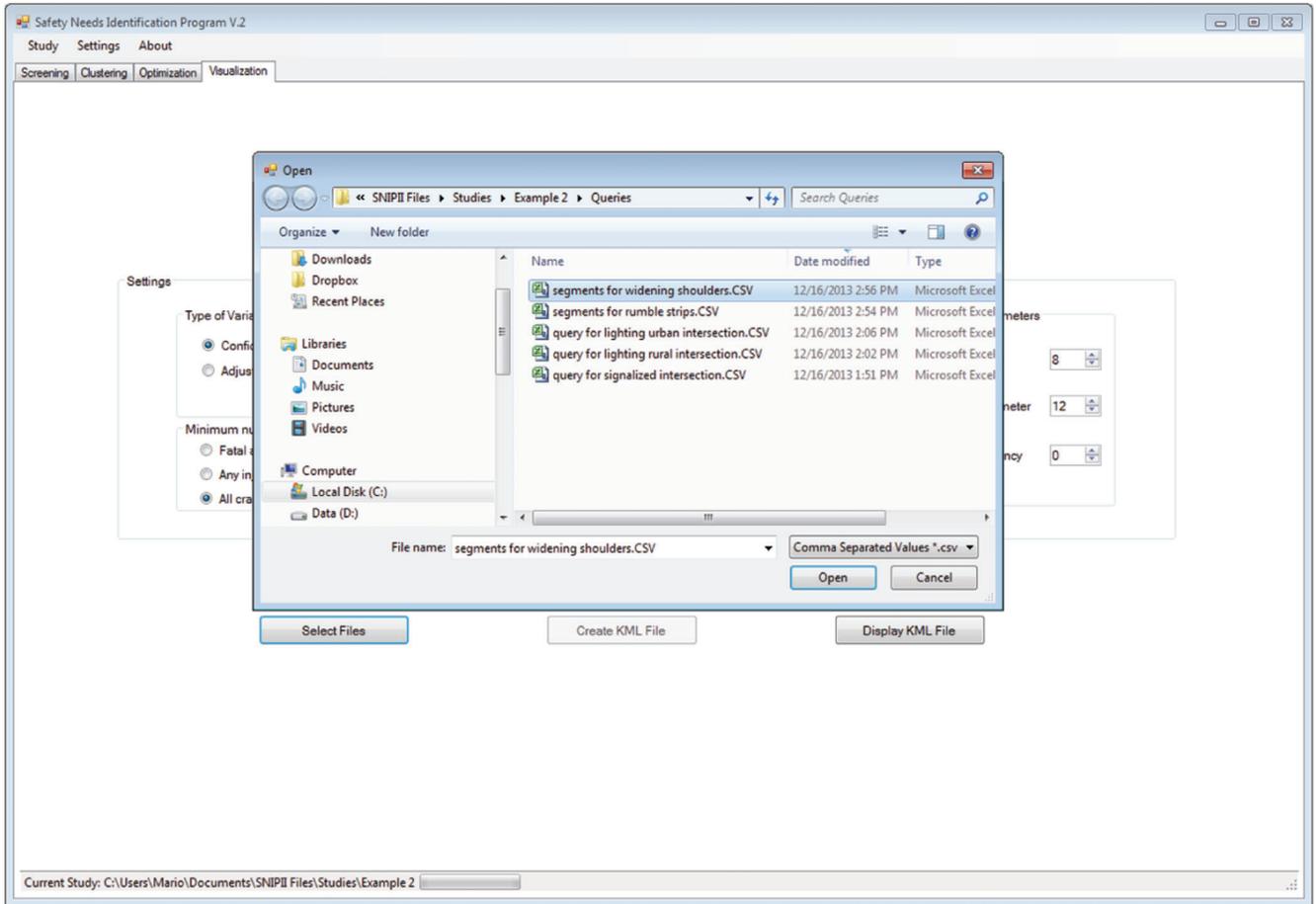


Figure 6.3 Selecting files for visualization.

situation happens, you need only to zoom in to display fewer elements at the same time.

By clicking on one of the displayed elements, a pop up window with the calculated indexes and confidence values and the element ID is shown (Figure 6.4).

You will find more information about using Google Earth at <http://www.google.com/earth/learn/>.

6.4 ArcGIS Tools

ArcGIS provides many tools that can accomplish the visualization task. Among the various features available in ArcGIS, the following three visualization tools are widely used:

- Symbology
- Labels
- Selection by attributes

Symbology refers to visualization of feature (i.e., a single element), categories of elements, quantities, etc., by colors or symbols. Symbology has a special procedure to prepare tools like bar charts or pie charts as part of an individual element. A step-by-step approach is presented as an example.

6.4.1 Joining Output with Shape File

Open the output file and join with the existing shape file that corresponds to it. Here, the segment shape file needs to be joined with the output table that ranks the rural two-way lanes having narrow shoulders. For joining purposes, the CFID from the segment file

should match with the CFID in the output file. (See Figure 6.5.)

6.4.2 Select Symbology Settings

After the output table is joined with the segment table, right click on the joined shape file and click on *Properties*.

Click on the *Symbology* tab, then *Quantities*, and finally *graduated colors*. Now the user needs to add the total crash field (or any other field to symbolize) as input for the Field. The user may choose a specific color ramp for symbology. For example red to green ramp can show the high crashes in red and low crashes in green. (See Figure 6.6.)

ArcGIS provides options for user-defined and standard classification for Symbology. In order to change the default classification scheme and values, click on the *Classify* button. The classification window will appear. In the drop-down list in the *Classification* window, there are a number of options for classification methods (e.g., Equal Interval, Quantile, Natural Breaks, etc.). Also there is an option called *Manual* for user-defined classification settings. The user can see the distribution of the *Index* in the window along with vertical lines showing the current classification (Figure 6.7).

Also, when the *Manual* option is selected, the user can slide the vertical line by dragging the mouse to change the upper/lower limits for a particular class.

Click *Apply* after all the parameters for symbology are selected. Now, the user will be able to see the high crash location in red. (See Figure 6.8.)

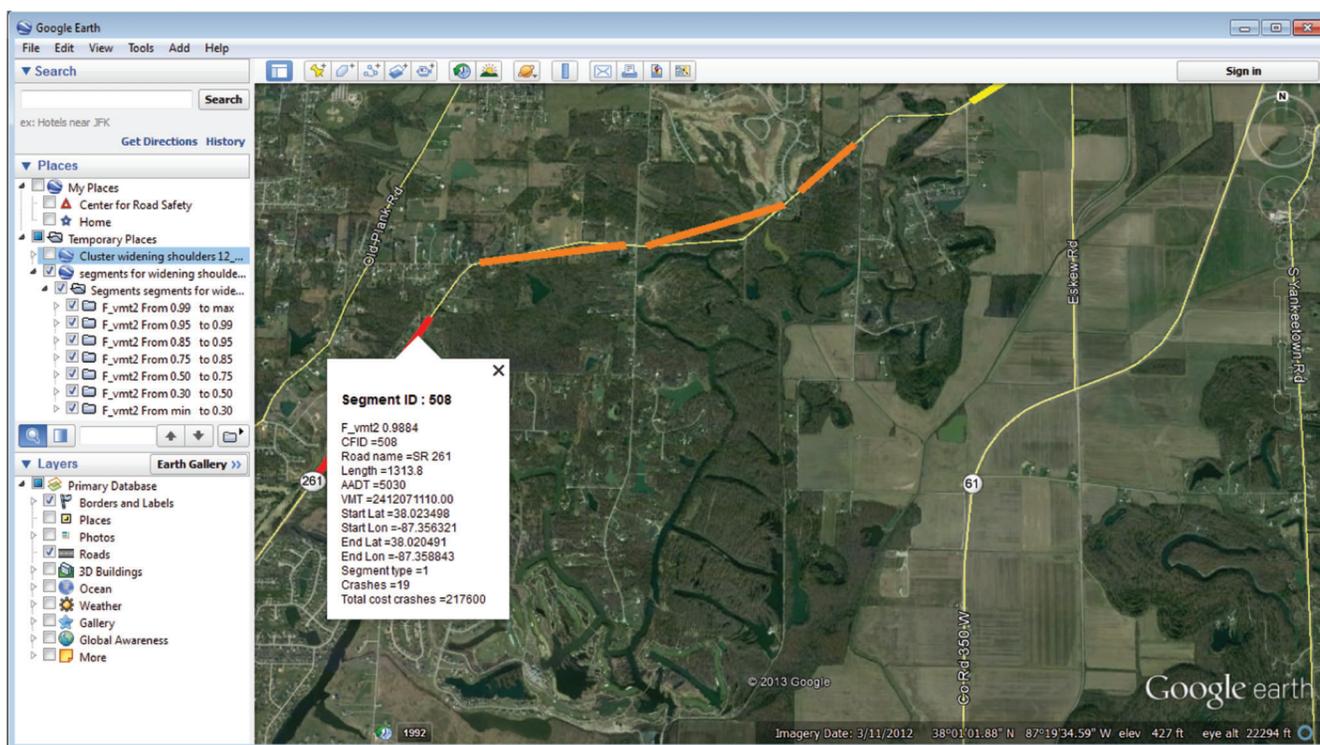


Figure 6.4 KML file displayed in Google Earth.

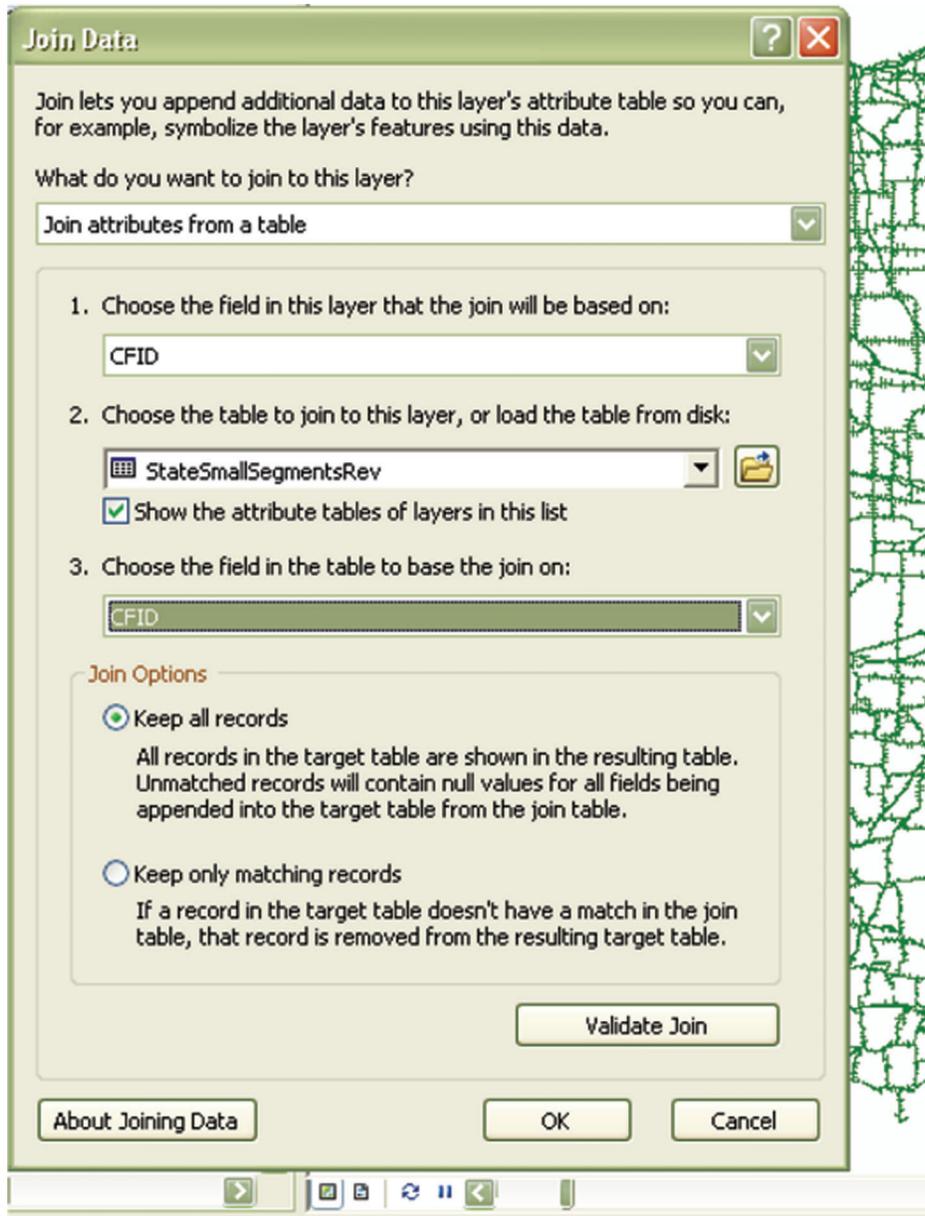


Figure 6.5 Joining shape file with output.

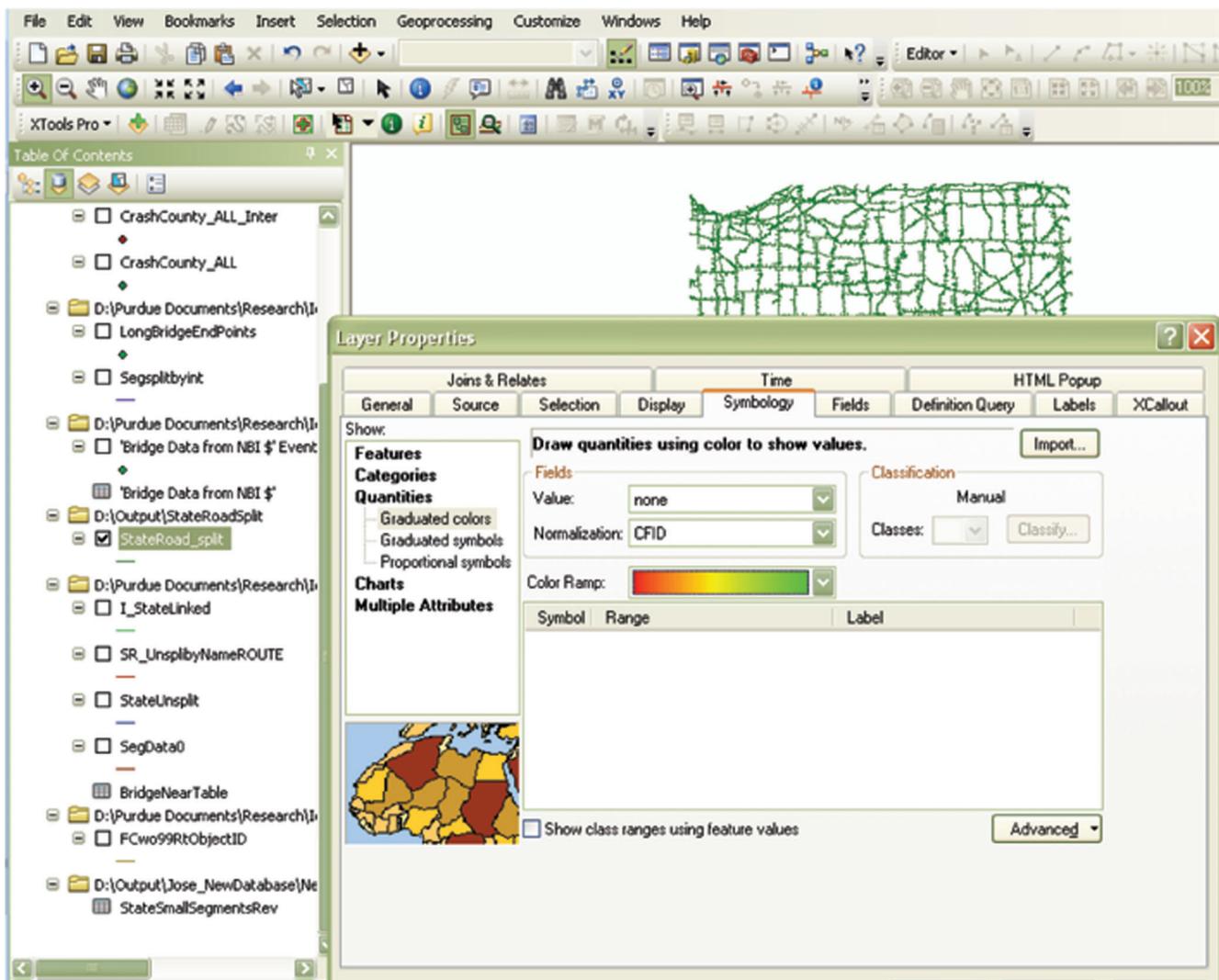


Figure 6.6 Choosing setting for symbology.

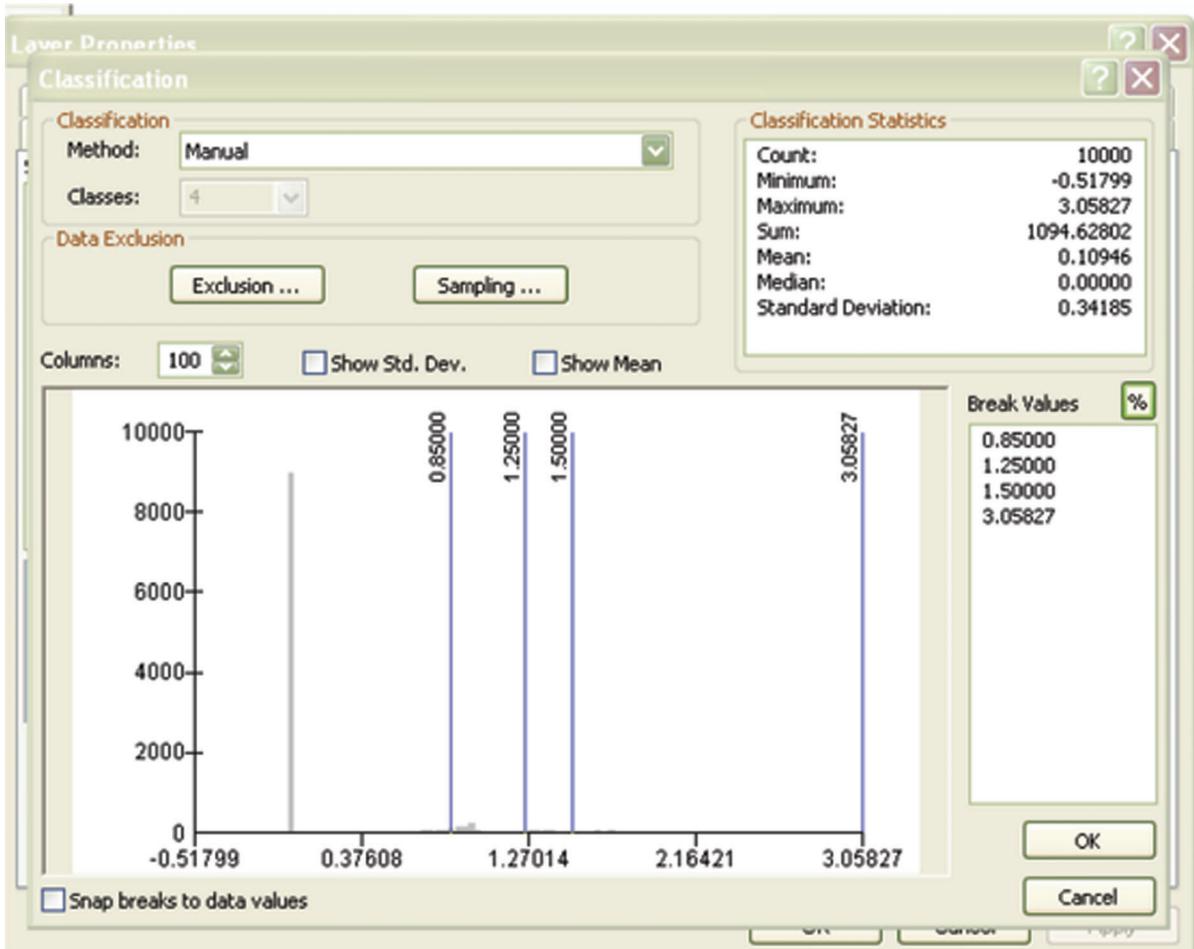


Figure 6.7 Classification settings in symbology.

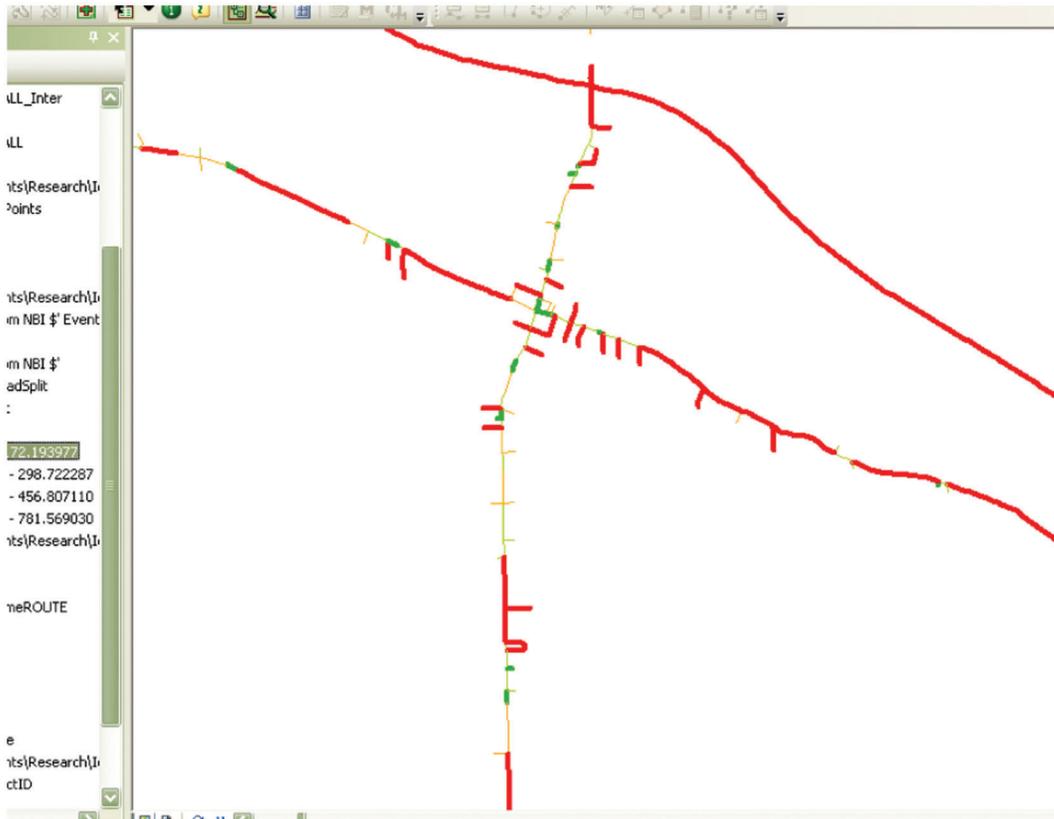


Figure 6.8 Final results of symbology in map.

7. OPTIMIZATION

The SNIP2 Optimization module selects the best combination of roads with safety needs and relevant countermeasures that produce the highest safety benefit within a limited budget. A *Catalog of Countermeasures* includes the needed inputs that represent the safety countermeasures to be considered in the optimization with corresponding selection criteria. In addition, the catalog also provides the unit cost of the countermeasures and their safety effectiveness expressed through the CMFs applied to the target crashes.

The first phase requires creation of the catalog of countermeasures and running it in a batch mode to prepare a set of roads that experience conditions that justify the countermeasure. In the second phase, the SNIP2 optimizer selects the best subset of roads from this initial set in the following three types of analysis:

1. *Safety plan optimization* selects the best combination of relevant safety countermeasures by maximizing the safety benefits.
2. *Safety plan cost prediction* predicts the lowest cost required to save the user-specified number of crashes in the study area.
3. *Safety plan evaluation* predicts the overall cost of a specific safety plan consisting of well-defined safety countermeasures.

Depending on the type of analysis, the user may need to specify the total budget and other constraints. The

matrix of conflicting countermeasures is needed regardless of the type of analysis. These three analyses are illustrated with an example in chapter Example: Safety Plan Optimization.

The following sections first explain the creation of a catalog of countermeasures and its use in a batch mode; then, the settings and execution of the optimization are discussed.

7.1 Catalog of Countermeasures

This feature is used to prepare catalogs of countermeasures that are collections of query definitions (or road and crash conditions associated with a countermeasure). In addition, each query is supplemented with a set of crash modification factors for different crash severities, and a unit cost of the associated countermeasure.

Catalogs can be created by going to the *Catalog of Countermeasures* tab, the first tab in the second tab row (Figure 7) under the *Optimization* tab. This tab presents a form that allows creating a new catalog or editing an existing one. The user may save or run a loaded or just created catalog. The interface also allows combining different catalogs into a new one. In SNIP2, catalogs of countermeasure files have a *ctl* extension (Figure 7.1).

One line in the catalog of countermeasures corresponds to one query (Figure 7.2). The user may use a number of queries to identify all the roads that seem to

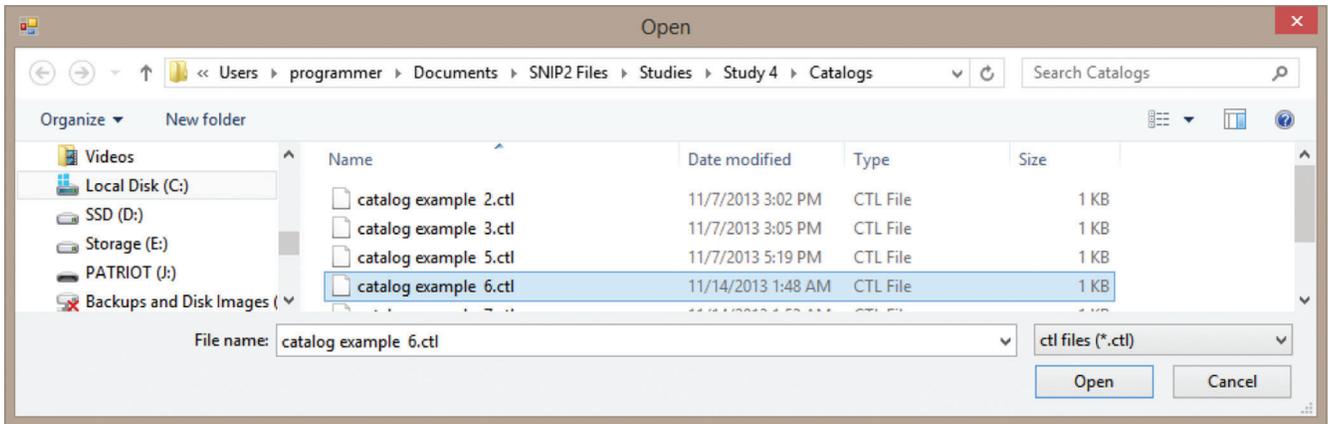


Figure 7.1 Catalog *ctl* files.

share the same type of safety needs that call for the same safety countermeasure. The first column should list the countermeasures. The same countermeasure can be present in multiple lines. The *Description* column should reflect this countermeasure but also should allow identification of the query.

The *Select Query* column opens a file dialog pointing to the query folder of the current study, from where

the user can select any query previously saved. It is important to remember that the user can always point to queries saved in any other folder, including other studies or the *Queries Library* folder. The CMF values for the different severity levels should also be entered, as well as the **annualized** unit cost in the applicable unit (e.g., per mile or per intersection). The user should leave the service life of the countermeasure blank. If the user

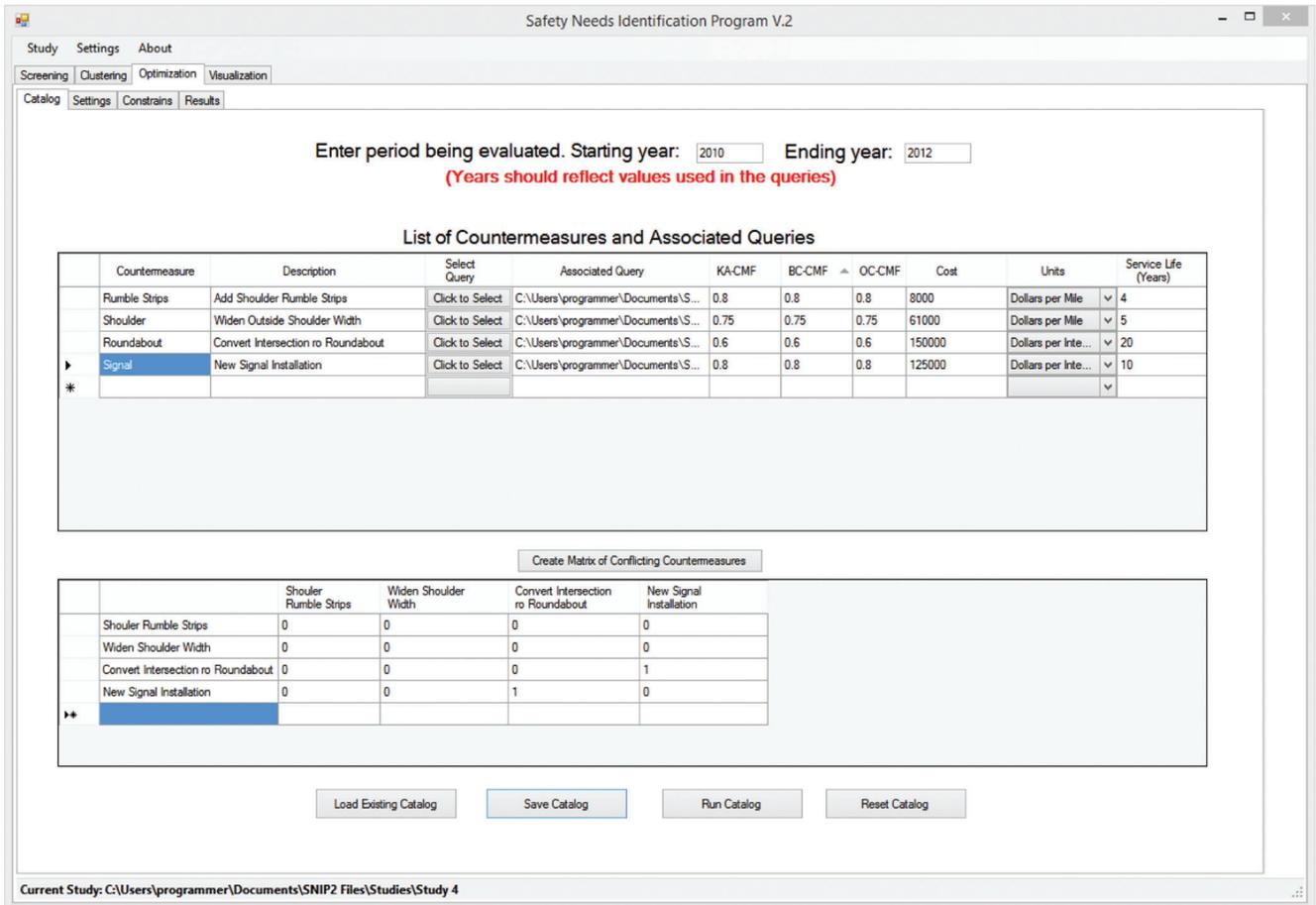


Figure 7.2 Catalog of countermeasures.

decides to enter the total cost of the countermeasure for its entire service life, then the service life expressed in years should be entered in the last column.

New lines are added by entering new names in the *Countermeasure* column. Lines can be deleted by placing the mouse button at the left margin of a line and clicking the “Del” key on the keyboard.

The *Matrix of Conflicting Countermeasures* is created by clicking the *Create Matrix of Conflicting Countermeasures* button. The countermeasures that cannot be implemented at the same location (e.g., convert intersection to roundabout and signalize intersection) are coded in the *Matrix* by typing one in the road and the column of conflicting countermeasures. SNIP2 automatically assures that the matrix stays symmetric.

The *Reset Catalog* button erases all the entered values. The *Save* button opens a file dialog so the user can save the edited catalog to the study folder or to the *Library of Catalogs* folder. When the catalog is saved to the current study, the *ctl* file is stored by default in the current study catalog folder. Multiple catalogs can co-exist in those folders.

In order to run the queries listed in the catalog, two conditions must be met: (1) the catalog must be created or loaded and be displayed on screen; and (2) a study folder must be selected because the results will be

exported to the selected study’s catalog folder. Once these two conditions are met, the catalog can be run by clicking the *Run Catalog* button. At that time, all queries in the catalog are run continuously one by one without the user’s participation.

A popup window is displayed while queries are being executed, and a status bar in red at the bottom of the window displays which query and respective query step are being executed (Figure 7.3). If any error occurs, the user is informed which query and query step caused the error (Figure 7.4) and the catalog continues to run.

The results from each query are exported to the catalog subfolder of the selected study. The results files will have the same names as their queries, but with an extension *csv* and not *qry*. A popup message is also displayed announcing the end of the catalog run (Figure 7.5).

7.2 Running Optimization

The optimization module identifies the best selection of safety countermeasures based on the results obtained from the catalog of countermeasures and the user settings, which include the basic settings and the constraints definition. This part of the optimization module is divided into three basic tabs: settings, constrains, and results. A

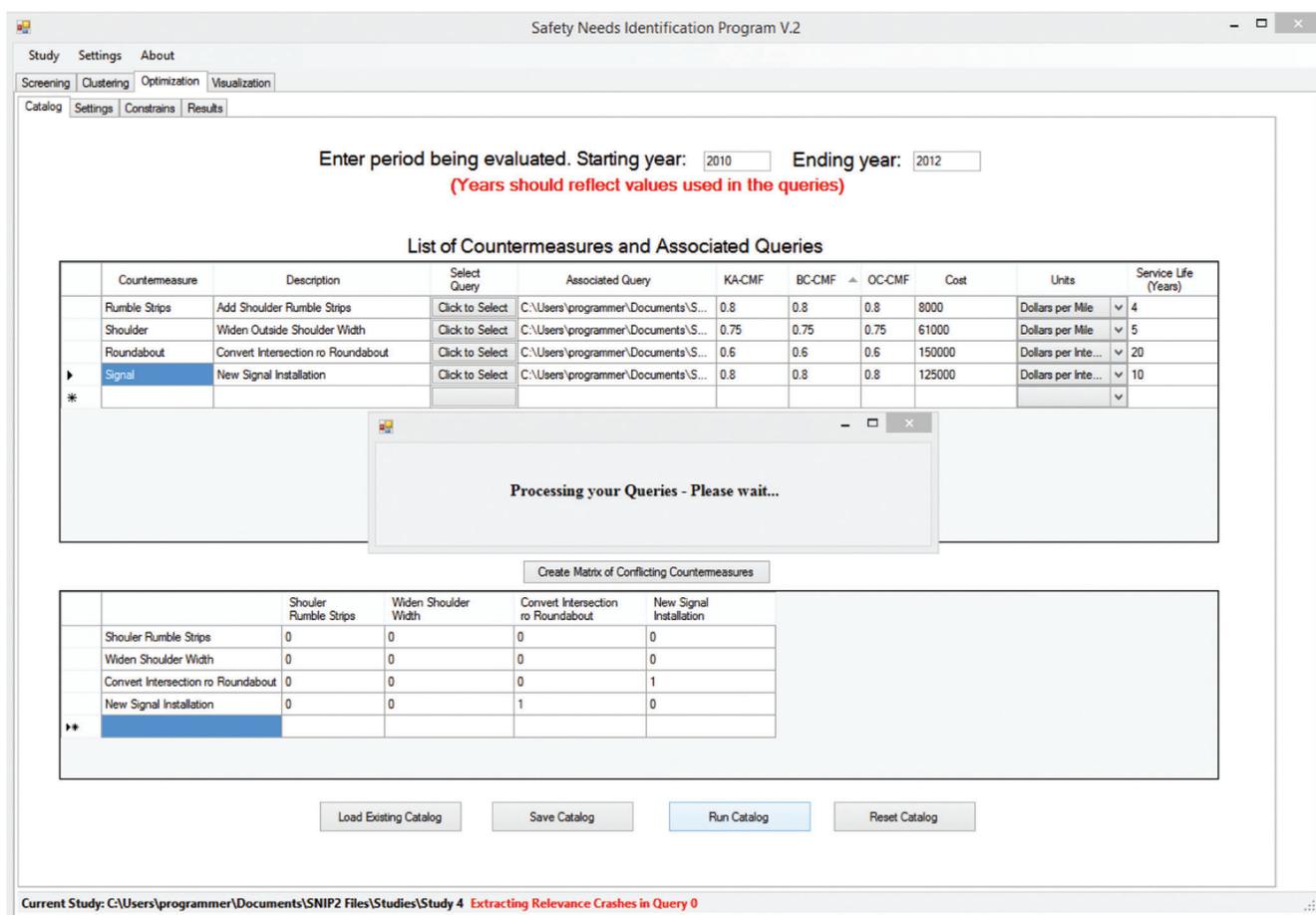


Figure 7.3 Running the catalog of countermeasures.

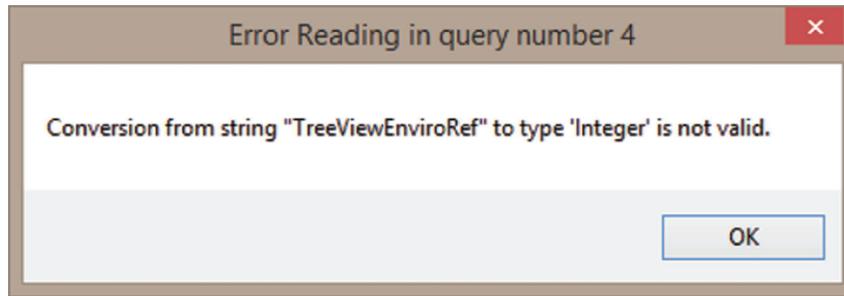


Figure 7.4 Catalog error message.

description of each necessary step to run the process is shown below.

7.2.1 Settings Tab

The Settings Tab allows the user to identify the optimization problem; select the catalog of countermeasures; modify the basic economic information such as interest rate and inflation rate; specify the corresponding years of the data and improvements; and set the exposure metrics and the type of variable in the selection criteria. The user also must define the minimum annual number of crashes for different crash categories and the minimum I or F values to allow road elements to be considered for potential improvements. The user can also save the current optimization problem or load an existing one. Finally, the user is allowed to define the type of analysis to perform and set the target annual number of crashes to save if applicable. (See Figure 7.6.)

7.2.1.1 Selecting Catalog. After running, in batch mode, the catalog of countermeasures, the user must select this catalog for the optimization process. Clicking on the corresponding open folder button shown in Figure 7.7 opens the file selection window where the appropriate file can be selected. Once the catalog is uploaded, the *Constrains* tab shows a list of the existing countermeasures in the catalog. Each of these countermeasures can be disabled or enabled for the optimization run.

7.2.1.2 Saving the Optimization Problem. At any time, the user can save all options in an optimization problem file that can be opened and edited later. Clicking on the *Save* button shown in Figure 7.8 opens the file selection window where the appropriate file can be saved.

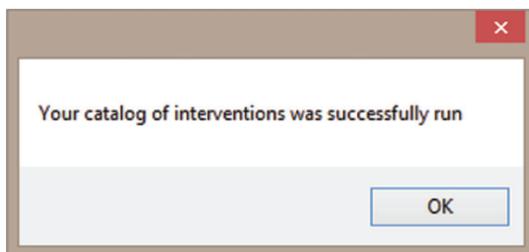


Figure 7.5 Catalog run completion message.

7.2.2 Constraints Tab

After loading the catalog of countermeasures and displaying the countermeasures in the *Constrains* panel (Figure 7.9), the user can define the budget-related constraints:

1. The total budget constraint is a single value that represents the total annual budget. This value is only needed for safety plan optimization and is disabled for other types of analysis.
2. The regional constraints are the maximum budgets for individual INDOT districts. These constraints are optional; and the user may select districts for which these constraints will be applied and then set the maximum budgets for the selected districts. This constraint may also be used to exclude a district from the analysis by setting the maximum budget for that district at zero.
3. The program constraints are the minimum expenditures guaranteed for user-defined safety programs. A safety program is a set of queries selected from one or more countermeasures. As with the regional constraints, the program constraints are also optional so the user may create programs and then corresponding program constraints or remove existing ones.

It is important to note that the mutually exclusive constraint is defined when the catalog of countermeasures is created.

7.2.3 Results Tab

The optimization problem can be executed if the following components are specified:

1. A file with the results of processing a catalog
2. An identification name for the optimization problem
3. The thresholds for including road elements in the optimization
4. The optimization constraints
5. The type of analysis and the target annual number of crashes to save if cost prediction analysis is selected

By clicking on the *Run* button, the optimizer starts its calculations. It might take several minutes to get the results, depending on the size of the problem determined by the type of analysis, the number of road elements included in the optimization, and the number of countermeasures considered.

As shown in Figure 7.10, a summary of the optimization problem is presented in the Results tab.

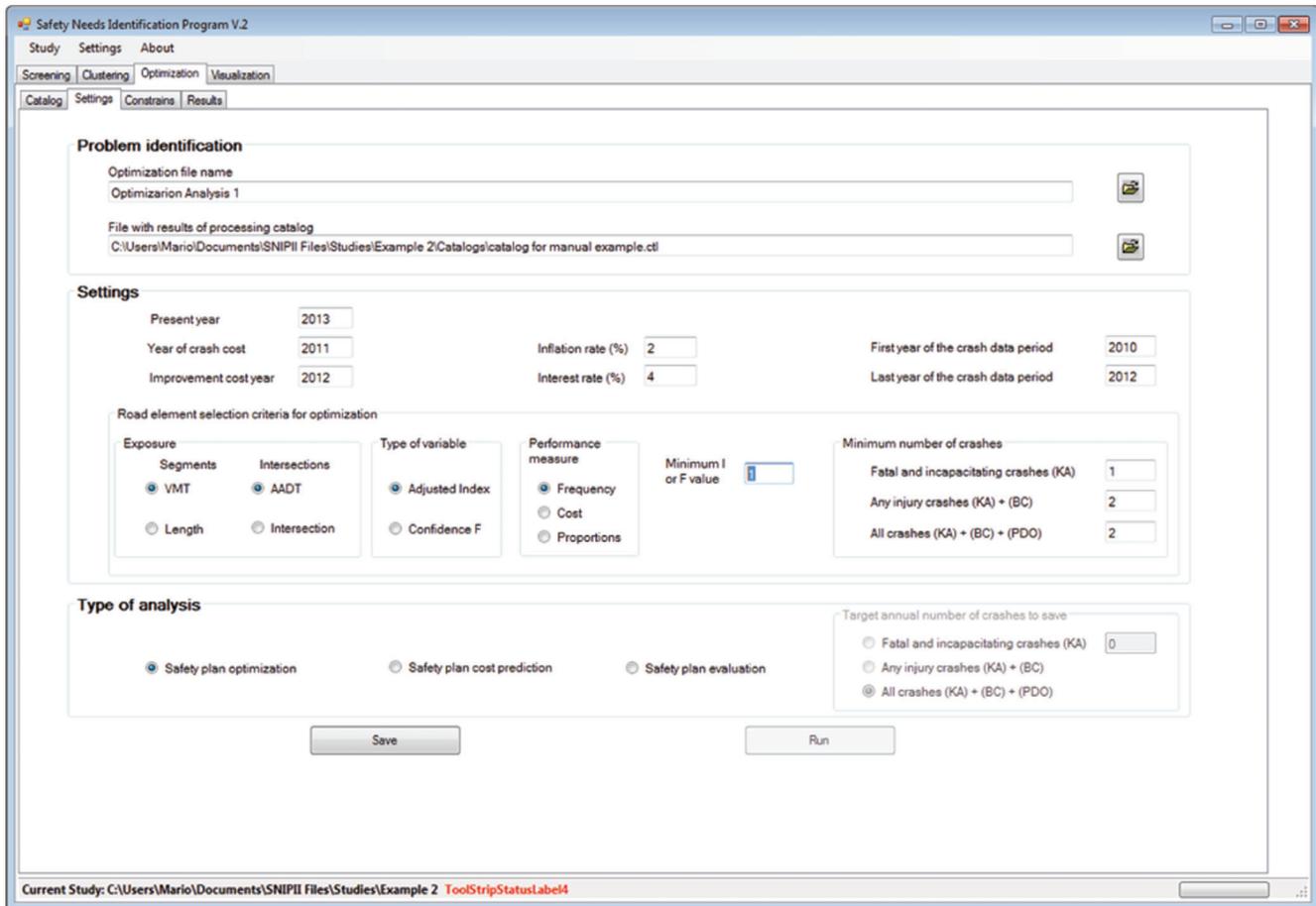


Figure 7.6 Input data.

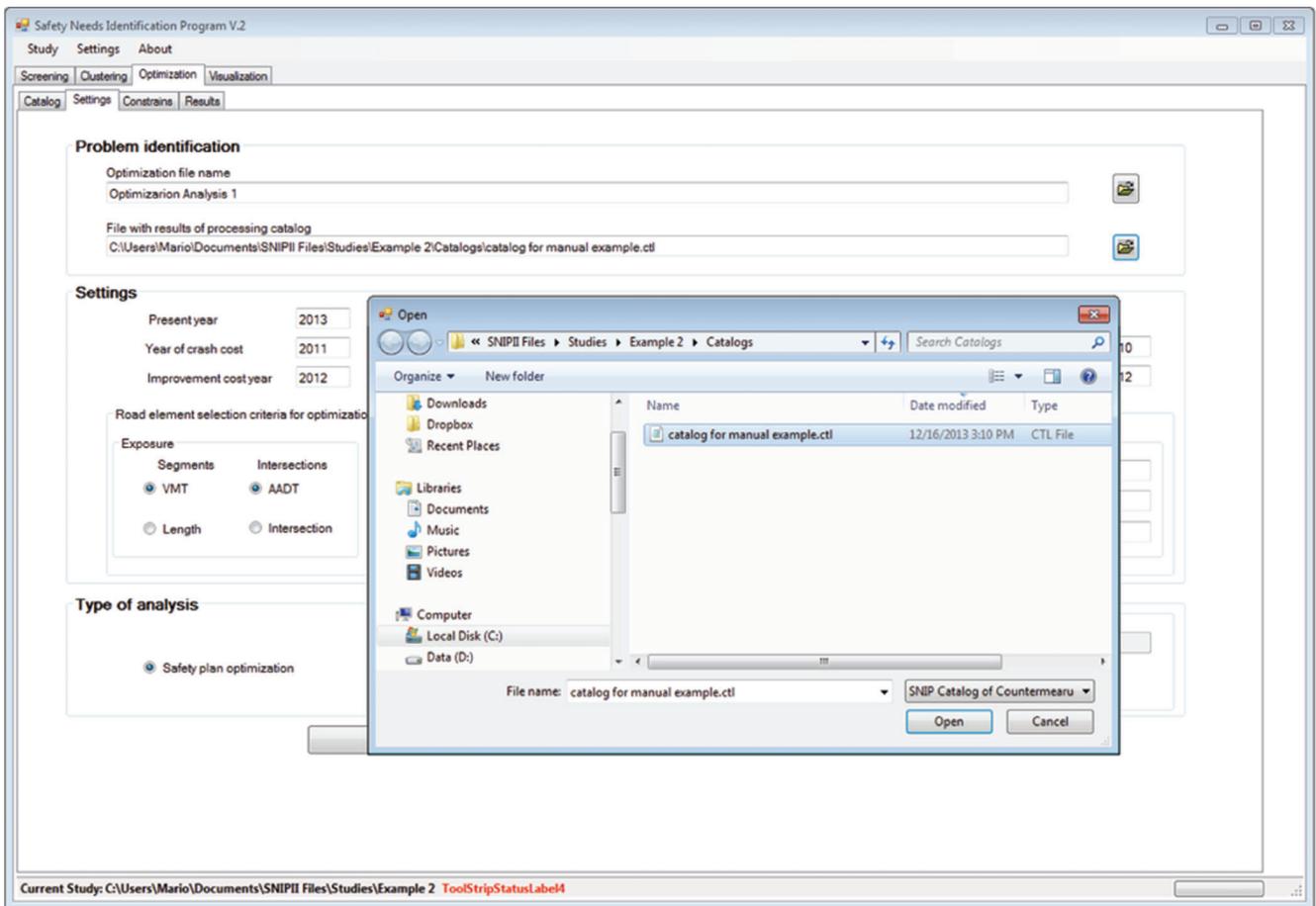


Figure 7.7 Opening the file with the results of the processing catalog.

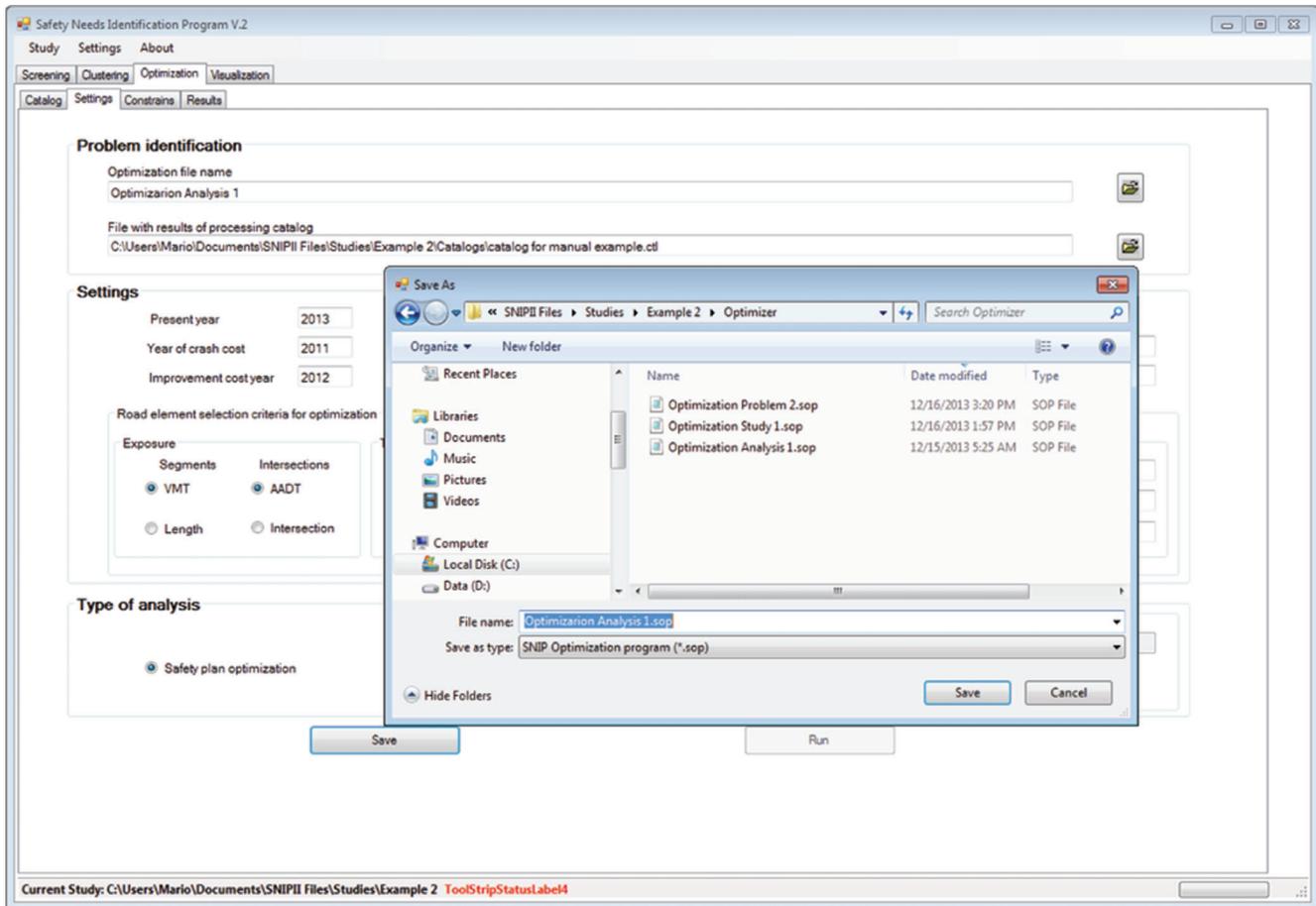


Figure 7.8 Saving the optimization problem input.

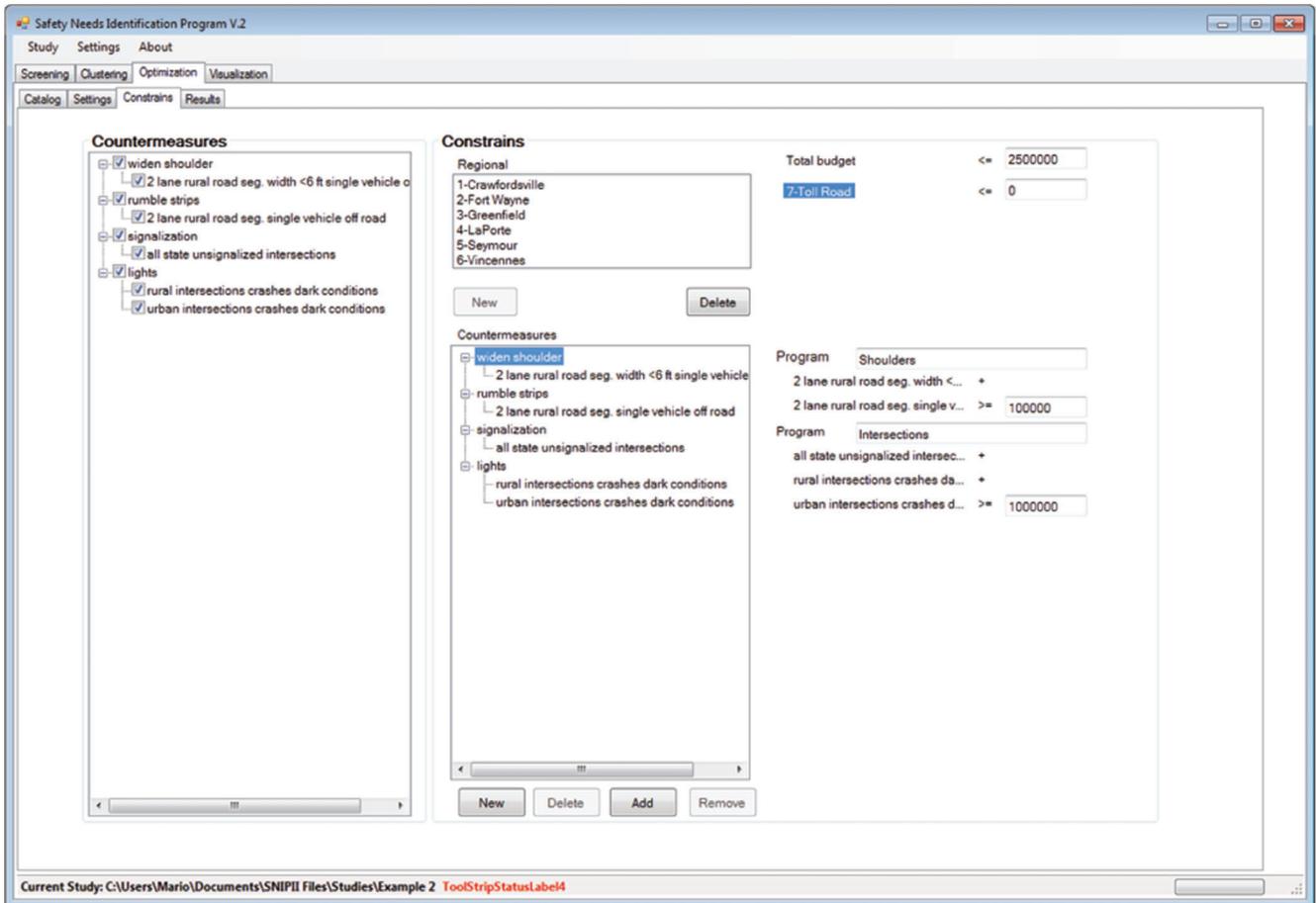


Figure 7.9 Constrains tab.

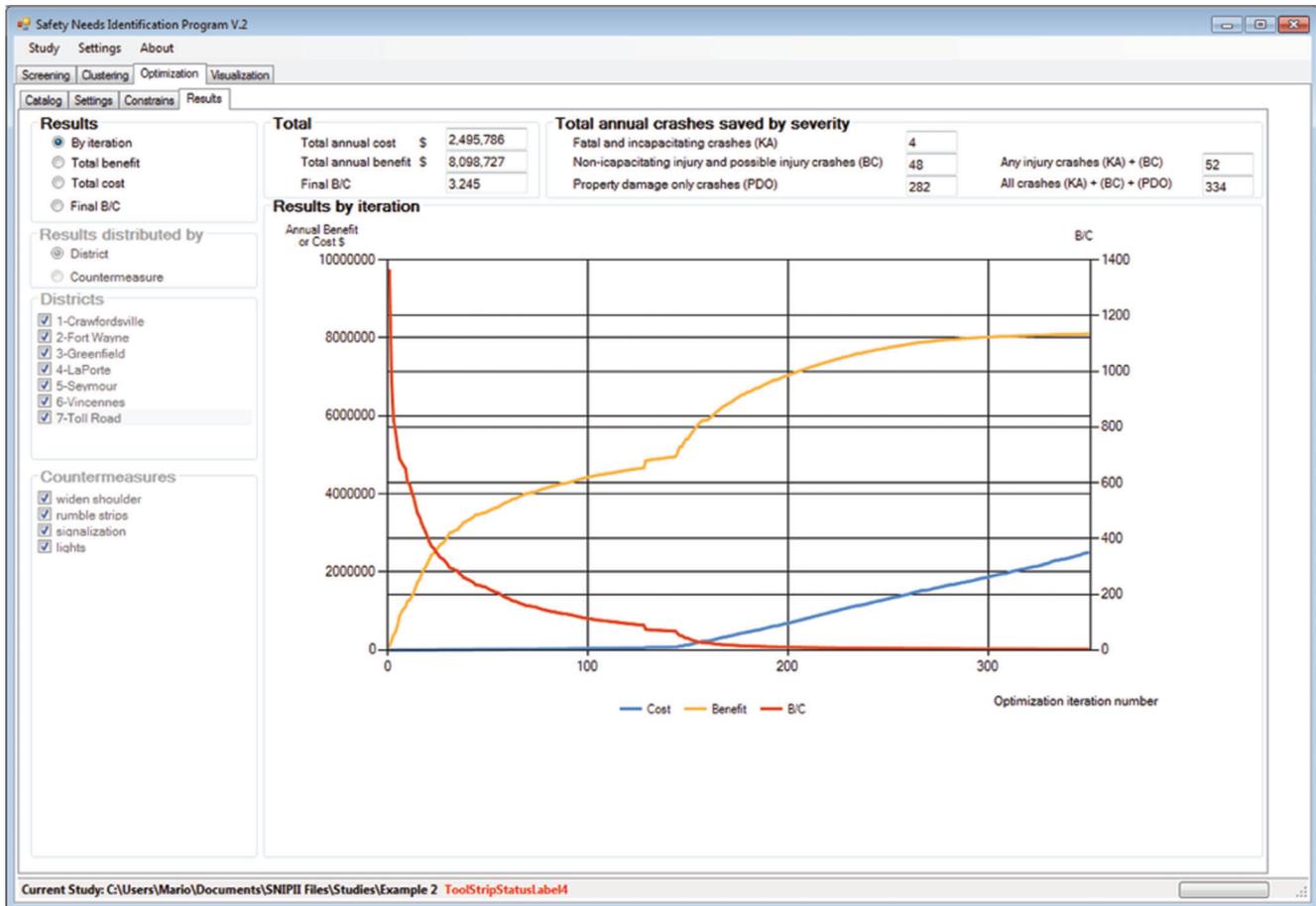


Figure 7.10 Optimization summary: general results.

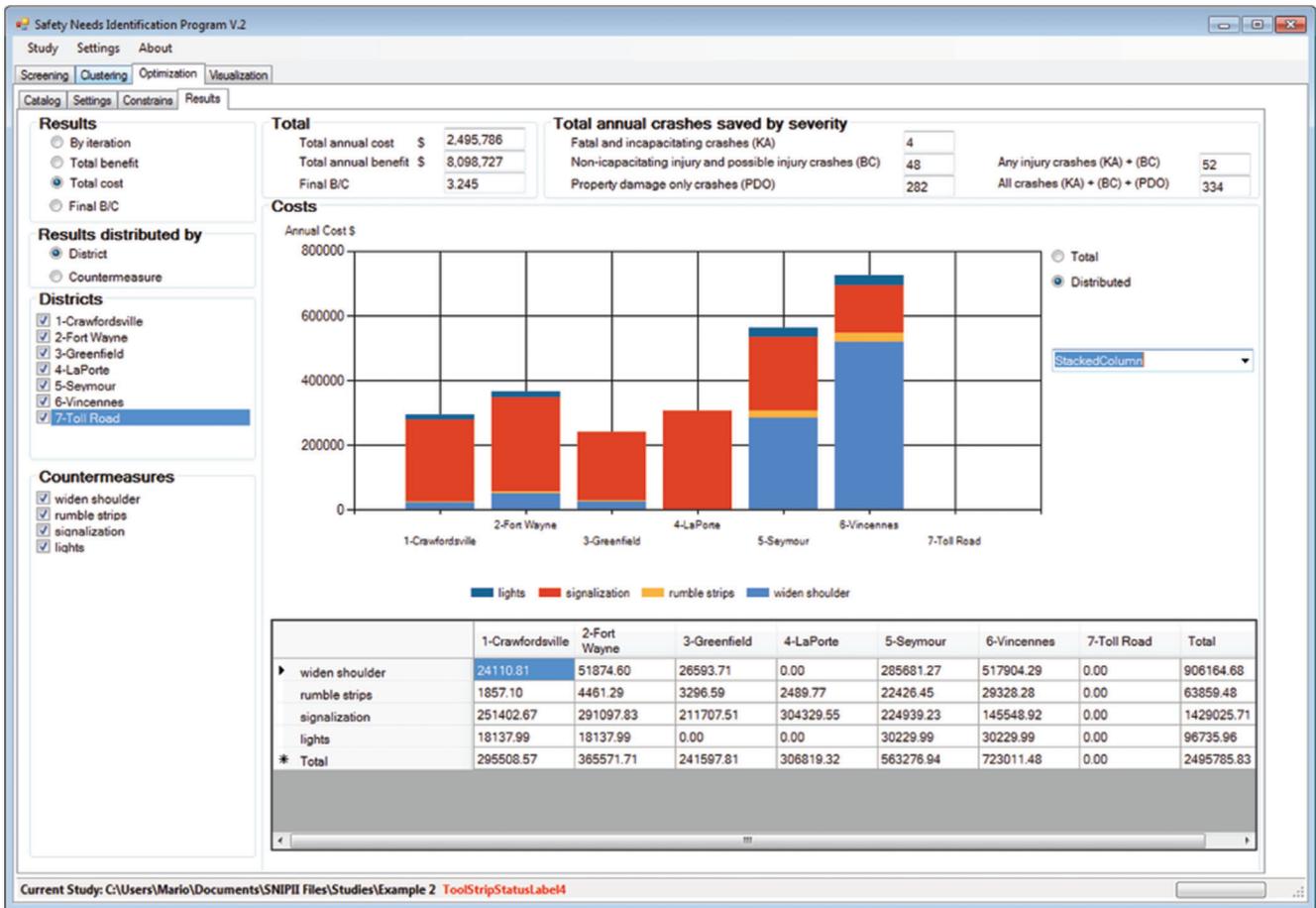


Figure 7.11 Optimization summary: total benefit.

It includes the total annual benefit, the total annual cost, the final B/C ratio, the number of crashes saved per year for different levels of severity aggregation, and a graph representing the results. The horizontal axis represents the number of selected elements; the left-vertical axis represents the monetary costs and benefits; and the right-vertical axis denotes the benefit/cost ratio values.

All results are saved in three csv files: one file is for the optimization results, the second is a log file that contains all the countermeasures not selected due to a violation of one or more constraints, and the third contains a summary by countermeasure and district.

The results displayed by SNIP2 include summary charts and tables for total benefit, total cost, and benefit-cost ratio, grouped by countermeasure or by INDOT district. The user can select the chart type and also a two-way distribution of the results by INDOT district and by countermeasure as shown in Figure 7.11. The user is also allowed to exclude some countermeasures or INDOT districts from the summary charts and tables.

8. EXAMPLE: WIDENING SHOULDERS

This chapter illustrates the features of SNIP2 applied to the first study focused on a single safety countermeasure: widening shoulders of rural roads. It demonstrates the screening, clustering, and visualization features.

8.1 Screening

The SNIP2 screening tool is used to select roads that meet user-specified road criteria, select crashes on those roads that meet user-specified crash criteria, and then calculate the safety performance measures and other statistics for the resulting combination of roads and crashes. The produced tables of roads with corresponding statistics are saved, and these tables are used as an input to the SNIP2 clustering and visualization tools.

In the present example, SNIP2 is used to identify road segments that are suitable for widening shoulders. The first step is to start the SNIP2 program. SNIP2 initially checks the database for any leftover tables from

previous analyses; and if any are found, they are purged in preparation for the new query. Then, the user creates a study folder called *Manual Examples* by clicking the menu option *Study/Create New Study Folder*.

The example query for widening segment shoulders is executed for Indiana (scope) and for the period 2010–2012 (years). Shoulders widening is considered only for rural two-lane roads with shoulders narrower than six feet. To avoid segments crossing populated areas, the non-existence of curbs is also included as part of the element selection criteria.

The crashes that can be linked with the narrow shoulders are single-vehicle and run-off-road crashes. The crash rates are calculated using the segment VMT. To calculate the crash proportions, the number of single-vehicle and run-off-road (*relevance crashes*) are related to the number of all crashes on the segment (*reference crashes*). The available Crash Modification Factors (CMF) for shoulder widening apply to all crashes on the segment (*target crashes*).

To enter the above query conditions, the user opens the Screening tab and then the Scope/Element tab in the second row of tabs (Figure 8.1).

Figure 8.1 shows the selected options for years, scope, infrastructure elements, and exposure that reflect the desired input for the widening shoulders query.

The next step is to open the Road Elements tab and selecting Rural Two Lane in the *Segment Type* group and ranges: <2 ft, 2–4 ft, and 4–6 ft in the *Shoulder Width* group. (Figure 8.2.)

Clicking the *Extract Element Data* button extracts the relevant segments. The process is complete when the processing query message disappears. To see the selected segments, the user must click the *View Extracted Elements* button. A table with segments that match the criteria is displayed.

Once the proper segments are extracted, the user should extract relevance crashes that match the crash criteria. This can be done by clicking the Relevance Crashes tab and selecting option 1 in the Number of Vehicles Involved group and t 1 and the Ran Off-Road option in the Manner of Collision group (Figure 8.3). Once the proper selections are made, the user should click the *Extract Relevance Crash Data* button. The relevance crashes are extracted when the processing message disappears.

Next, the user should open the Reference Crashes tab. By default, the selections made in the Relevance Crashes tab are repeated in the Reference Crashes tab. Since the reference crashes are all crashes, the user should clear any selection in the Reference Crashes tab (Figure 8.4) and then, click the *Extract Reference Crashes* button. After the extraction process, the user has the choice viewing the extracted crashes or moving to the Target Crashes tab. By default, all the crash options in the Target Crashes tab are blank, which corresponds to selecting all the crashes. There is no need to change anything in this tab because the CMFs apply to all the crashes. The user should now press the *Extract Target Crashes* button to process the extraction.

Once the segments and the crashes are extracted, the user can execute the calculation of the crash frequencies, rates, proportions, costs, and other statistics for the selected segments. The user should go to the Results tab and click the *Start Screening Calculations* button. It will take a few minutes to create several intermediate tables. The user will see the pop-up message that the processing is taking place. Once the message disappears, the final results are displayed on screen (Figure 8.5).

The table with the results should be exported as a comma-separated-values (csv) file by clicking the *Export Results* button that is located next to the *Start*

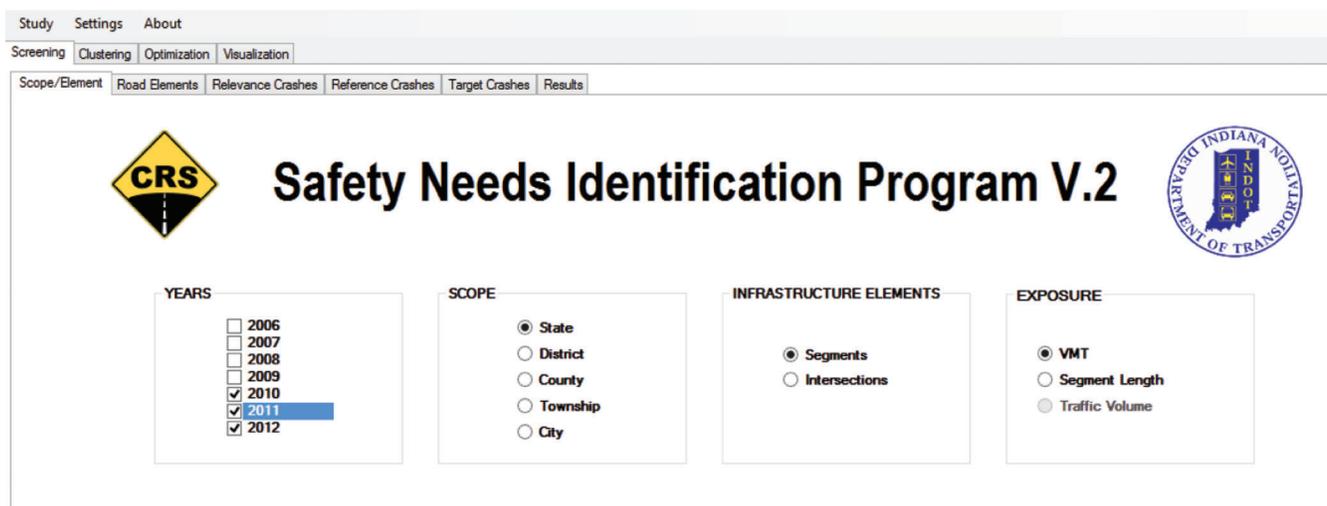


Figure 8.1 Beginning the screening process.

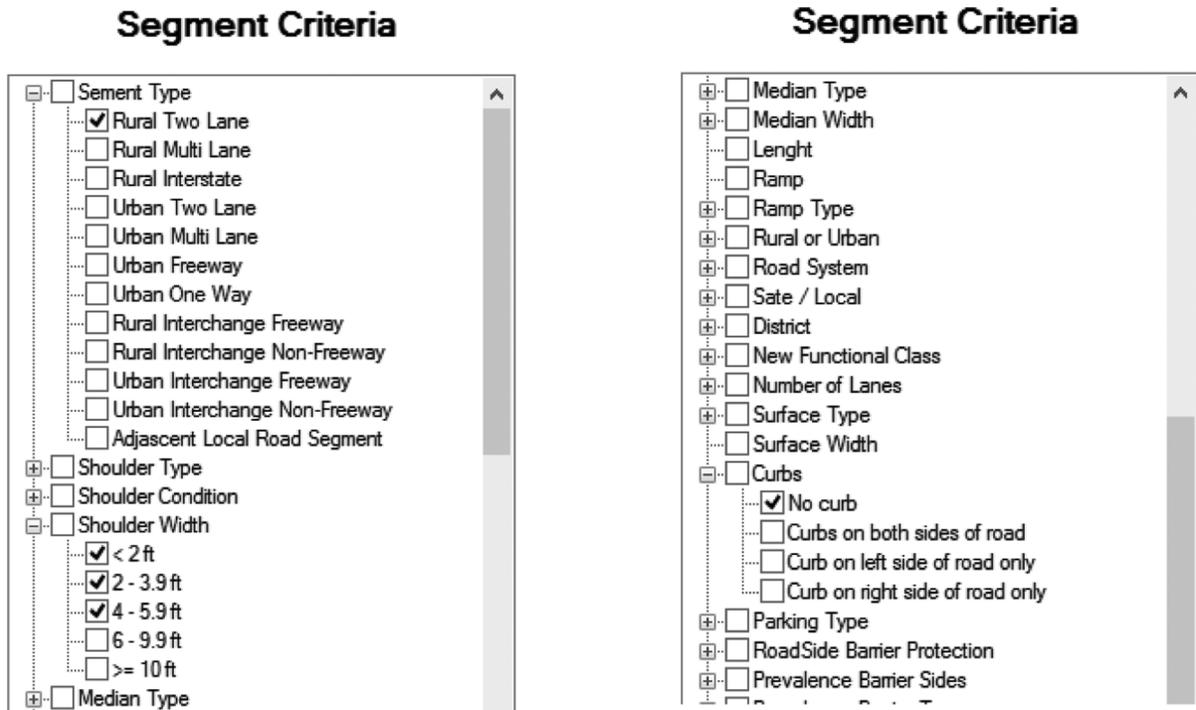


Figure 8.2 Segment type = rural two-lane and shoulder width ≤ 6 ft and with no curb.

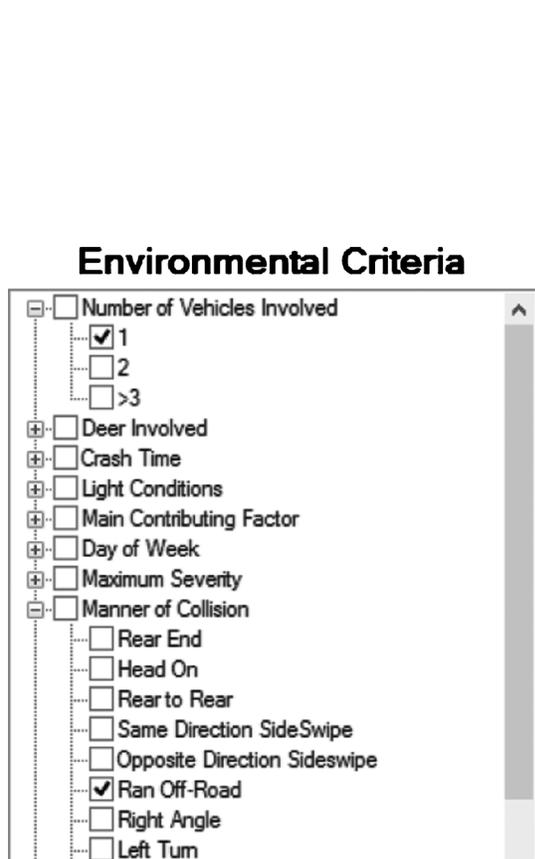


Figure 8.3 Choosing the relevance crashes.

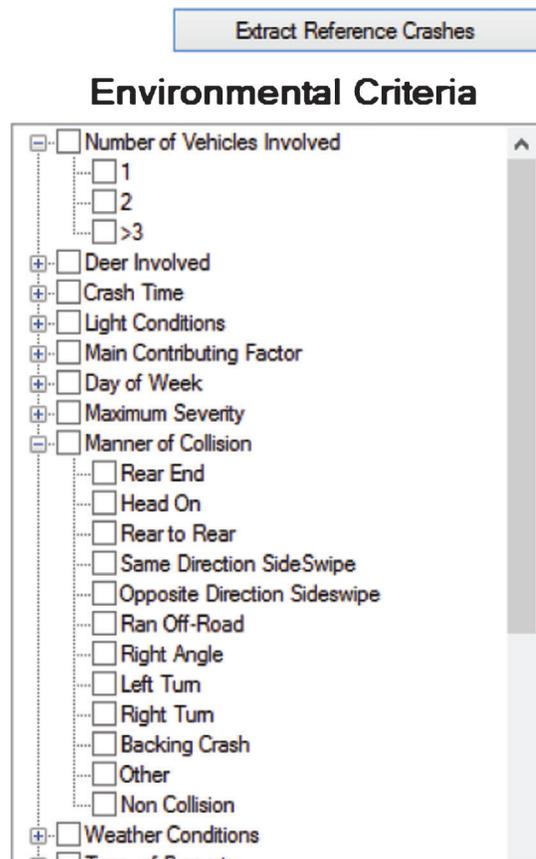


Figure 8.4 Selection tree for the reference and target crashes.

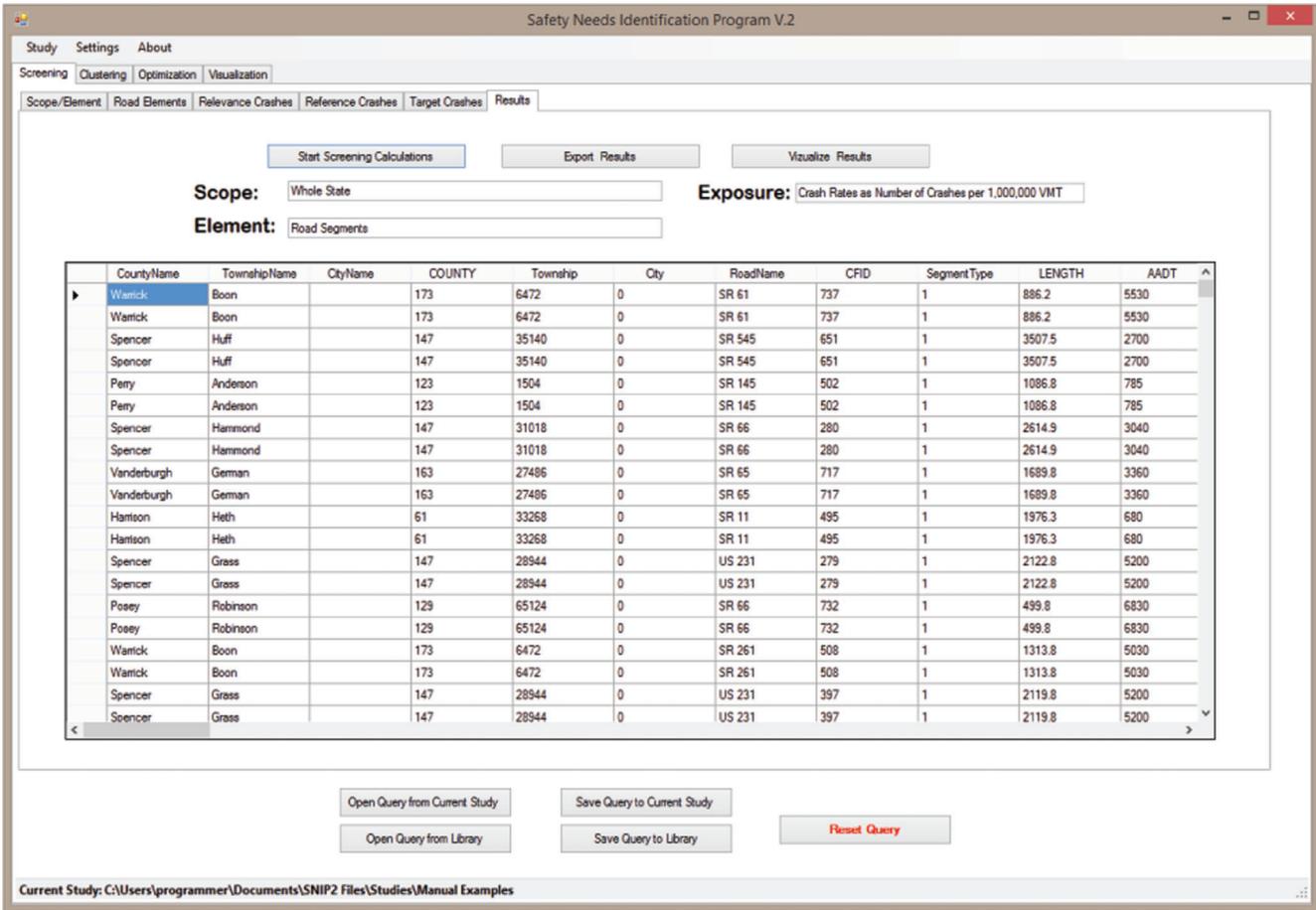


Figure 8.5 Final results table.

Screening Calculations button. When the button is clicked, the file is saved by default to the current study query folder; in our case, it the *Manual Examples/Queries* folder. (See Figure 8.6.)

Some successfully run queries might be needed for future use, such as inclusion in a catalog of countermeasures. To save such a query, the user must click the

Save Query to Current Study button at the bottom of any of the tabs under the Screening tab. It opens a file save dialog that by default prompts the current study folder (*Manual Examples/Queries* folder). It should be noted that although the results of a query (table) is exported as a csv file, the query definition is saved in a file with extension *qry* (Figure 8.7.) It is important to

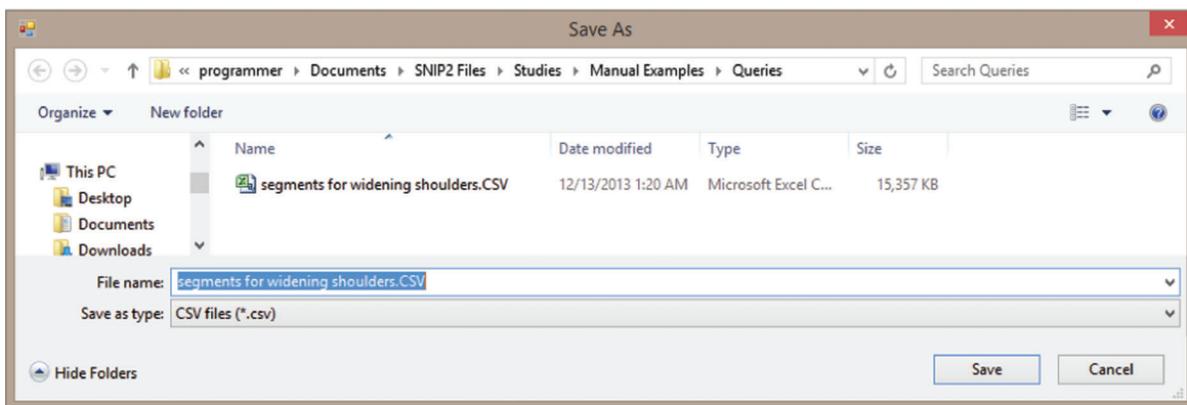


Figure 8.6 Exporting results to current study query folder.

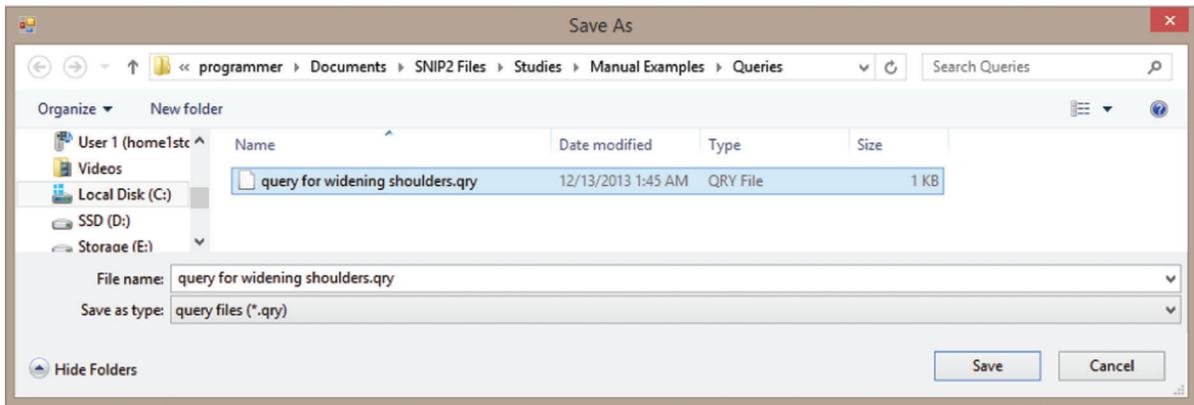


Figure 8.7 Saving query definition to current study query folder.

remember that a query definition can be saved at any time, even before all the criteria are selected in the Screening tab, which means that an incomplete query will be saved and then restored when opened. This feature is useful when the user wants to stop and resume later. Only complete queries should be included in a catalog of countermeasures. Otherwise, processing

of the catalog for countermeasures will yield incorrect results.

8.2 Clustering

The initial values needed by the clustering process are shown in Figure 8.8. The required parameters are set as

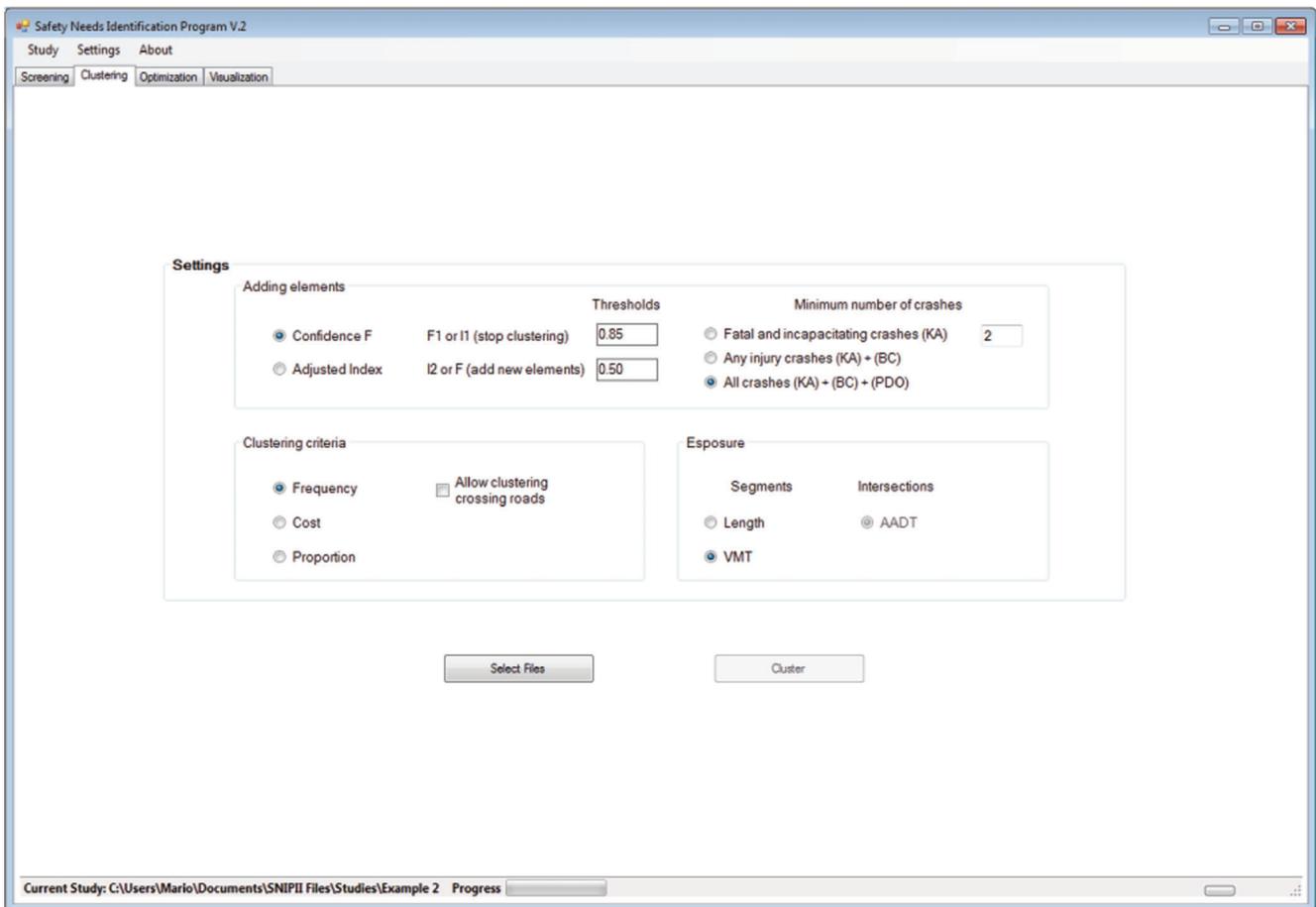


Figure 8.8 Selecting files for clustering.

follows: The *Confidence F* option for controlling adding elements; $F_1=0.85$; $F=0.5$. The *Frequency* option is selected as the clustering criterion and the *Allow cluster to switch roads* option is disabled. The *VMT* option is selected as exposure for segments. The type of crashes is set as all crashes and the minimum number of crashes is set as two.

Then, the file with the query results for widening shoulders of two-lane rural segments from the screening tool is selected to be processed as shown in Figure 8.9.

Once the data import is completed, the clustering process can be run. At the end of the clustering process, a *Summary* window appears with a summary report showing the total number of clusters found and the total segments included (Figure 8.10).

8.3 Visualization Using KML files

The initial values used by the visualization tool are set at their default values as shown in Figure 8.11.

For the purposes of the example, both the initial screening results and the clustering results were selected to create the KML files. Figure 8.12 shows the screening results of the high-crash two-lane rural segments with shoulders less than six foot wide (narrow shoulders) in Indiana.

Figure 8.13 shows the high-crash two-lane rural segments with narrow shoulders in Perry County and Figure 8.14 shows the same KML near Derby, IN in Perry County. Figure 8.15 presents the clusters obtained for this area.

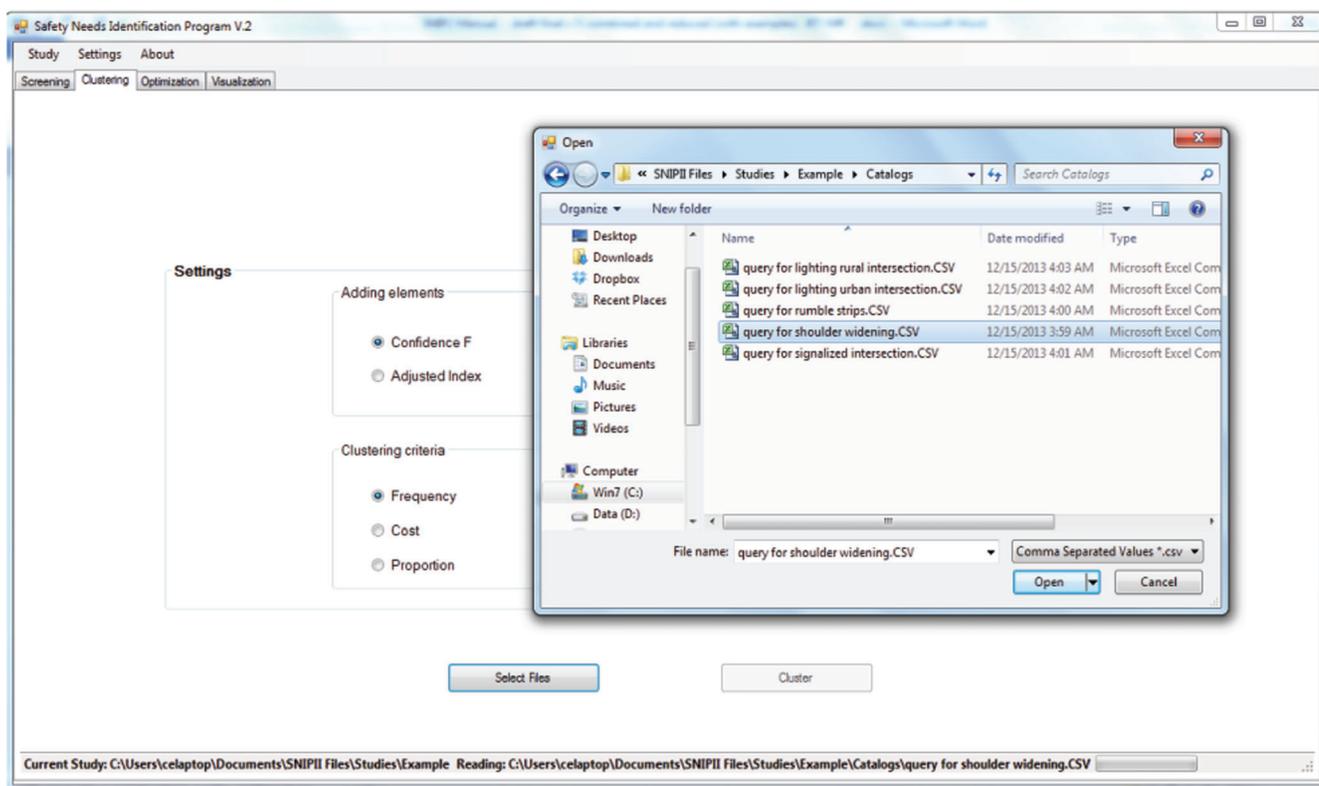


Figure 8.9 Clustering selected files.

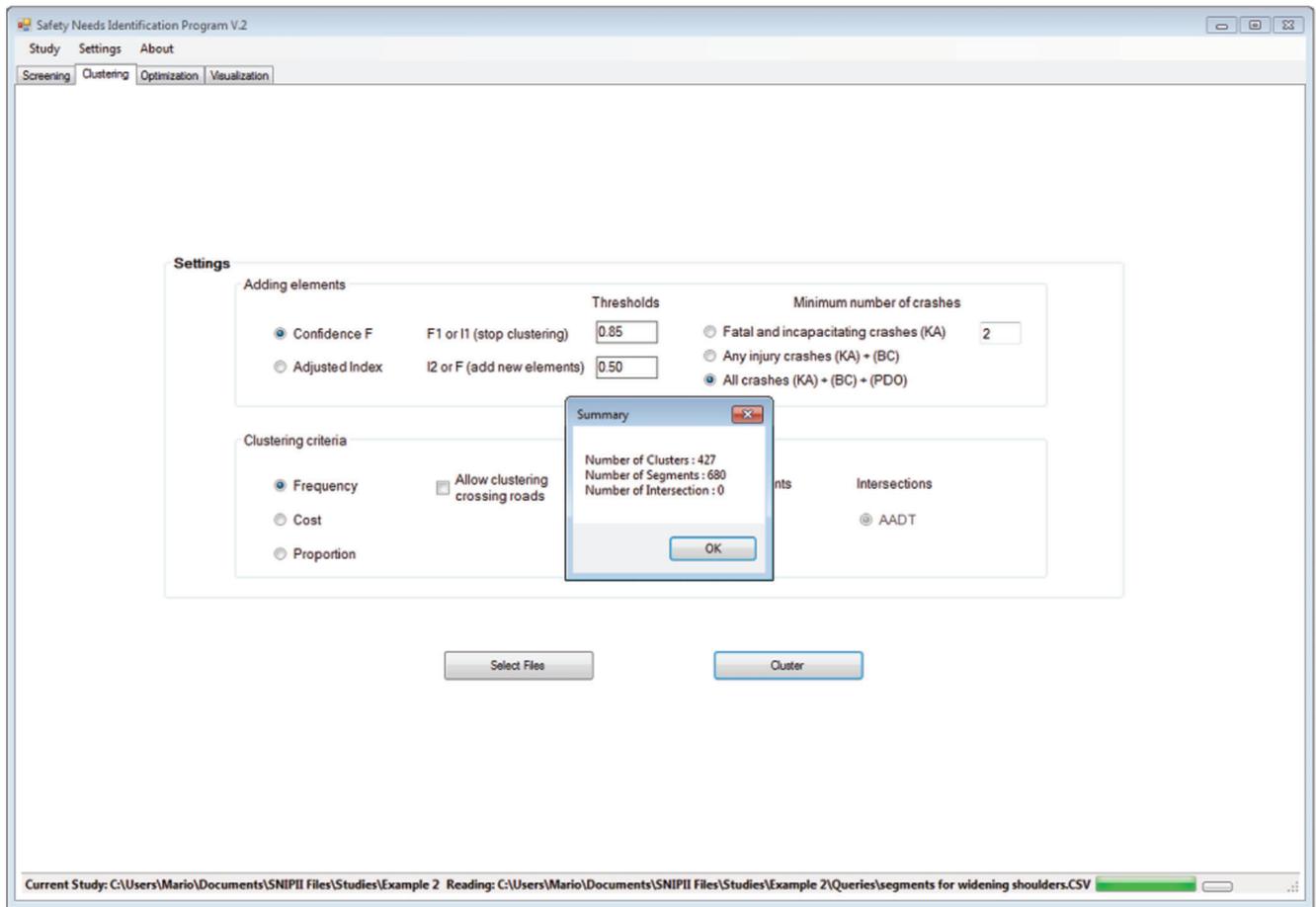


Figure 8.10 Clustering summary.

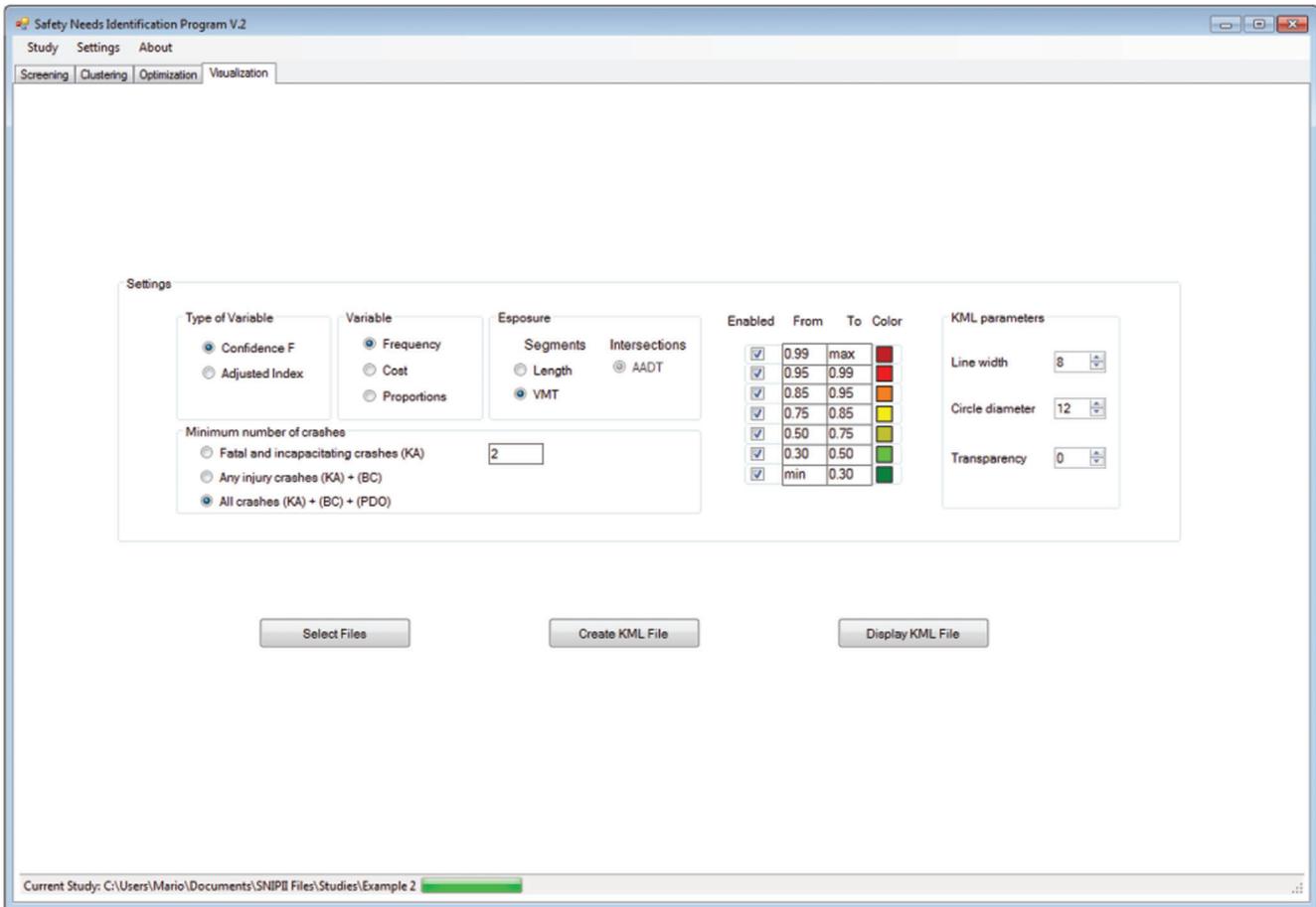


Figure 8.11 Visualization tool settings.

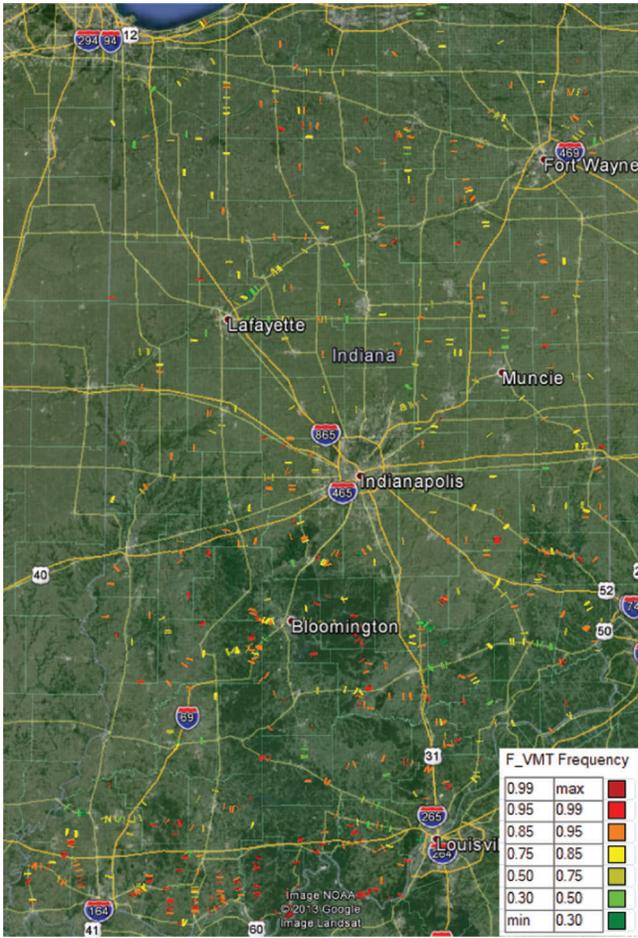


Figure 8.12 High-crash rural two-lane segments with narrow shoulders in Indiana.

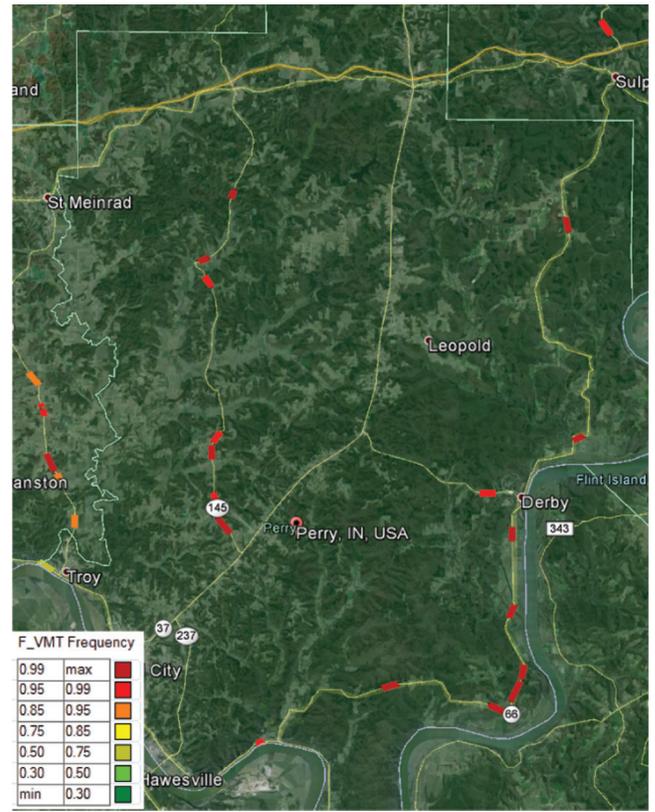


Figure 8.13 Brown County: High-crash rural two-lane segments with narrow shoulders.

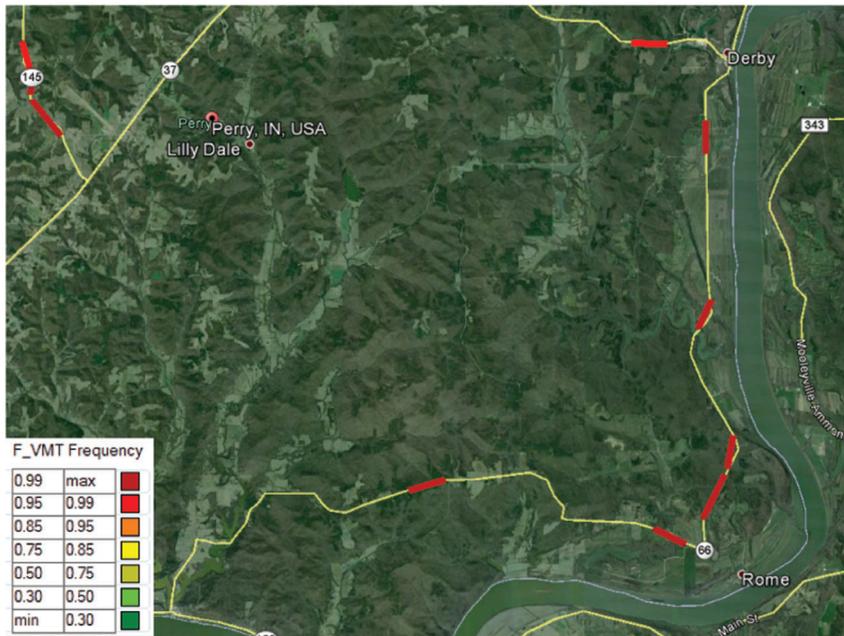


Figure 8.14 High-crash rural two-lane segments with narrow shoulders on SR 135 near Gnaw Bone, Indiana.

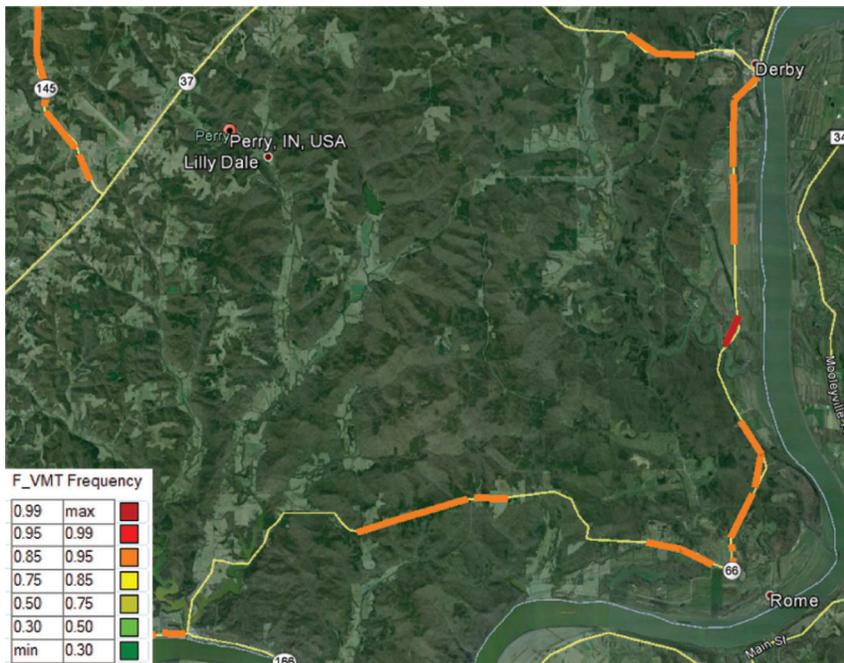


Figure 8.15 Clusters of high-crash rural two-lane segments with narrow shoulders on SR 135 near Gnaw Bone, Indiana.

9. EXAMPLE: SAFETY PLAN OPTIMIZATION

The second example is focused on optimizing a safety plan that includes four safety countermeasures: widening shoulders and installing shoulder rumble strips on segments and signalization and lighting at road intersections. The road screening for four countermeasures is presented. Then, the four countermeasures (and corresponding five queries) are placed in a catalog of countermeasures to run three different types of analyses aimed to optimize or evaluate the example safety plan for Indiana.

9.1 Screening

In this example, SNIP2 is used to identify roads (segments and intersections) that are suitable for four safety countermeasures: widen shoulders on road segments, shoulder rumble strips on segments, traffic signals at intersections, and lighting at intersections.

The first step is to start the SNIP2 program. SNIP2 checks the database for any left-over tables from previous analyses. If any are found, they are purged in preparation for the new query.

Then, the user creates a study folder called *Manual Examples* by clicking the menu option *Study/Create New Study Folder*. First, the user runs the screening for shoulder widening, and then for the remaining countermeasures, one by one.

This and the other example queries are made for Indiana (scope) and for the period 2010–2012 (years). Shoulder widening is considered only for rural two-lane roads with shoulders narrower than six feet. The crashes that can be linked with the narrow shoulders are single-vehicle and run-off-road crashes. The crash rates are calculated using the segment VMT. To calculate the crash proportions, the number of single-vehicle and run-off-road (*relevance crashes*) are related to the number of all crashes on the segment (*reference crashes*). The available CMFs for shoulder widening apply to all crashes on the segment (*target crashes*).

9.1.1 Shoulder Widening

Details of screening the Indiana road segments for widening shoulder are provided in chapter Example: Widening Shoulders. The reader is requested to refer to that chapter. Additional comments and explanation about screening for the other three countermeasures are provided in this chapter.

9.1.2 Shoulder Rumble Strips

Once the query of road segments for shoulder widening is executed and saved, the user should repeat similar steps to extract segments for shoulder rumble strips.

Similar query conditions apply to this countermeasure except the type of segments; all two-lane rural road segments without curbs should be considered regardless of the shoulder width. The assumption is made for this example that Indiana state roads have shoulders

sufficiently wide for installing rumble strips and to avoid segments crossing populated areas; particularly after implementing the shoulder-widening countermeasure. The user may utilize the previous query by reading it, modifying, and saving under a different name. The modification is limited to removing the condition for shoulder width under or equal to 6 ft (Figure 9.1); other conditions are applicable without changes.

At the end of the process, another results table is created. This file should also be exported as a csv file using the *Export Results* button. The query definition should also be saved. Since the current study folder did not change, the new results file and the query definition will both be saved in the *Manual Examples/Queries* folder. (See Figure 9.2.)

9.1.3 Intersection Signalization and Lighting

The specific conditions for intersections that may benefit from traffic signalization include the lack of signalization. Other conditions include: 2010–2012 period, entire Indiana state road system, reference and target crashes are all crashes, and traffic volume used as the measure of exposure.

The steps for the analyses of intersections are identical to the ones executed with the analyses of segments, except that the infrastructure element chosen is *Intersections* and the only measure of exposure available is *Volume of Traffic* (Figure 9.3). In addition, once intersections are selected as the infrastructure elements, the *Intersection Criteria Selection Tree* becomes active in the Road Elements tab (Figure 9.4), while the Segment Criteria become inactive.

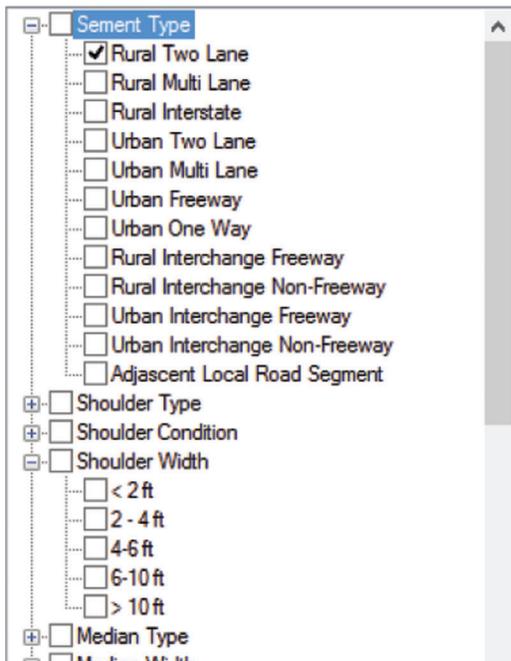
The conditions for intersections that may benefit from installation of lighting include the lack of lighting and the excessive number of crashes during nighttime hours. Although all intersections are considered, screening is done separately for urban and rural intersections with two queries (Figure 9.5). This separation is needed because the cost of adding lighting to rural intersections is on average different from the cost of adding lighting to urban intersections. The two queries are represented by two rows in a catalog of countermeasures, which allow entering two different countermeasure costs. For the signalization countermeasure, the element extraction should be confined to intersections without signals, and the relevance crashes criteria focus on those which occur at right angles as seen in Figure 9.4.

As with the previous example, after each screening is completed, the results should be exported to csv files and the query definitions to *qry* files for future use in a catalog of countermeasures.

Five files with saved query definitions should be available at the end of the screening (Figure 9.6):

1. two-lane rural road segments considered for widening shoulders
2. two-lane rural road segments considered for rumble strips
3. intersections considered for signalization
4. rural intersections considered for lighting
5. urban intersections considered for lighting

Segment Criteria



Segment Criteria

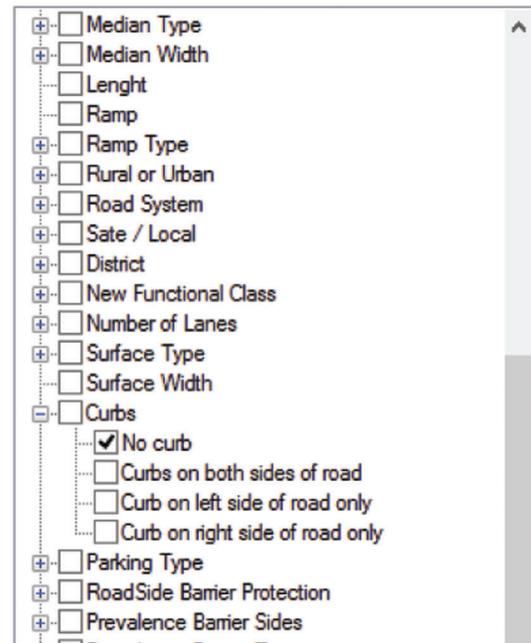


Figure 9.1 Segment criteria for adding rumble strips.

9.2 Optimization

A two-step process is required to select the best combination of roads with safety needs and relevant countermeasures that produce the highest safety benefit. A catalog of countermeasures is created and the optimizer is run.

9.2.1 Catalog of Countermeasures

A catalog of countermeasure is a collection of saved queries; each query is associated with a safety countermeasure, the CMFs for different severities, and the unit cost of a countermeasure.

To create a catalog, the user should go to the Catalog tab, the first tab under the Optimization tab (Figure 9.7). Figure 63 displays the example catalog of the four countermeasures. To generate this catalog, the user must enter the name of the countermeasure in the first column, followed by a brief description of the query that helps identify it and associate it with the countermeasure. Clicking the *Select Query* button in the catalog opens the “open file” dialog that displays the queries saved in the previous steps (Figure 9.8.) The user selects the correct query. Then, the values for the CMFs and the unit costs of the countermeasure should be entered. Finally, the user enters the service life for the countermeasure, in years, in the last column.

Once all the countermeasures are added, the user clicks the *Create Matrix of Conflicting Countermeasures*

button to display a matrix filled with zero values. If a certain pair of countermeasures cannot be applied to the same road, then value 1 should be entered in the cell corresponding to the two countermeasures. The user can enter 1 in the lower or upper part of the matrix. SNIP 2 will fill the other cell of the matrix automatically. In the example, none of the countermeasures are conflicting and all the cells have value 0.

At this point, the user should save the catalog by clicking the *Save Catalog* button. The catalog file will be saved in the *Manual Examples/Catalogs* folder as the current one (Figure 9.9). The catalog saved can be opened at any time by clicking the *Load Existing Catalog* button.

Clicking the *Run Catalog* button executes all the queries in the catalog one by one in a batch mode without involvement of the user. The results of running each query are automatically exported to a csv query results file in the *Catalog* folder. The query results file has the same name as the associated query. In addition, a single csv query names file is generated that includes all the names of the query results files. The query names file is later read by the SNIP2 optimizer to locate all the query results files whose contents are used in the optimization.

9.2.2 Running the Optimizer

Processing the catalog of countermeasures produces a set of roads that experience conditions that justify the countermeasures included in the catalog. In the second

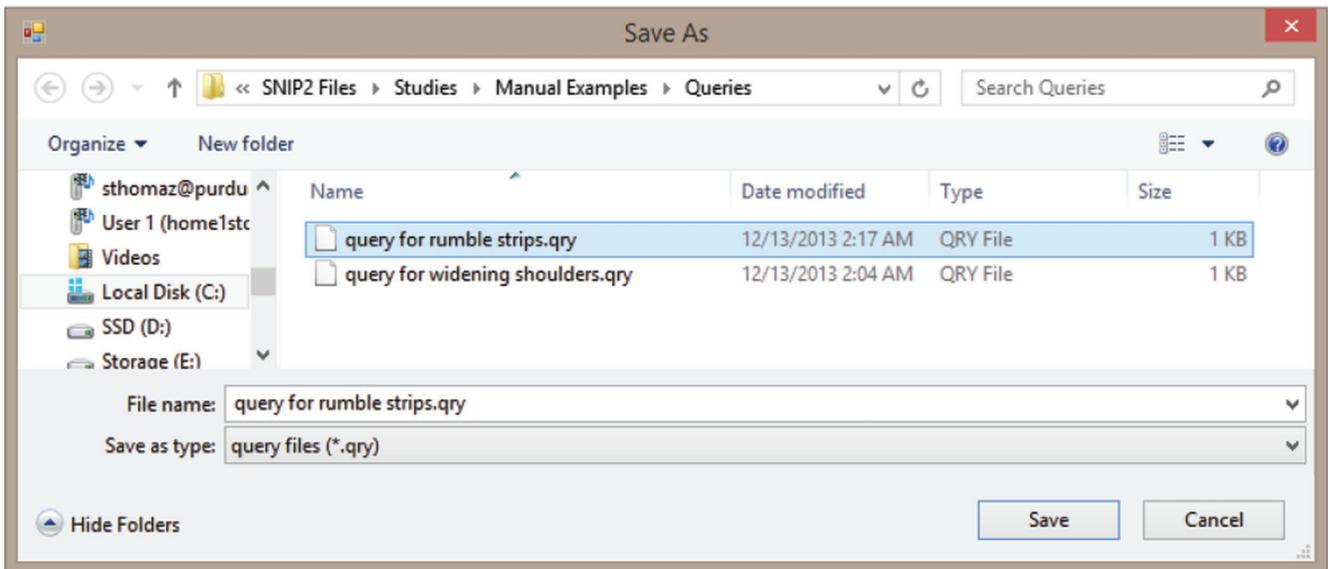
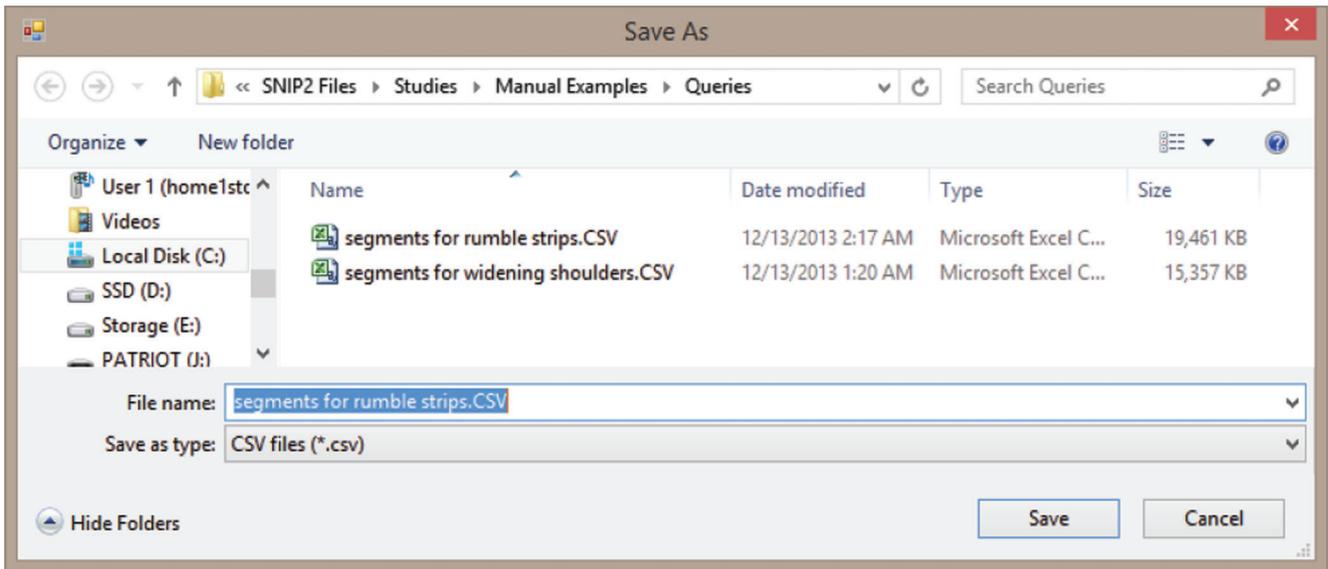


Figure 9.2 Saving the query results and definitions in the current study query folder.

YEARS <input type="checkbox"/> 2006 <input type="checkbox"/> 2007 <input type="checkbox"/> 2008 <input type="checkbox"/> 2009 <input checked="" type="checkbox"/> 2010 <input checked="" type="checkbox"/> 2011 <input checked="" type="checkbox"/> 2012	SCOPE <input checked="" type="radio"/> State <input type="radio"/> District <input type="radio"/> County <input type="radio"/> Township <input type="radio"/> City	INFRASTRUCTURE ELEMENTS <input type="radio"/> Segments <input checked="" type="radio"/> Intersections	EXPOSURE <input type="radio"/> VMT <input type="radio"/> Segment Length <input checked="" type="radio"/> Traffic Volume
--	--	--	---

Figure 9.3 Choosing intersections will assign exposure to “traffic volume.”

Intersection Criteria

- Intersection Type
- Rural or Urban
 - Urban
 - Rural
- State / Local
- Control type
 - Signals
 - Flashers
 - Non-signalized
- Number of Legs
- State Legs Number
- Local Legs Number

Environmental Criteria

- Manner of Collision
 - Rear End
 - Head On
 - Rear to Rear
 - Same Direction SideSwipe
 - Opposite Direction Sideswipe
 - Ran Off-Road
 - Right Angle
 - Left Turn
 - Right Turn
 - Backing Crash
 - Other
 - Non Collision
- Weather Conditions
- Type of Property

Figure 9.4 Intersections without signals and right angle crashes.

Intersection Criteria

- Intersection Type
- Rural or Urban
 - Urban
 - Rural
- State / Local
- Control type
 - Signals
 - Flashers
 - Non-signalized
- Number of Legs
- State Legs Number
- Local Legs Number

Intersection Criteria

- Intersection Type
- Rural or Urban
 - Urban
 - Rural
- State / Local
- Control type
 - Signals
 - Flashers
 - Non-signalized
- Number of Legs
- State Legs Number
- Local Legs Number
- Major Road Median Type

Figure 9.5 Urban and rural intersections processed separately.

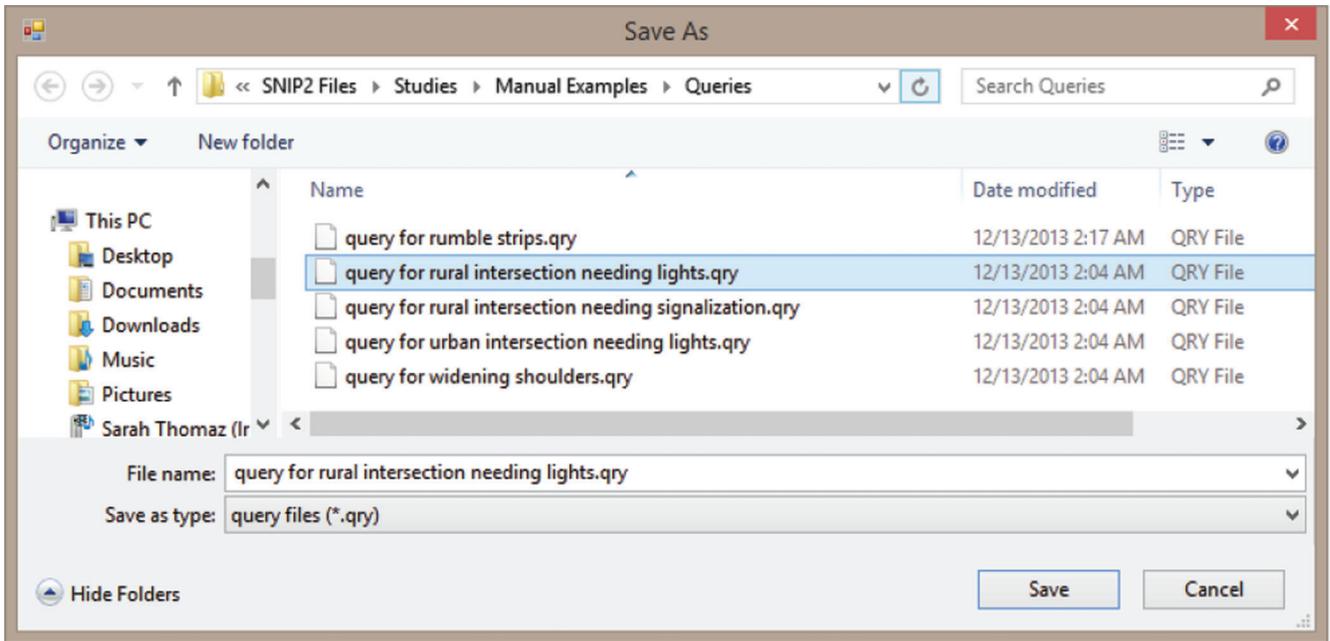


Figure 9.6 Resulting five queries for this example.

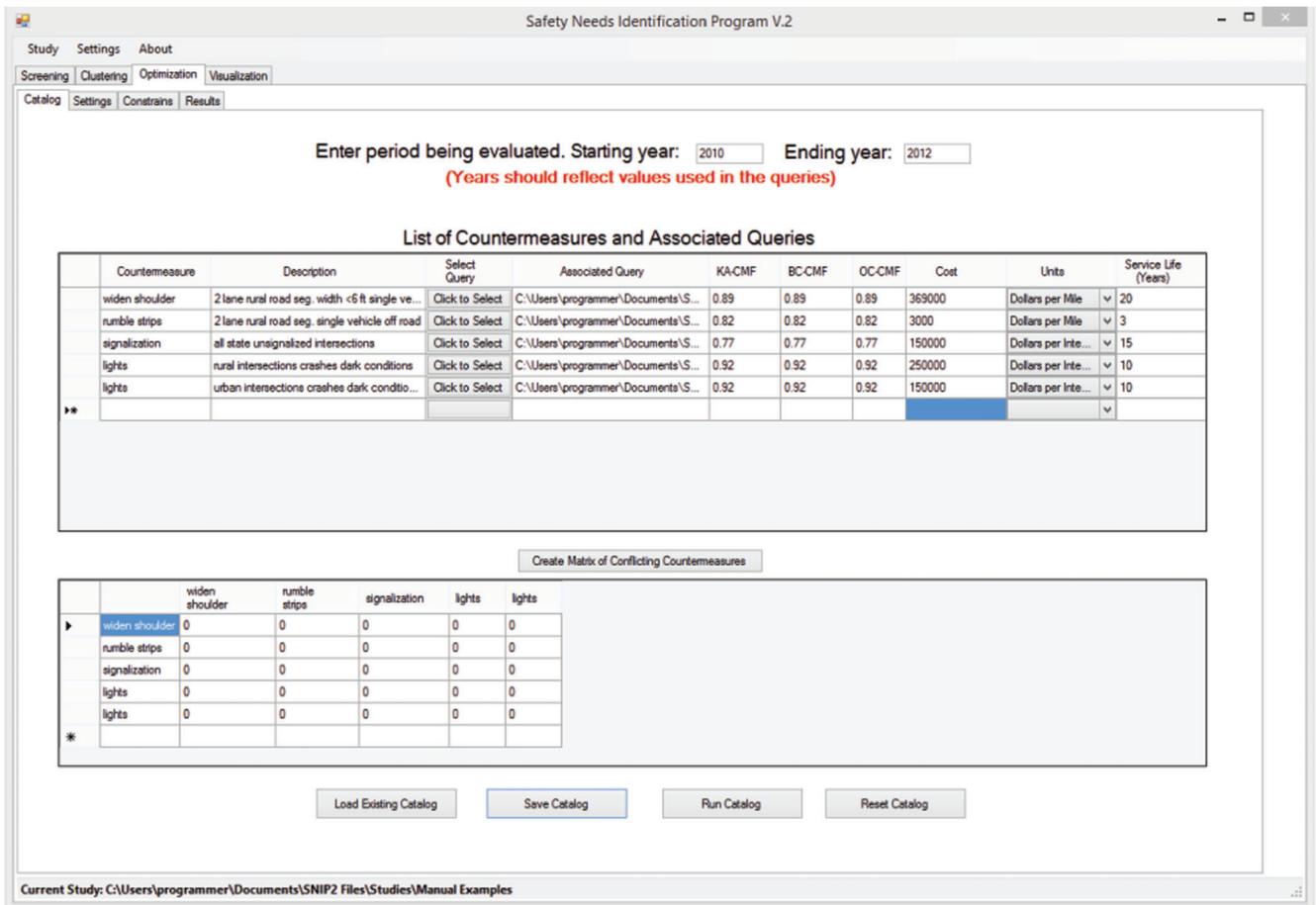


Figure 9.7 Assembled catalog of countermeasures.

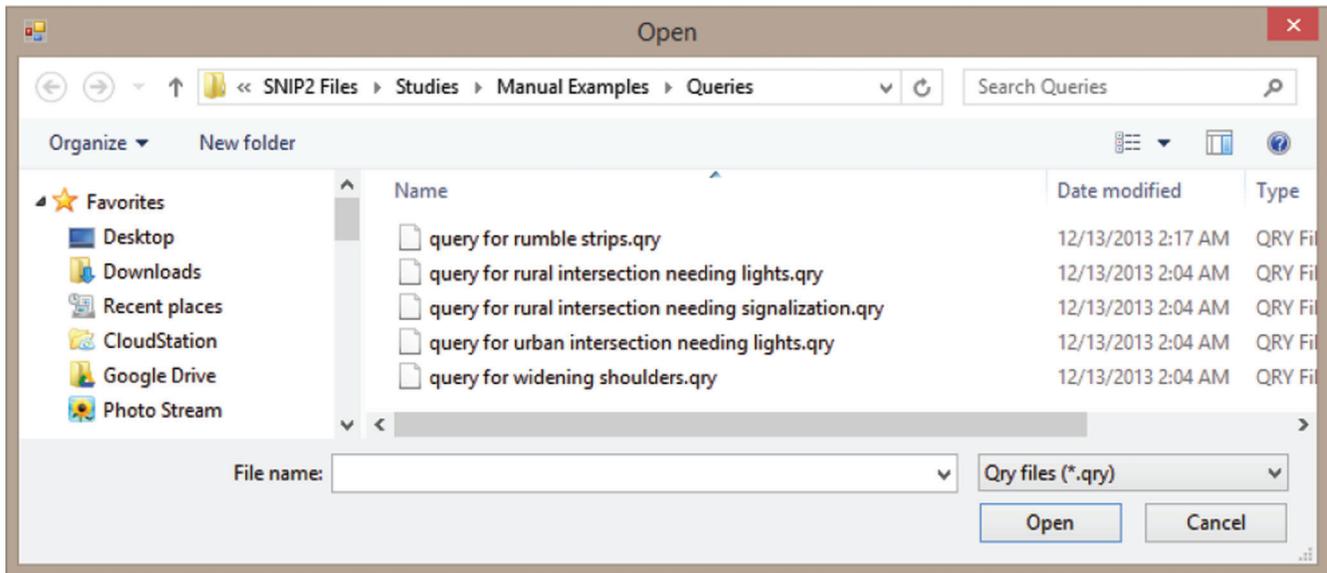


Figure 9.8 Selecting associated queries.

phase, the SNIP2 optimizer selects for each countermeasure the best subset of roads from the initial set of candidate roads. This selection of roads can be performed in three types of analyses: evaluation of a safety plan, prediction of the annual expenditure needed to save a certain number of crashes, or optimization of a safety plan within the given budget. The three types of analyses are discussed below.

9.2.2.1 Safety Plan Evaluation. This type of analysis aims to predict the annual safety benefit measured in the number of crashes saved by severity level and the associated annual expenditure if all the countermeasures included in the catalog are applied subject to constraints other than the total budget. The first step is to enter the name of two files:

(1) where the optimization results will be saved, and (2) the query names file that allows the optimizer locate all the files with query results (Figure 9.10).

Next, the input to the economic calculations should be entered: the interest rate of 4% and the inflation rate of 2%. The present year is 2013, the year of crash cost is 2011, the year of cost improvement is 2012; and the first and last years of the crash data period are 2010 and 2012, respectively (Figure 9.11).

Then, the exposure for segments and intersections are selected as VMT and AADT, respectively. The road element selection criteria includes the Adjusted Index of Frequency with the threshold value of 0.85 and the minimum number of crashes set at one for KA crashes, two for KA+BC crashes, and two for KA+BC+PDO

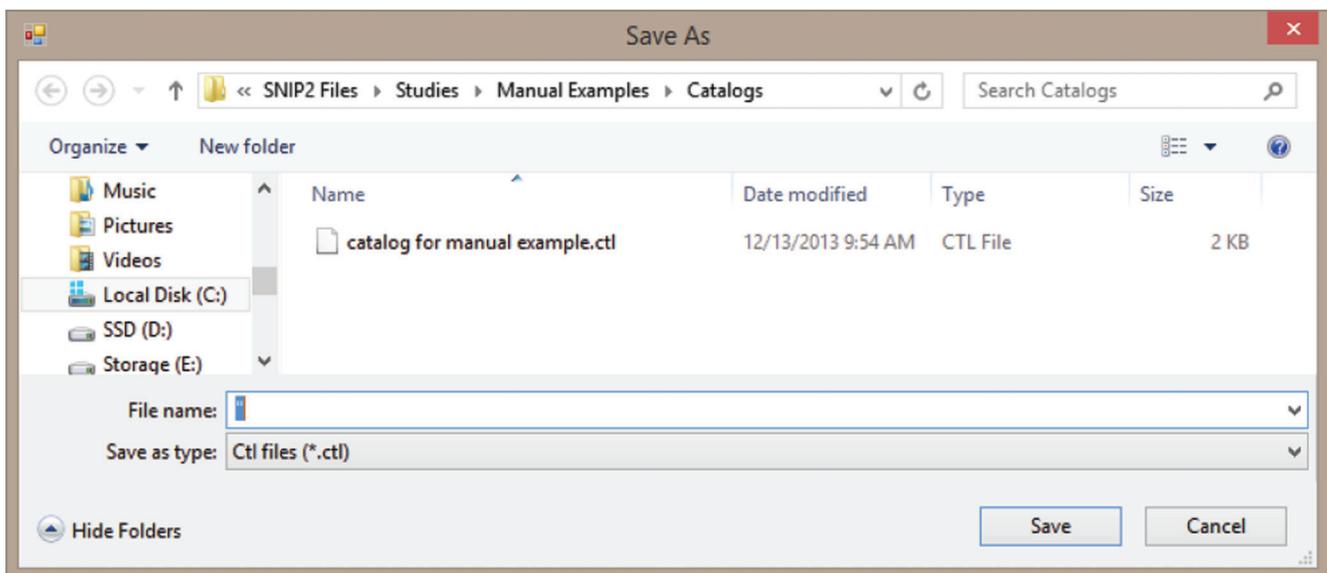


Figure 9.9 Saving example catalog of countermeasure.

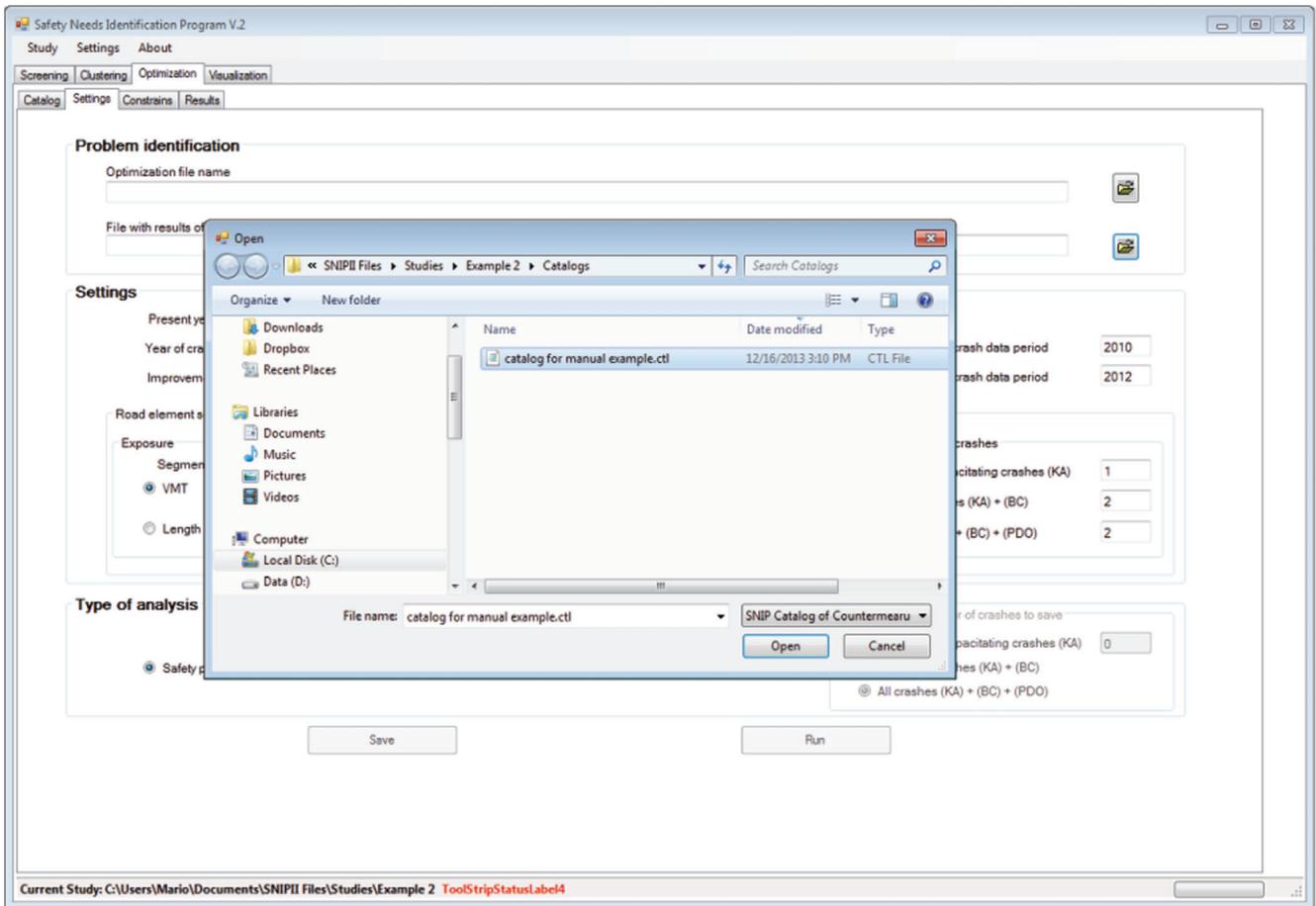


Figure 9.10 Selecting the catalog of countermeasures results file.

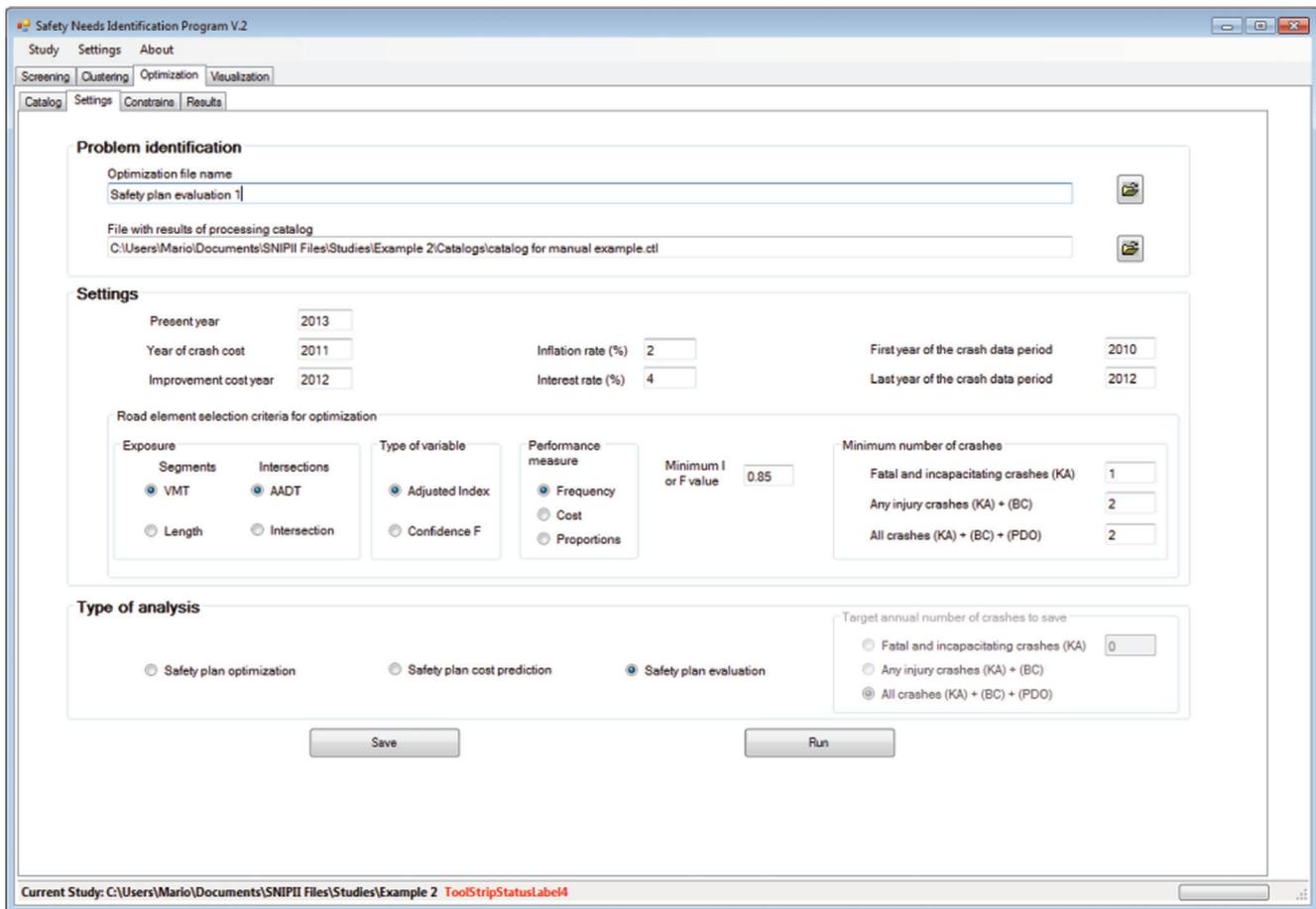


Figure 9.11 Optimization settings data.

crashes. Relaxing the minimum crash criteria will lead to saving a higher number of crashes but with a lower B/C ratio because road elements with the largest room for improvement are those with a high number of crashes and high values of I or F. Finally, the safety plan evaluation type of analysis needs to be selected, which is the Safety Plan Evaluation.

Although SNIP2 allows setting the regional constraints and program constraints of the safety plan evaluation plan, no constraints are set; only the conflicting countermeasures matrix is always required and it is filled with zeroes in the considered case.

Figure 9.12 shows the results. Applying the evaluated safety program will cost nearly \$3million per year, and it will save annually more than 330 crashes with the annual benefit over \$8.1 million. A detailed results file

that includes all the selected countermeasure and road elements is saved as a csv file (Figure 9.13).

9.2.2.2 Safety Plan Cost Prediction. The safety plan evaluation has shown that \$3million will buy four KA crashes, 48 BC crashes, and 282 PDO crashes saved a year if the entire safety plan is implemented. Another good question is about the annual cost of saving a certain number of crashes on the studied roads with the set of countermeasures at hand but applied in the most cost-effective way. Let us assume that the desired reduction is 50 KA+BC crashes. All the previous inputs remain except the type of analysis, which now is safety plan cost prediction and the target crash reduction, which is 50 injury crashes (KA or BC). This input is shown in Figure 9.14.

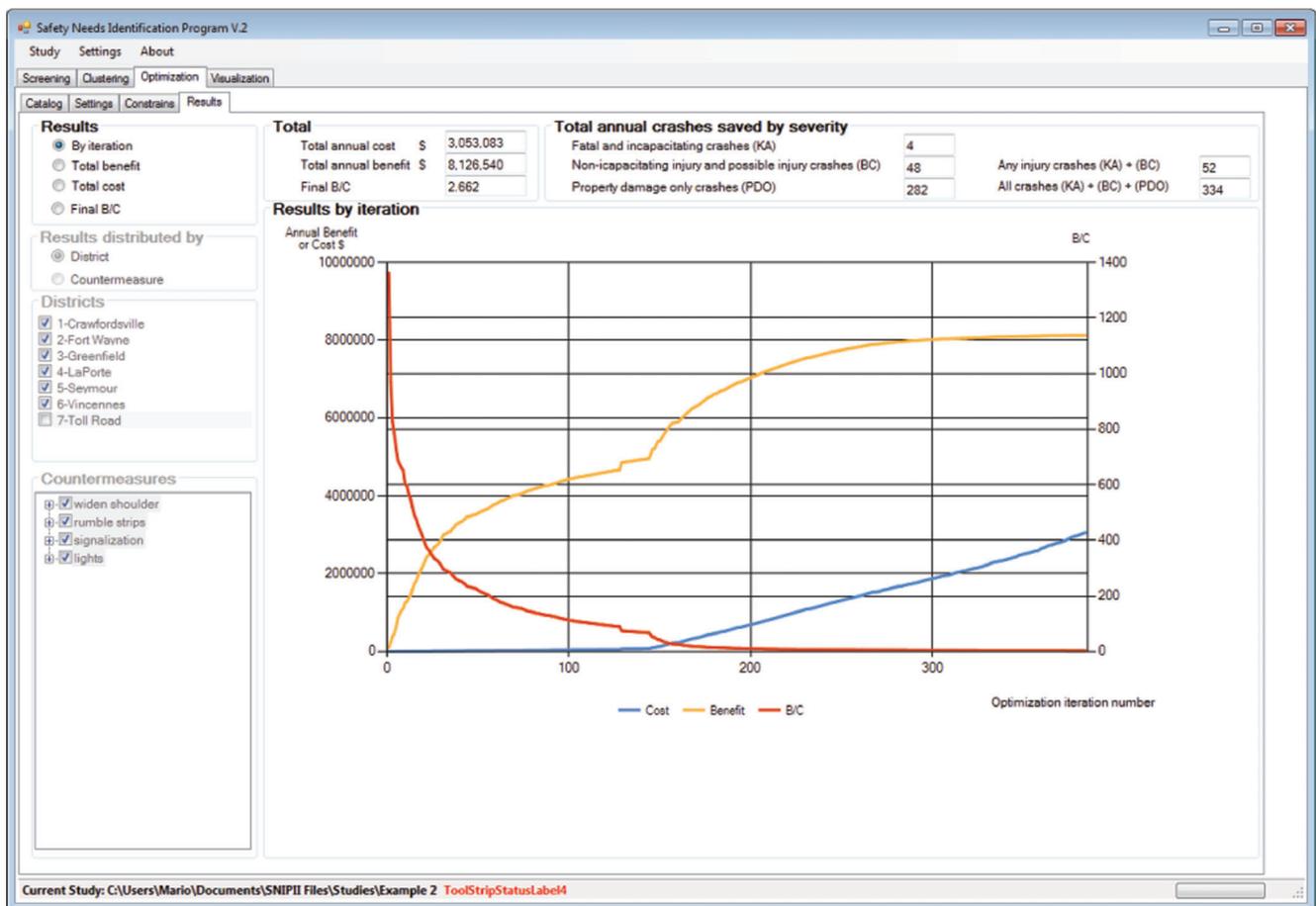


Figure 9.12 Safety evaluation program results.

The analysis executed by clicking the *Run* button produced the results shown in Figure 9.15. Saving 50 injury crashes would cost \$1.3 million per year.

9.2.2.3 Safety Plan Optimization. Unlike the previous types of analysis, the safety program optimization takes into account the total annual budget when selecting the best combination of roads and relevant countermeasures. In this case, a total budget of \$1.25 million was used.

Figure 9.16 shows the distribution of the expenditure between the INDOT districts as a result of running the safety plan optimization for the four countermeasures. To reduce expenditures in the LaPorte, Fort Wayne and Vincennes districts, new constraints were added to limit the total expenditure in each of the two districts to

\$230,000 (Figure 9.17). A minimum expenditure of \$200,000 was also set for a widening shoulders program as shown in Figure 9.17. Figure 9.18 shows the distribution of the expenditure between the INDOT districts after running the safety plan optimization with the extra constraints.

A log file in the csv format is generated as a result of the optimization. This file contains the list of non-selected road elements and the reason why they were not selected. Figure 9.19 shows an example of this file.

9.2.2.4 Saving the Optimization Problem. At any time, all the settings and inputs of the optimization problem can be saved in a csv file by clicking on the *Save* button (Figure 9.20). This file can be opened and edited later by the user.

1	Comments	Countermeasure	Accumulated Cost	Accumulated benefit	Accumulated B/C	Element Type	ID	Vari	KACrashesB	BCCrashesB	PDOCrashesB	Length	Exposure	O Benefit	Benefit	Cost	B/C	DI
140	Selected	2 lane rural road seg. single vehicle off road	70921.91	4924406.99	69.434	Segment	5773	2.121456046	2067.614368	2453.744272	1243.983404	0.43592803	680497605	5765.342044	5765.342044	462.1954312	12.47981877	
141	Selected	2 lane rural road seg. single vehicle off road	71364.23	4929838.63	69.08	Segment	14275	1.766790516	3257.481078	2132.192003	41.97047537	0.41717803	1105480963	5431.943556	5431.943556	442.315626	12.28001734	
142	Selected	2 lane rural road seg. single vehicle off road	71531.38	4936646.92	68.63	Segment	8908	1.575376595	4195.345813	1265.186116	1347.249105	0.534924242	1459762996	6808.281033	6808.281033	567.1557867	12.00423092	
143	Selected	2 lane rural road seg. single vehicle off road	72413.4	4942388.4	68.252	Segment	40532	1.538203168	4807.154357	1291.216914	43.11222916	0.454621212	1542016960	5741.4835	5741.4835	482.0149946	11.91142094	
144	Selected	2 lane rural road seg. single vehicle off road	73134.69	4949874.83	67.682	Segment	9012	1.801755312	4218.17219	1268.080041	2000.254726	0.68030303	1468409600	7486.434957	7486.434957	721.2955594	10.37915021	
145	Selected	2 lane rural road seg. single vehicle off road	73780.55	4956049.71	67.173	Segment	3137	1.316234966	3652.833185	1190.758635	1331.286677	0.609147727	1252604117	6174.880497	6174.880497	645.8527026	9.60818547	
146	Selected	all state unsignalized intersections	87012.27	5081983.04	58.405	Intersection	10889	1.041124478	35420.76875	34991.93832	55520.61949	1	41200	125933.3266	125933.3266	13231.71957	5.157532914	
147	Selected	all state unsignalized intersections	100243.99	5207187.85	51.945	Intersection	10921	2.230271596	27925.20681	63569.59735	33710.00577	1	28524	125204.8099	125204.8099	13231.71957	4.62474566	
148	Selected	2 lane rural road seg. single vehicle off road	100800.92	5211544.86	51.701	Segment	72366	1.781090016	3196.862392	1118.257722	41.88983602	0.525284091	1083190425	4357.009951	4357.009951	556.9357556	7.82318231	
149	Selected	all state unsignalized intersections	114032.64	5309954.42	46.565	Intersection	17022	2.193358827	7933.576603	50211.48313	40264.49953	1	18700	98409.55926	98409.55926	13231.71957	7.437397589	
150	Selected	all state unsignalized intersections	127264.36	5399947.87	42.431	Intersection	5127	2.235346631	4333.853862	64527.45574	21132.14352	1	9000	89993.45312	89993.45312	13231.71957	6.804732231	
151	Selected	2 lane rural road seg. single vehicle off road	127749.09	5403003.97	42.294	Segment	6361	1.752656217	2120.068717	896.2269742	39.80295554	0.45717803	698691285.5	3056.098287	3056.098287	484.7258772	6.207818763	
152	Selected	all state unsignalized intersections	140980.81	5485408.63	38.909	Intersection	16775	2.079237486	5792.36688	44021.8972	32590.3962	1	12638	82404.66028	82404.66028	13231.71957	6.227811874	
153	Selected	all state unsignalized intersections	154212.53	5564157.92	36.081	Intersection	15165	1.644297566	7812.050419	29217.50668	41719.73246	1	18330	87849.28956	87849.28956	13231.71957	5.951553698	
154	Selected	all state unsignalized intersections	167444.24	5631723.36	33.633	Intersection	1011	1.441175903	4430.423654	27930.64837	35204.36375	1	9230	67565.43577	67565.43577	13231.71957	5.106323135	
155	Selected	all state unsignalized intersections	180675.96	5698981.51	31.543	Intersection	274	1.357739773	8908.282324	40924.26412	17425.60903	1	10540	67238.15547	67238.15547	13231.71957	5.083100129	
156	Selected	all state unsignalized intersections	193907.68	5764322.45	29.727	Intersection	12399	2.209849964	5966.708351	30633.64772	29400.58541	1	5880	65340.94148	65340.94148	13231.71957	4.938204828	
157	Selected	all state unsignalized intersections	207139.4	5824740.06	28.12	Intersection	23294	1.611692102	8144.419057	22873.27279	29599.31022	1	19350	60417.60207	60417.60207	13231.71957	4.566118692	
158	Selected	all state unsignalized intersections	220371.12	5877109.49	26.669	Intersection	20816	1.374983608	8919.126175	30679.81638	12770.48612	1	21830	52369.42868	52369.42868	13231.71957	3.957870206	
159	Selected	2 lane rural road seg. single vehicle off road	220855.43	5879002.62	26.619	Segment	6235	2.209914098	1229.221616	627.898204	35.94860776	0.456780303	396138150	1893.130044	1893.130044	484.3041843	3.908969001	
160	Selected	2 lane rural road seg. width <6 ft single vehicle	225324.95	5895682.62	26.165	Segment	7897	3.83337075	11722.90637	16925.65664	64014.99954	0.167840909	1788750390	92663.56256	92663.56256	4469.520285	3.7318172	
161	Selected	2 lane rural road seg. single vehicle off road	225943.23	5897926.68	26.104	Segment	59555	2.041883129	1489.820345	717.3443439	37.45845844	0.583143939	483249050	2244.623148	2244.623148	4669.520285	3.630419555	
162	Selected	all state unsignalized intersections	239174.95	5944784.99	24.855	Intersection	23945	1.353555443	6339.646346	26825.10054	13693.56352	1	14100	46858.31041	46858.31041	13231.71957	3.541862115	
163	Selected	all state unsignalized intersections	252406.67	5991627.37	23.738	Intersection	9222	2.126795636	7473.636615	32207.43991	7161.303412	1	8610	46842.37994	46842.37994	13231.71957	3.467358154	
164	Selected	all state unsignalized intersections	265638.39	6037506.64	22.728	Intersection	4440	1.368838407	4896.060301	35009.77055	5973.441302	1	10360	45879.27215	45879.27215	13231.71957	3.467370352	
165	Selected	all state unsignalized intersections	278870.11	6082708.3	21.812	Intersection	15087	1.767997796	4219.69734	32418.60061	8563.355966	1	8730	45201.65391	45201.65391	13231.71957	3.416158698	
166	Selected	all state unsignalized intersections	292101.83	6126427.65	20.974	Intersection	8996	1.299993918	14859.81126	22666.56318	6462.982403	1	19330	43719.35684	43719.35684	13231.71957	3.304126664	
167	Selected	all state unsignalized intersections	305333.55	6169912.19	20.207	Intersection	7125	1.709605299	4695.101912	34278.15707	4511.278292	1	9668	43484.53728	43484.53728	13231.71957	3.286385948	
168	Selected	all state unsignalized intersections	318565.27	6212526.52	19.502	Intersection	900	2.476071339	4023.909944	25564.74683	13025.67675	1	4360	42614.33592	42614.33592	13231.71957	3.120619458	
169	Selected	all state unsignalized intersections	331796.99	6254233.98	18.85	Intersection	18657	0.888541683	17016.08582	16542.30637	8149.063784	1	23760	41707.45598	41707.45598	13231.71957	3.152081311	
170	Selected	2 lane rural road seg. width <6 ft single vehicle	337278.23	6271150.73	18.593	Segment	502	4.102883336	2315.184302	10718.7755	80948.01283	0.205833333	311395730	93981.97262	16916.75507	5481.262548	3.086299306	

Figure 9.13 Optimization output data file.

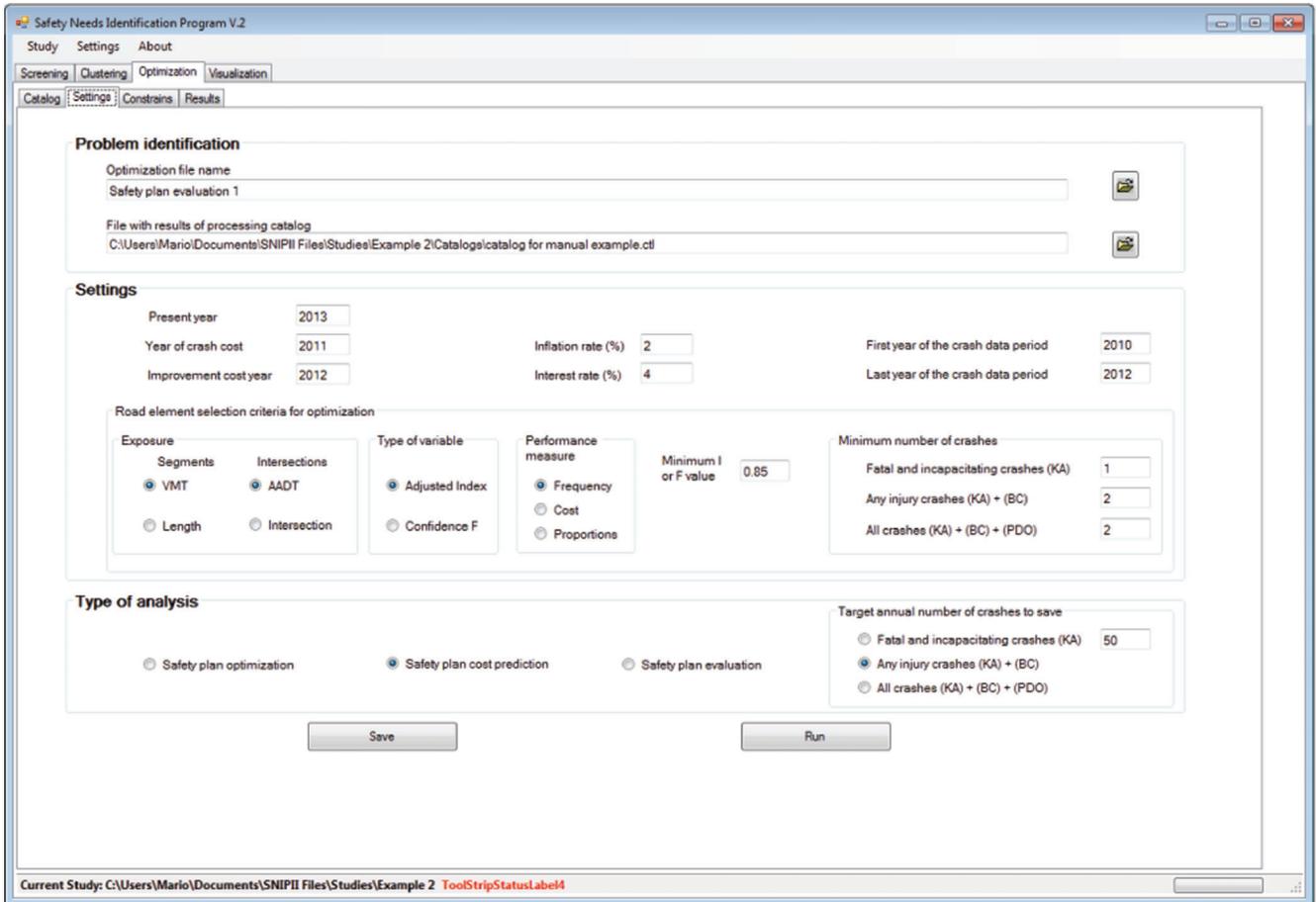


Figure 9.14 Safety plan cost predication settings.

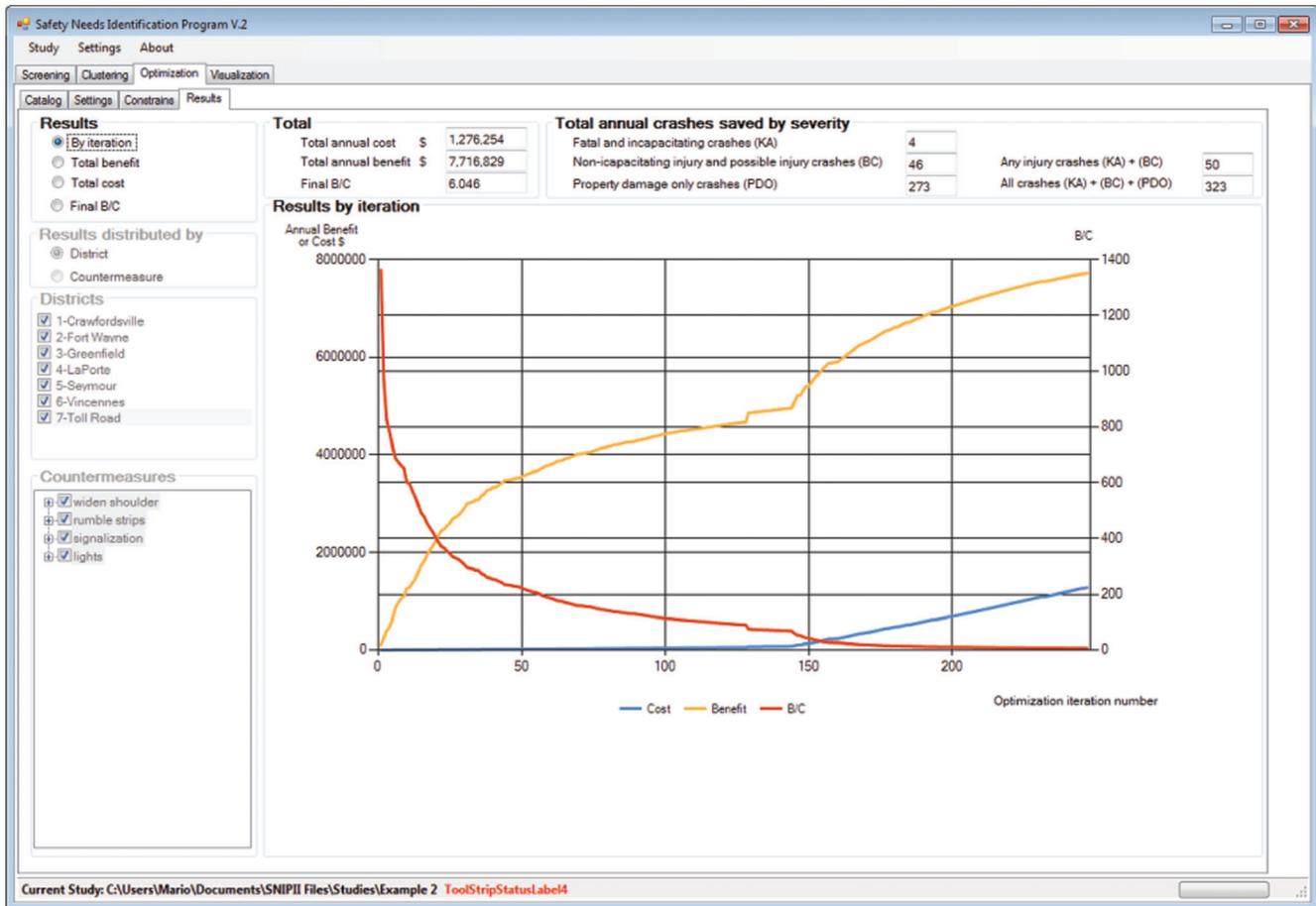


Figure 9.15 Safety plan cost prediction results.

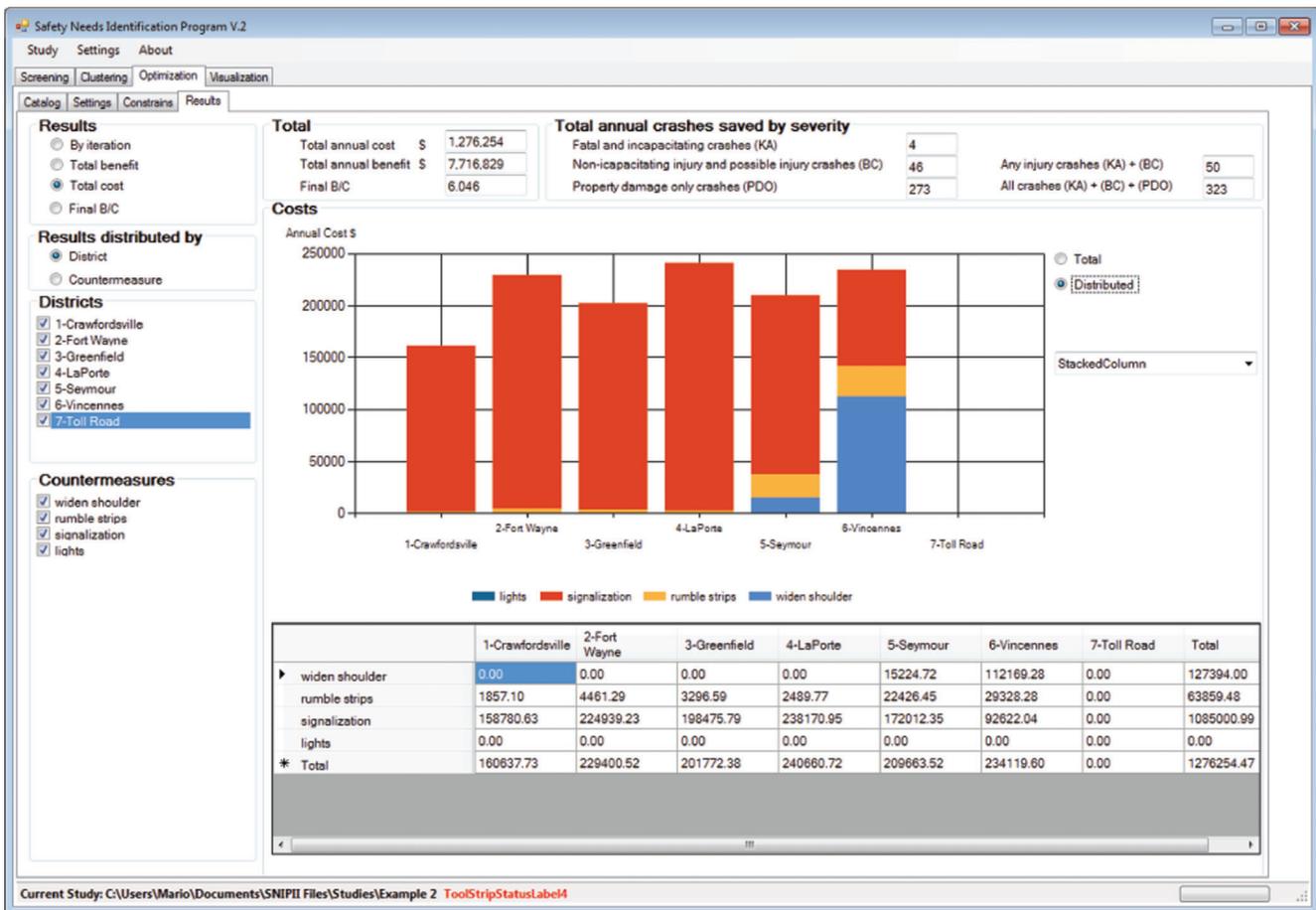


Figure 9.16 Safety plan optimization results: total cost.

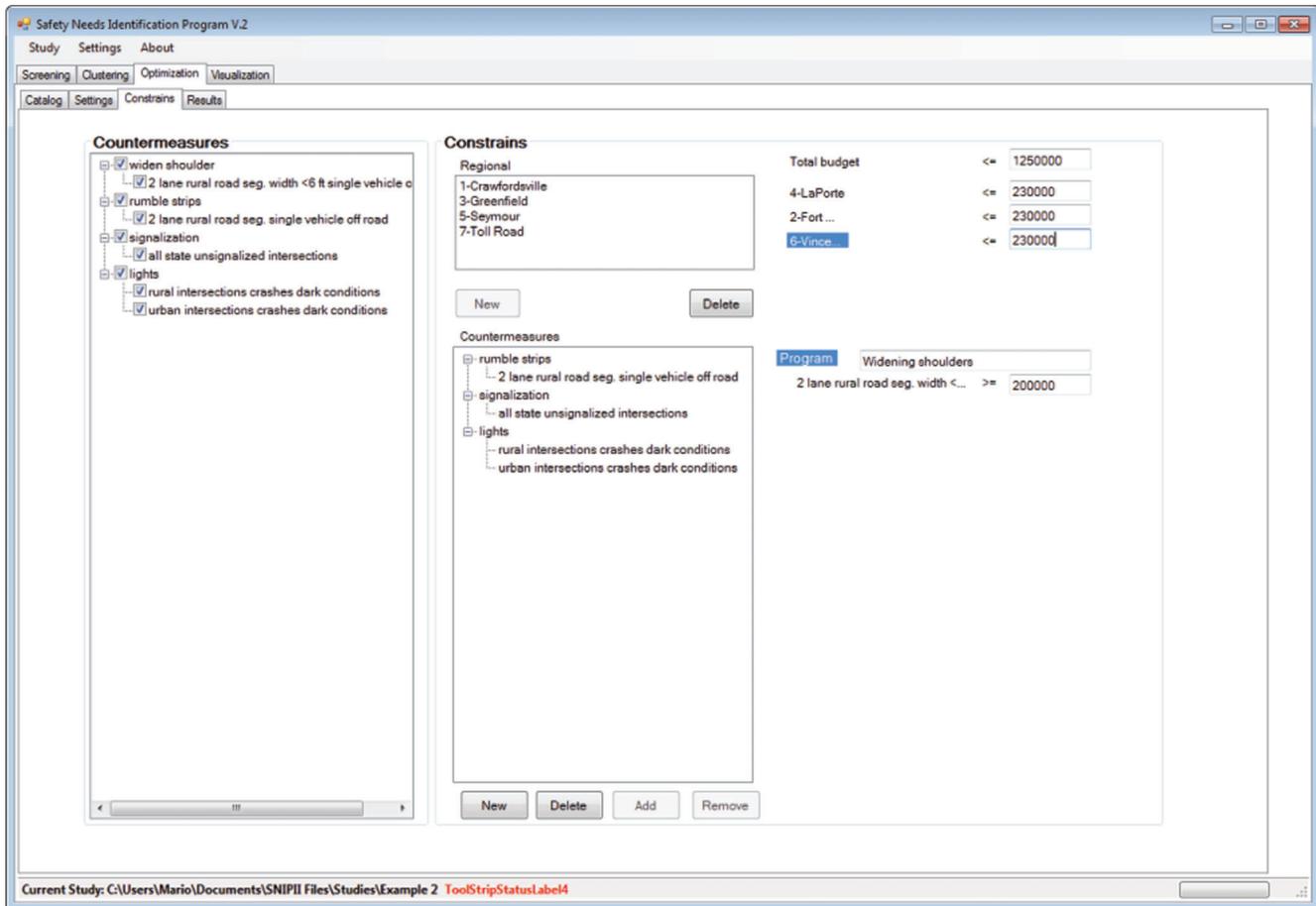


Figure 9.17 Safety plan optimization constraints.

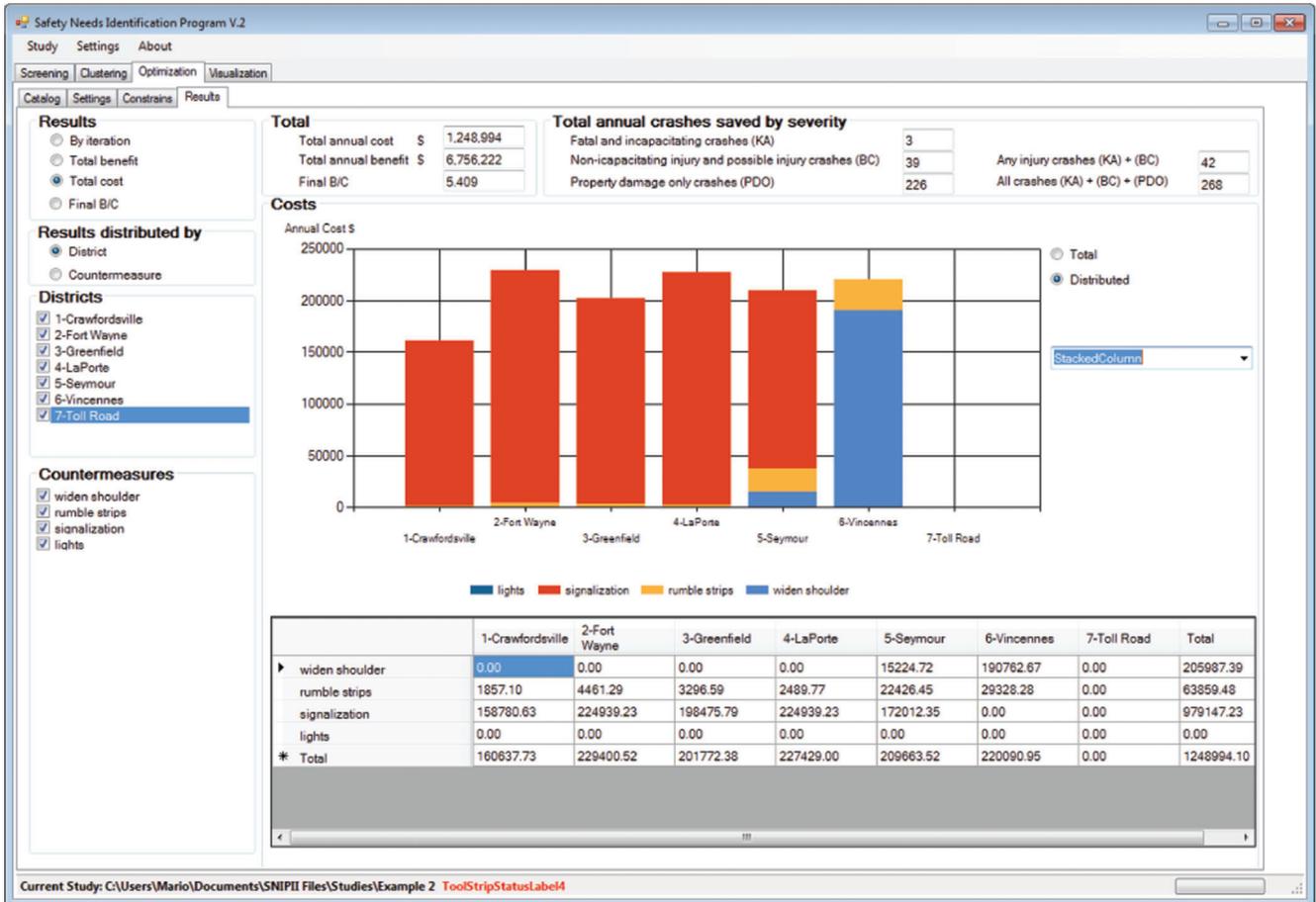


Figure 9.18 Safety plan optimization results with extra constraints: total cost.

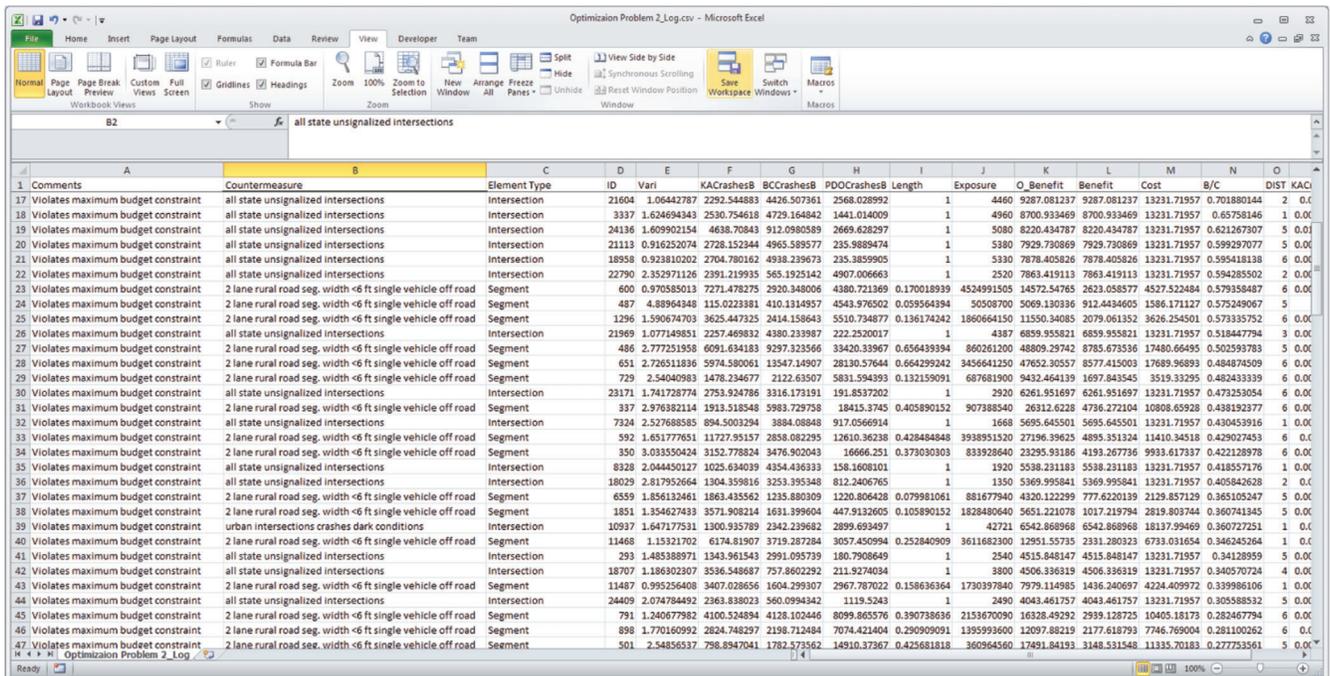


Figure 9.19 Optimization log file.

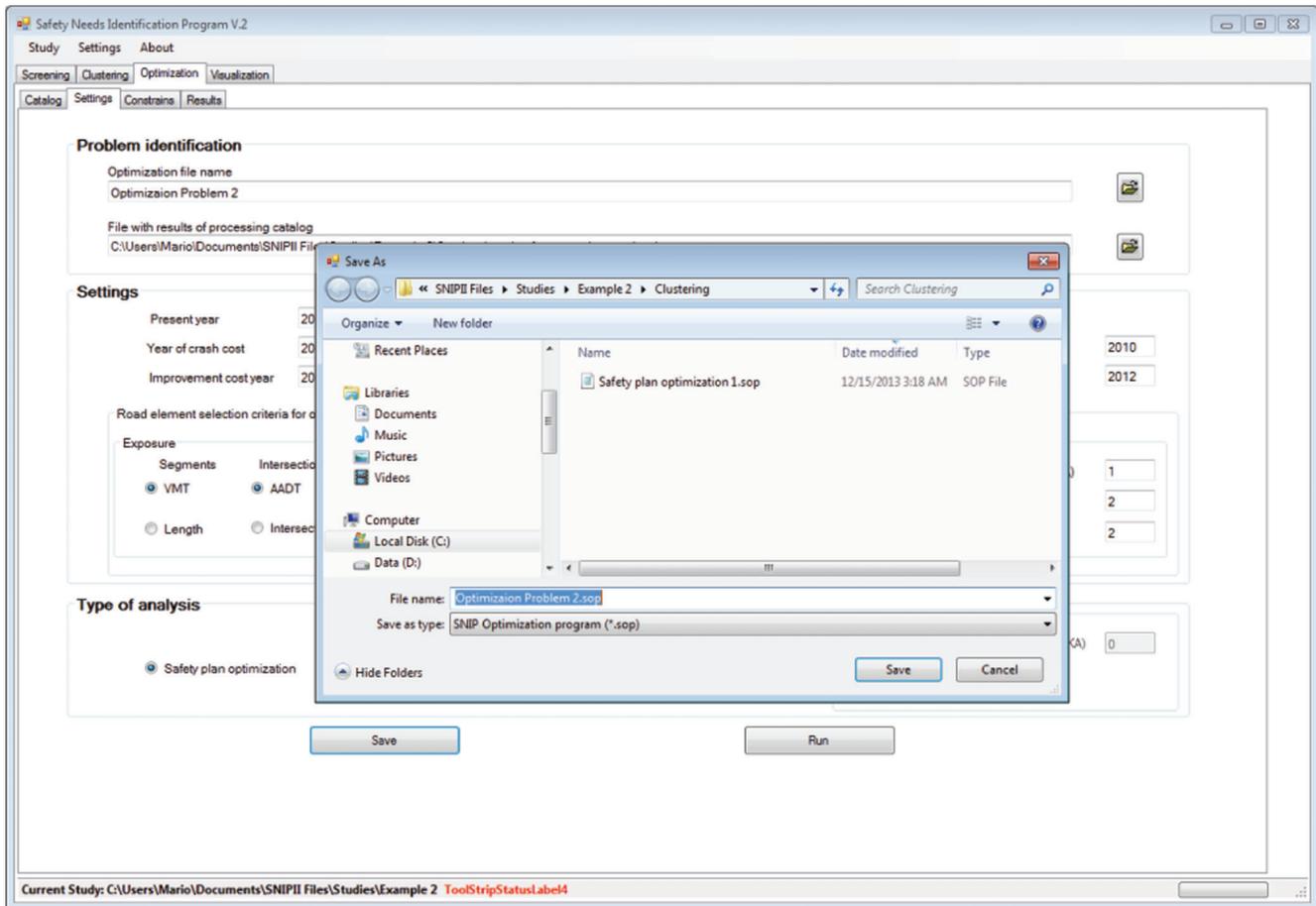


Figure 9.20 Saving the optimization problem input.

REFERENCES

- Lacy, J. K. (2001). *Recommended procedure for combining crash reduction factors*. Chapel Hill, NC: Highway Safety Research Center.
- Tarko, A. P., Mingyang, L., Romero, M., & Thomaz, J. (2014). *A Systematic Approach of Identifying Traffic*

Safety Needs and Intervention Programs for Indiana: Volume I—Research Report (Joint Transportation Research Program Publication No. FHWA/IN/JTRP-2014/03). West Lafayette, IN: Purdue University. <http://dx.doi.org.10.5703/1288284315497>.

APPENDIX A. SCREENING CONCEPTS

IDENTIFICATION METHOD

The safety identification method includes components designed to fulfill the agency's need to systematically investigate a particular problem. The following items are the core components necessary for a successful identification method:

1. Scope, elements, and selection criteria
2. Safety performance measures
3. Exposures measures
4. Statistical evaluation measures

Scope, Elements, and Selection Criteria

Scope

The scope or domain is the geographic unit in which the user wishes to conduct the screening. In the safety screening tool, three scopes have been defined: state, county, and city/town. The scope can be limited to a particular county/township or multiple counties/townships, but should always be greater than the elements in geographic extent.

Element

An element is the smallest unit of aggregation level that a user wishes to investigate. Elements can be the facility type (e.g., segments, intersection points, intersection, ramps, or bridges) or can be a smaller geographic unit within the scope. Therefore, the scope or domain is the group of elements an agency wishes to investigate.

Selection Criteria

After defining the scope and elements, it is important to define the selection criteria. The selection criteria basically facilitate obtaining a subset of the elements within the scope. Within the conceptual framework of safety screening, the *Screening* tab is used to define the selection criteria, which can be of two types:

- Crash selection
- Element selection

Crash Selection Criteria

Crash selection criteria are considered in order to investigate a specific type of crash. For example, an agency might be interested in only fatal or incapacitating injury types of crashes or only nighttime crashes. An example might be obtaining only alcohol-related crash locations for targeted enforcement purposes. The crash selection criteria are mainly dependent on the crash variables and their availability.

Element Selection Criteria

Element selection criteria also have a very specific purpose. Since the Indiana road inventory is embedded in the master record sets, a user might be interested in the crash propensity for a specific design condition (e.g., a particular roadway with a specific median type/width). Combining the crash and element selection criteria can serve as a great tool for choosing candidates for a specific program. Figure A.1 shows the interaction among the scope, element, and selection criteria in the overall safety screening process.

TABLE A.1
Exposure measures for different road elements

Element of Investigation	Exposure
County	Population, VMT, registered vehicle, area
City	Population, VMT, registered vehicle, area
Township	Population, VMT, registered vehicle, area
State segment	Link volume (ADT, VMT), Length
State-state intersection	Total approach volume (ADT, VMT)
State-local intersection	State (major) road volume
Ramp	Link volume (ADT, VMT), length
Bridge	Link volume (ADT, VMT), length

Safety Performance Measures

After a user defines the scope, element, and selection criteria, it is important to define the unit of identification. The identification unit is analogous to the *measures of safety* which can have three basic types:

- *Crash Frequency.* Crash frequency is the crash counts of all crashes or a specific subset of crashes as determined by the user.
- *Crash Cost.* Crash cost applies to all crashes or a specific subset of crashes as determined by the user.
- *Crash Rate.* The crash rate is the crash frequency/exposure and can vary based on the type of elements selected.
- *Proportions of Crash.* The proportion is the ratio of relevance crashes (studied crashes) to reference crashes (e.g., the proportion of rear-end crashes to the total number of crashes).

Exposure Measures

Exposures are used to estimate crash rates as the ratio of the crash count and a specific measure of exposure reflecting the analyzed period. They can be AADT, VMT, or road length, depending on the element under investigation (see Table A.1).

Statistical Evaluation

Notation

Basic variables:

c= number of studied crashes during the analysis period

w= cost of crashes on road element during the analysis period

m= estimate of the expected crashes or cost during the analysis period and for the exposure

v= variance of the m estimate

Variables needed to calculate w, m, and v:

e= exposure on road element (AADT, length L, VMT during the analysis period)

r= number of reference crashes on road element during the analysis period

u= unit crash cost

N= number of road elements in the group of roads

S= total number of studied crashes in the group of roads during the analysis period

R= total number of reference crashes in the group of roads during the analysis period

E= total exposure in the group of roads during the analysis period

W= total cost of crashes in the group of roads during the analysis period

sub k= severity level k

Two distributions are used to evaluate the statistical significance of the safety problem: Gamma distribution and Negative Binomial Distribution. Gamma distribution has parameters α and β , such that mean is $\alpha\beta$, and variance is $\alpha\beta^2$ and density:

Interrelationship of Scope, Element & Selection Criteria in Safety Screening

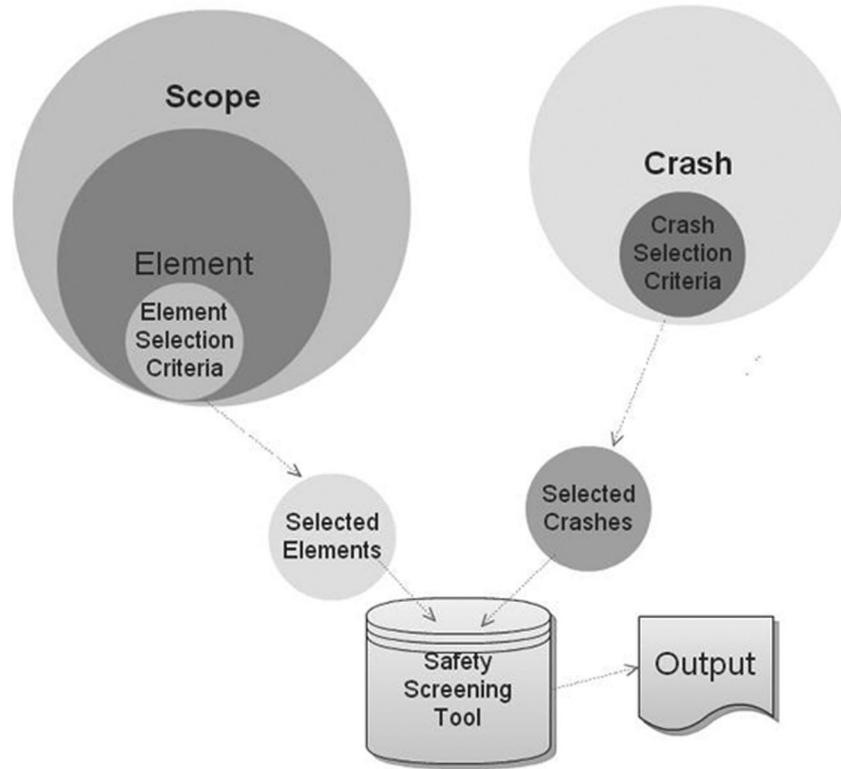


Figure A.1 Scope, elements, and crashes and their selection/screening criteria.

$f(\lambda) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\lambda/\beta)$. The Negative Binomial distribution can be viewed as a mixture of Poisson distributions with the Poisson parameter λ distributed according to the Gamma. The parameters of Negative Binomial the distributions are inherited from the Gamma. The mean is $\alpha\beta$, and the variance is $(\alpha\beta + \alpha\beta^2)$, and density: $P(c) = \frac{\Gamma(\alpha+c)}{\Gamma(\alpha)c!} \left(\frac{\beta}{1+\beta}\right)^c \left(\frac{1}{1+\beta}\right)^\alpha$. The MS Excel parameterization of the Gamma distribution is as introduced above while the Negative Binomial distribution uses parameters: $r = \alpha + 1$ and $p = 1/(1+\beta)$.

Concepts

Let c be the recorded number of crashes of a certain type used to evaluate a road element's safety during the analysis period. An agency wants to know if this number of crashes indicates that there is a safety problem on the considered road element. The safety problem is confirmed if the number of crashes c is significantly higher than the number expected for the exposure e on the considered road element.

There are a number of exposure measures (e.g., the traffic volume entering an intersection or the vehicle-miles travelled along a road segment). The number of crashes m expected for the exposure is the product of the average crash rate S/E in the considered group of roads and the exposure e on the studied road element.

The segment length can be used if the traffic volume is missing. This option is reserved for local roads that typically do not have traffic volumes measured. The number of crashes m expected for the exposure is the product of the average crash density in the group of roads and the length of the studied road element.

Checking if the considered crashes constitute too large of a proportion of a wider category of crashes (reference crashes) is another important safety test. For example, all intersection crashes may serve as reference crashes for a proportion of right-angle crashes. The number of crashes m expected in this case is the product of the reference proportion S/R (average proportion of intersection crashes that are right-angle crashes in the group of considered intersections) and the reference crashes r at the studied intersection.

Crash Frequency (Count)

The first step is to estimate the crash rate S/E in the considered group of roads, where S is the total number of considered crashes in the group of roads and E is the total exposure in that group. The expected number of crashes m on the considered road element is the product of the exposure e on this road element and the crash rate S/E in the road group: $m = e \cdot S/E$. The variance of this estimate is caused by the varying number of S crashes scaled with e/E . The estimate m is distributed according to the Gamma distribution $G(\alpha=S, \beta=e/E)$ with the variance $v = \alpha\beta^2 = S(e/E)^2$.

The test of whether the actual number of crashes c is larger than the number m expected for the exposure e is done through checking whether the crash count c is sufficiently far into the right tail of the distribution of the crash counts around the uncertain Gamma-distributed mean m . This test calls for using the Gamma-mixture of Poisson distributions, thus for using the Negative Binomial distribution $NB(\alpha=S, \beta=e/E)$. The crash count c indicates that the current safety on the road element is worse than expected for the exposure if the cumulative distribution NB at c takes a high value (i.e., higher than 0.95). This value is called *Confidence* F_{CF} – the probabilistic measure which varies between 0 and 1.

Using the Excel notation, the calculation of confidence F is:

$$F_{CF} = \sum_{x=0..c} \text{NegBinom.Dist}(x, r=\alpha+1, p=1/(1+\beta)) = \text{Beta.Dist}(1/(1+\beta), \alpha, c+1, 1),$$

or more specifically: $F_{CF} = \text{Beta.Dist}(1/(1+e/E), S, c+1, 1) =$

Another method of statistical significance is the *Index I_{CF}*, the quality control measure that indicates the difference between the estimated safety and the target safety (expected for the exposure) measured with the standard deviation of the difference estimate. A high value of index I, for example, higher than 2, indicates a safety problem.

$$I_{CF} = \frac{c-m}{\sqrt{v}}$$

The value I_{CF} may be questionable and inconsistent with the significance F_{CF} if the underlying distribution is strongly skewed to the right (Gamma and Negative Binomial distributions tend to be skewed if the expected value is close to zero). It may lead to an I-based ranking that is inconsistent with the F-based ranking. Since agencies may prefer using index I_{CF} , an adjusted I_{AF} value is proposed that is determined based on the calculated F_{CF} value. It uses an “equivalent” normal distribution which preserves the original m , c , and F_{CF} values. The equivalent parameter σ_e needs to be calculated. Given that the standardized cumulative normal distribution can be closely and conveniently approximated with the logistic function:

$$\Phi_1(c) \cong 1 - \frac{1}{1 + \exp\left(1.7 \cdot \frac{c-m}{\sigma}\right)}$$

the equivalent σ_e that provides the same F value as the Gamma distribution is:

$$F_{CF} \cong 1 - \frac{1}{1 + \exp\left(1.7 \cdot \frac{c-m}{\sigma_e}\right)}$$

or:

$$\sigma_e = \frac{1.7(c-m)}{\ln\left(\frac{F_{CF}}{1-F_{CF}}\right)}$$

The adjusted I_{AF} value can be calculated as follows:

$$I_{AF} = \frac{s-m}{\sigma_e} = \frac{\ln(F_{CF}) - \ln(1-F_{CF})}{1.7}$$

To control the overflow error, small values of F_{CF} and $1-F_{CF}$ should not be used. Instead, an assumed small negative value of

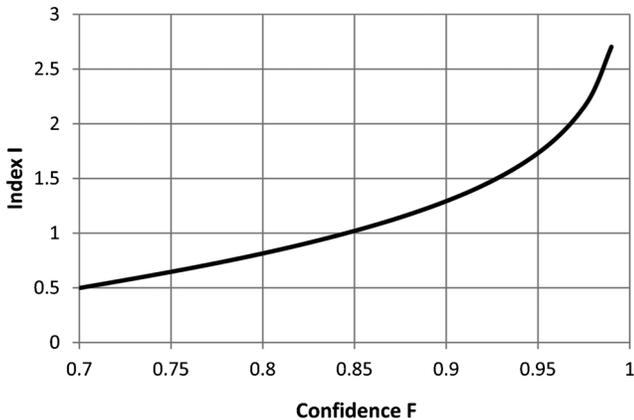


Figure A.2 Relationship between the index of Frequency I and the significance of Level F.

$\ln(F)$ and $\ln(1-F)$ (e.g., -99) should be used. Since the equation for I_{AF} is an approximation (although a close one), I_{AF} should be set at a value 0 of $s=m$ to avoid an obviously counterintuitive result.

The relationship between index I_{AF} and significance F_{CF} is shown in Figure A.2 and summarized in Table A.1. It can be concluded that I_{AF} lower than 1.25 indicates no or weak statistical evidence of a safety problem ($F_{CF} < 0.90$), and I_{AF} between 1.3 and 1.7 indicates considerable evidence (F_{CF} between 0.90 and 0.95), and I_{AF} between 1.7 and 2.7 indicates strong evidence (F_{CF} between 0.95 and 0.99), and I_{AF} larger than 2.7 indicates very strong evidence.

Proportion of Crashes

The reference proportion is the estimated proportion of studied crashes S in the reference crashes R in the group of roads: S/R . The expected number of crashes at a single road element which corresponds to the reference proportion is calculated as: $m=r \cdot S/R$, where r is the number of reference crashes on the road element and S/R is the proportion of studied crashes in the reference crashes in the group of roads (reference proportion). The variance of the estimate m is caused by the variability of all the component crash counts: r , S , and R . These counts are not independent from each other as explained in Figure A.3. To estimate the variance of the m estimate, the crashes in the group of considered roads have been divided into four counts: c , c_1 , c_2 , c_3 , and c_4 in a way that these counts vary independently and can be used to calculate the counts r , S , and R : $R=c+c_1+c_2+c_3$, $S=c+c_1$, and $r=c+c_2$. The derivation of the variance of estimate m ; $v=(2crSR+r^2SR+rS^2R-3r^2S^2)/R^3$ is described below.

The variance of $m=rs/d$ is calculated as the variance of $m=(c+c_1)(c+c_2)/(c+c_1+c_2+c_3)$ with four independent sources of Poisson variance: c , c_1 , c_2 , and c_3 . The variance is derived from the following equation:

$$\text{var } m(c, c_1, c_2, c_3) = \left(\frac{\partial m}{\partial c}\right)^2 \cdot c + \dots + \left(\frac{\partial m}{\partial c_3}\right)^2 \cdot c_3. \quad (\text{A.1})$$

The validity of the derived variance and of the assumption of Gamma distribution applied to this criterion has been evaluated using a simulation of 10,000 values of the m estimates for two distinct sets of values of c , r , s , and d . The simulated distribution of the m estimates and corresponding Gamma distributions with the parameters calculated in steps 2, 3, and 4 are shown in Figure A.4 (for $(c=10, s=210, r=210, d=510, m=45.3, v=18.1)$) and Figure A.5 for $(c=1, s=6, r=3, d=18, m=0.44, v=1.0)$. The simulation-based evaluation confirms the validity of the method for estimating right-hand distribution tails of m estimates.

Estimation of the significance F_{CP} is made using equation

$$F_{CP} = \text{Beta.Dist}(1/(1+\beta), \alpha, c+1, 1),$$

where $\alpha=m^2/v$ and $\beta=v/m$, thus

$$F_{CP} = \text{Beta.Dist}(1/(1+v/m), m^2/v, c+1, 1).$$

The adjusted Index I_{AP} is calculated as

$$I_{AP} = \frac{\ln(F_{CP}) - \ln(1-F_{CP})}{1.7}$$

Cost Criterion

Traffic volume, AADT, and segment length are useful in calculating the expected cost of crashes on the studied road element. The expected cost of crashes can be obtained by multiplying the crash cost rate per unit exposure averaged for the studied road network with the exposure values for the studied road element.

The expected cost of crashes on a road segment or at an intersection exceeds the expected cost under the given exposure. An estimate of the expected number of crashes at severity level k is distributed according to Gamma with parameters $\alpha=c_k$ and $\beta=1$. Thus, the mean value is $m_k=c_k$ and the variance is $v_k=c_k$. The

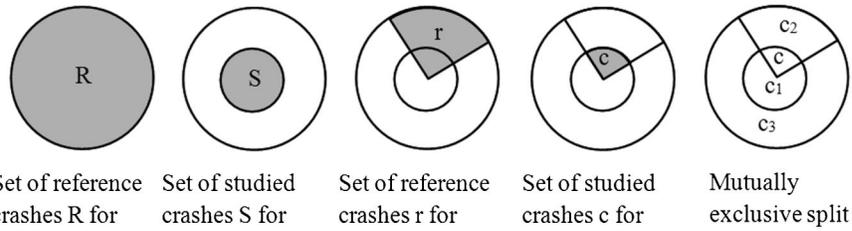


Figure A.3 Dependence between crash counts.

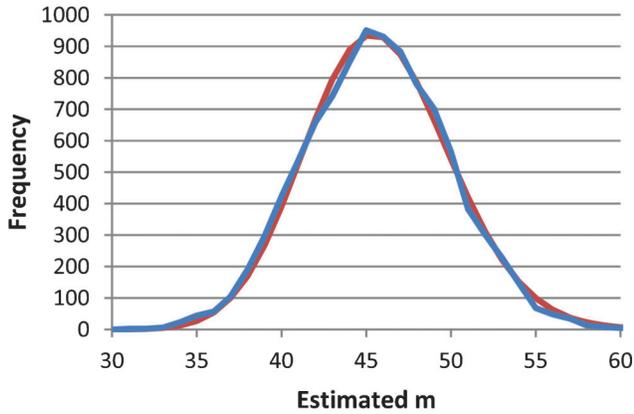


Figure A.4 Case 1: Simulated (blue) versus calculated (red) distributions of m estimates.

scaling property of the Gamma distribution allows assuming that the cost of all crashes of severity k at the location is also distributed as Gamma with parameters: $\alpha=c_k$, $\beta=u_k$. The corresponding mean $m_k=c_k u_k$, and variance $v_k=c_k u_k^2$. Thus, the cost of all crashes on the road is:

$$w = \sum_k c_k u_k$$

and the close approximation of the variance of cost estimate (confirmed with Monte Carlo experiments) is:

$$v_1 = \sum_k c_k u_k^2$$

If the cost of crashes on the road expected for the exposure can be calculated as: $m=e \cdot W/E$, where e is the exposure on the considered road element, W is the total cost of crashes in the group of roads, and E is the total exposure in the group of roads.

The estimate m has variance $v_2 = \sum_j \frac{v_{wj}}{E^2}$ which is the total cost variance and E is the total exposure in the road group. The variance of the difference between the w and m estimates is

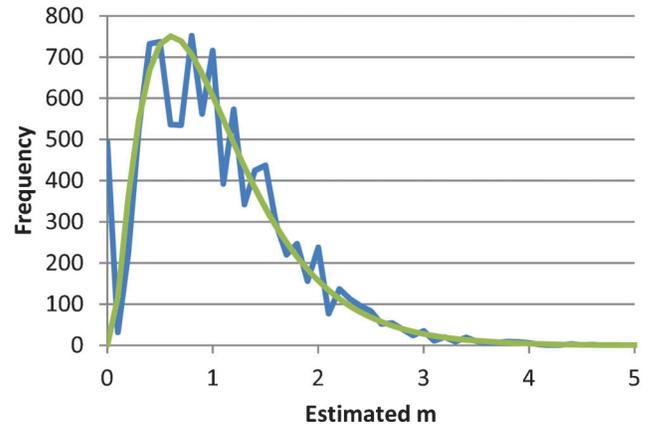


Figure A.5 Case 2: Simulated (blue) versus calculated (green) distributions of m estimates.

approximated with the sum of the two component variances v_1 and v_2 . The test is this time based on the Gamma distribution:

$$F_{CC} = 1 - \text{Gamma.Dist}\left(w, \alpha = \frac{m^2}{v_1 + v_2}, \beta = \frac{v_1 + v_2}{m}, 1\right)$$

and index I_{AC} is calculated as before:

$$I_{AC} = \frac{\ln(F_{CC}) - \ln(1 - F_{CC})}{1.7}$$

COMPUTATIONS

See Table A.2, "Confidence F and index I for the three screening criteria."

TABLE A.2
Confidence F and index I for the three screening criteria

Screening Criterion	Crashes/Cost on Road Element	Crashes/Cost Expected on Road Element m	Variance v	Significance F	Index I	Index I _A
Crash frequency	c	$m = e(S/E)$	$v_1 = c$ $v_2 = S(e/E)^2$ $v = v_1 + v_2$	$F_{CF} = \text{Beta.Dist}(1/(1+v_2/m), m^2/v_2, c+1, 1)$	$I_{CF} = (c-m)/v^{1/2}$	$I_{AF} = (\ln(F_{CF}) - \ln(1 - F_{CF}))/1.7$
Crash proportion	c	$m = r(S/R)$	$V = (2crSR + r^2SR + rS^2R - 3r^2S^2)/R^3$	$F_{CP} = \text{Beta.Dist}(1/(1+v_2/m), m^2/v_2, c+1, 1)$	$I_{CP} = (c-m)/v^{1/2}$	$I_{AP} = (\ln(F_{CP}) - \ln(1 - F_{CP}))/1.7$
Crash cost	$w = \sum c_k u_k$	$m = e(W/E)$	$v_1 = \sum c_k u_k^2$ $v_2 = \sum d_k m_k^2 u_k^2$ $v = v_1 + v_2$	$F_{CC} = \text{Gamma.Dist}(w, m^2/v, v/m, 1)$	$I_{CC} = (w-m)/v^{1/2}$	$I_{AC} = (\ln(F_{CC}) - \ln(1 - F_{CC}))/1.7$

Notations in alphabetic order: e = number of crashes on road element, c_k = number of studied crashes of severity k, d = over-dispersion parameter of SPF, e = exposure on road element (AADT, length L, VMT), E = total exposure in the group of roads, m = estimate of the crashes or cost expected for the exposure, r = number of reference crashes on road element during, R = total number of reference crashes in the group of roads, S = total number of studied crashes in the group of roads, SPF(e) = estimated expected crash count on a segment for the exposure e, u_k = unit cost of crash of k severity, w = cost of crashes on road element, W = total cost of crashes in the group of roads, and v = variance of the m estimate.

APPENDIX B. ROAD CLUSTERING

The screening tool identifies which road elements experience an excessive number of crashes. Clustering these elements into longer road sections may reveal useful spatial regularities that may be useful to INDOT engineers in scoping corridor improvement studies and other safety-oriented programs.

It is important to note that elements with safety needs should be clustered based on the safety performance measures in order to obtain relevant road clusters from the safety management point of view. The following text describes the statistical basis of clustering and the clustering method itself.

STATISTICAL BASIS

There are three basic safety measures that can be used to identify road elements with excessive numbers of certain categories of crashes: crash frequency, crash rate, and crash proportion. Crashes are subject to a strong random fluctuation over time and two safety performance indices, Confidence F and Index I, are proposed to estimate the level of statistical confidence indicated in the detected excessive number of crashes as a systematic issue rather than the effect of random fluctuation.

Significance F is the probability of a safety level equal to or better than the one observed during the period of analysis if the expected safety level in the long run is average for the type of location and under the given exposure. The higher the significance of F is, the stronger the evidence is that the location experiences a real safety problem. Values of F=0.90 and higher are typically used.

Index I is the difference between the safety observed during the period of analysis and the safety expected given the location type and exposure divided by the standard deviation of the difference estimate. It is a simplified measure of Significance F. Values of I=1.5 and higher provide sufficient evidence that the location experiences a real safety problem.

Significance F is calculated as $\text{BetaDist}(1/(1+am), 1/a, c)$, while Index I is calculated as: $(c-m)/(c+am^2)^{1/2}$, where: a is the over-dispersion parameter, m is the average crash count in a long run, and c is the actual crash count in the period of analysis. Equations for calculating the values of m and a for different safety measures are shown in Table B.1.

TABLE B.1
Levels of statistical evidence

Statistical Evidence of Safety Problem	Confidence F_{CF}	Index I_{AF}
None or very weak	<0.80	<0.8
Weak	0.80–0.90	0.8–1.3
Considerable	0.90–0.95	1.3–1.7
Strong	0.95–0.99	1.7–2.7
Very strong	>0.99	>2.7

CLUSTERING METHOD

One of the important operations of clustering road elements is evaluation of the safety level in the current clusters to ensure that the obtained clusters experience excessive numbers of crashes. A practical method of updating safety evaluations in clusters is aggregation of the safety measures of the individual network elements included in the cluster. The exact method based on Significance F is statistically and computationally troublesome because summing two Gamma variables does not yield a Gamma variable, and the convenient equivalency between Negative Binomial and Beta distribution cannot be used. Therefore, Index I, which is easy to update for clusters, is calculated instead. The following equation is used to calculate Index I for clusters of multiple road elements:

$$I = \frac{\sum_i (c-m)_i}{\sqrt{\sum_i (c+am^2)_i}} \quad (\text{B.1})$$

where the values of $(c-m)_i$ and $(c+am^2)_i$ are known for any road element i . The clustering algorithm is shown in Figure B.1. It is important to note that the clustering process is controlled by two user-selected threshold values: I_1 and I_2 . The recommended ranges are: (1.25–2) for I_1 and (0–1.25) for I_2 with the recommendation that $I_1 > I_2$. The user can restrict the clusters' building only along the same routes to follow the common practice in scoping road studies. Other restrictions may be added to the algorithm as needed. A list of clusters and their elements is obtained based on the screening results, the network topology, and the parameters set by the user.

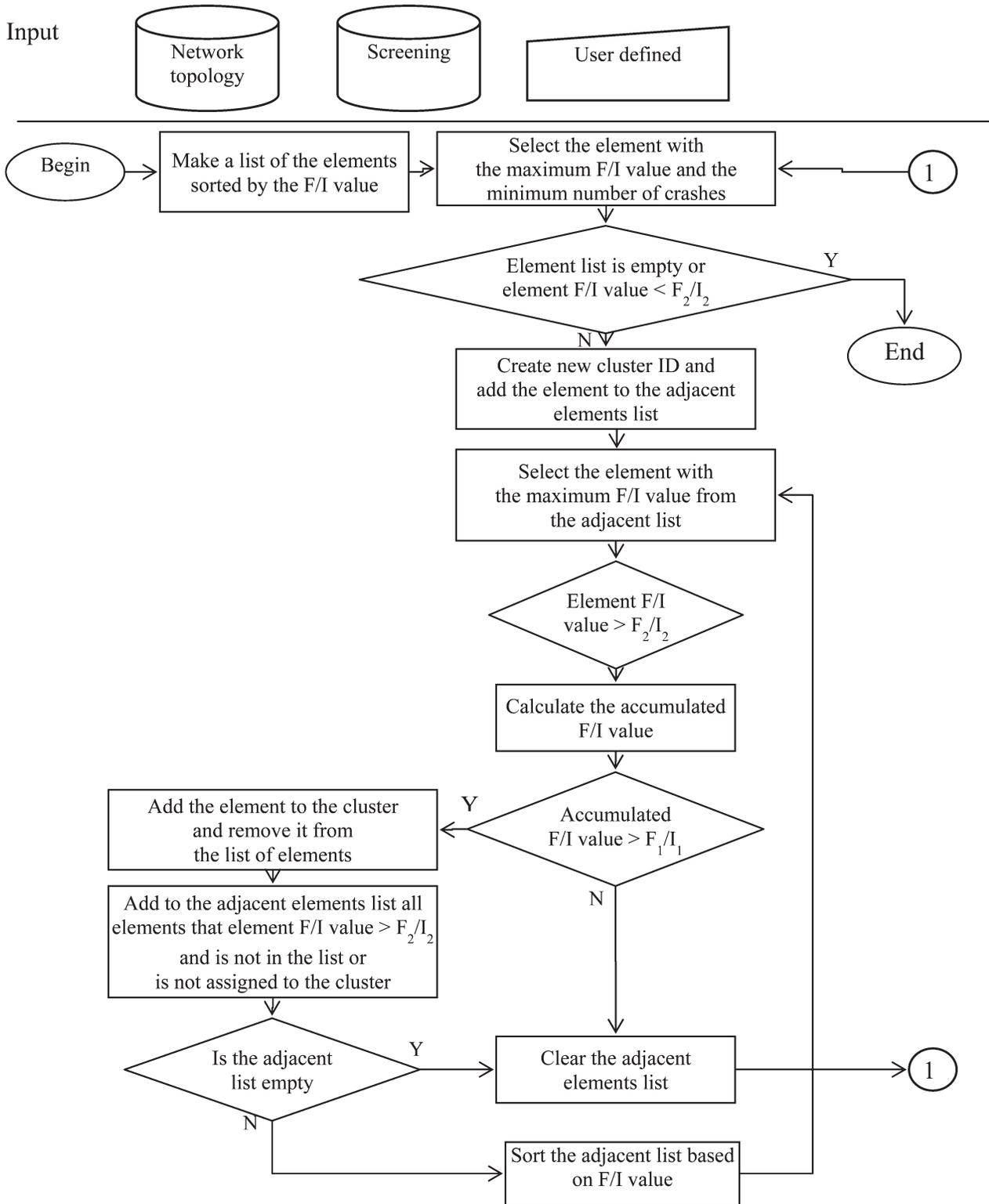


Figure B.1 Clustering algorithm flowchart.

APPENDIX C. RESULTS PRESENTATION

Result presentation is the final step of the screening process. In this phase, the user can visualize the results obtained in the standard screening or clusters/special studies. ArcGIS provides many visualization tools that can accomplish this job. Among the various features available in ArcGIS, the following three visualization tools are widely used:

- Symbology
- Labels
- Selection by attributes

Symbology refers to visualization of a feature (i.e., a single element), categories of elements, quantities, etc., by colors or symbols. Symbology has a special procedure for preparing

charts like bar charts or pie charts to be shown as part of the individual element. Figure C.1 shows a sample Symbology selection window.

Labels is useful in displaying a name or a value of a particular attribute on a map; for example, individual roads on a map can be labeled with their names to enhance visualization or to help identify a specific feature on the map. Figure C.2 shows a network layer with and without labels.

Selection by attribute can highlight particular elements of interest on a map. For example, intersections having more than 10 crashes per year can be easily selected and marked. Figure C.3 shows an example of visualization made by “select by attribute tool” in which the highlighted local roads were found to have signalization. The user can easily scan through the map once the features are selected. Details about the results display and visualization are discussed in the User Manual.

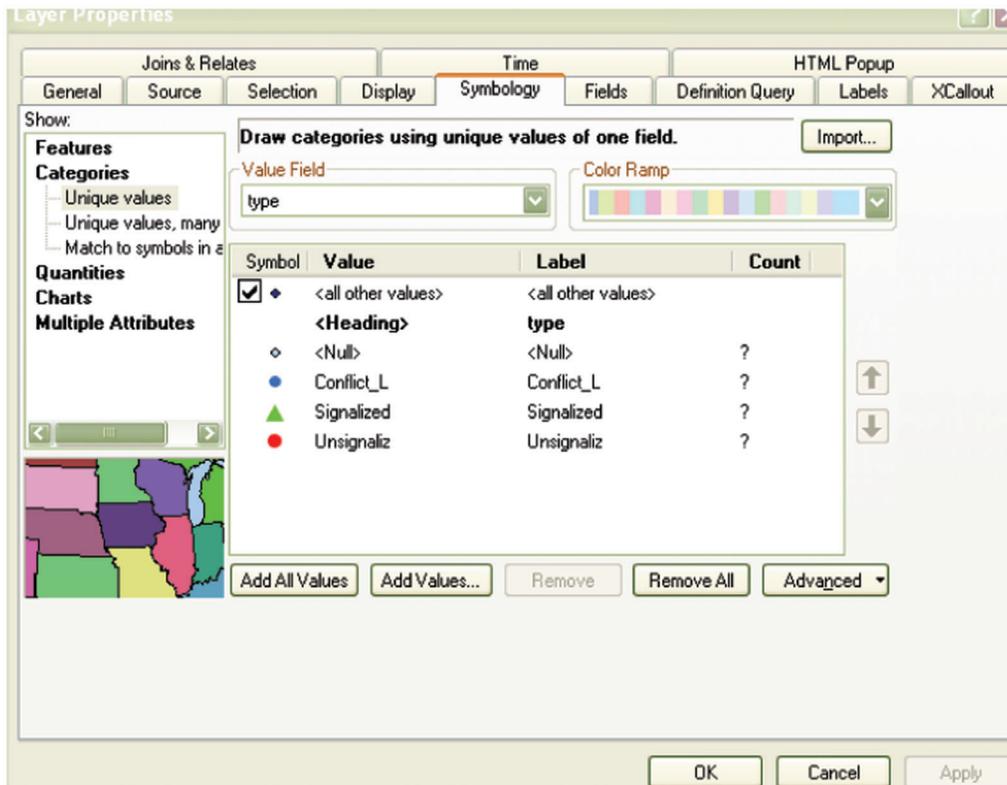


Figure C.1 Window in ArcGIS.



Figure C.2 Labeling feature (left: no labeling; right: labeling).

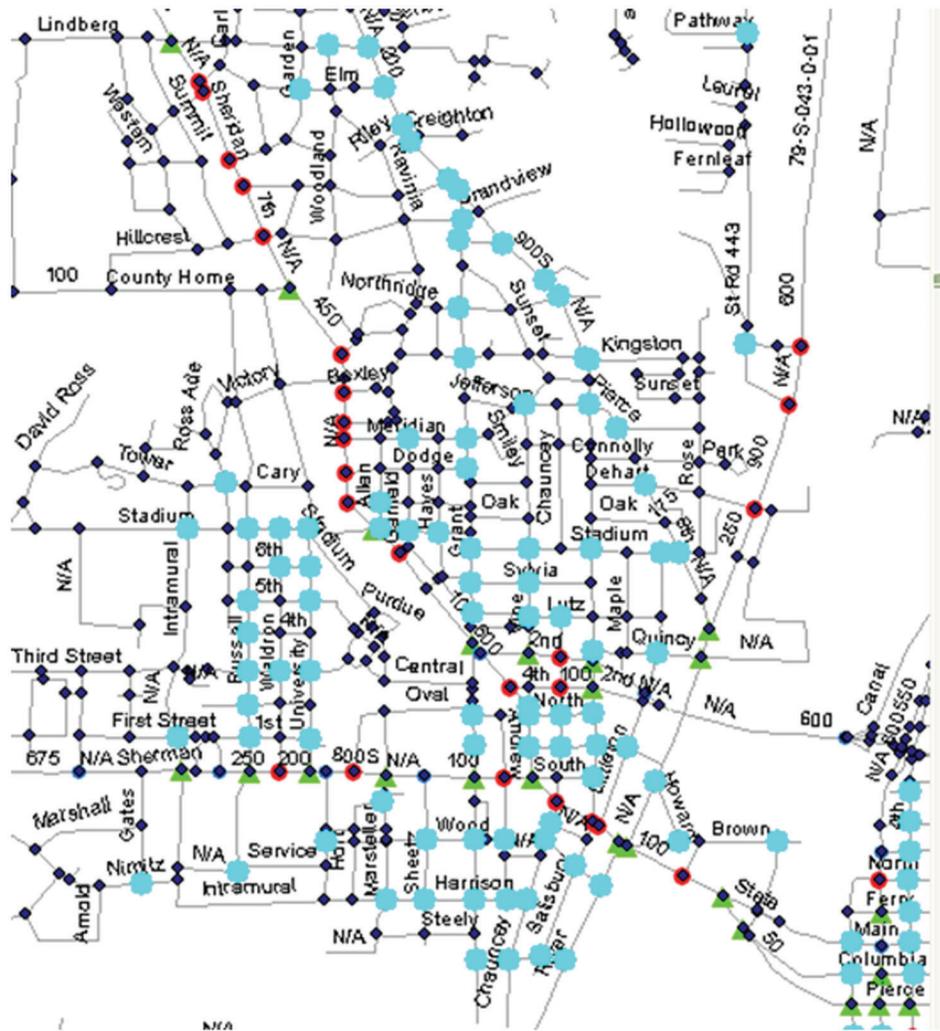


Figure C.3 Selection by attribute.

APPENDIX D. OPTIMIZATION

PROBLEM DEFINITION

Based on the available data, which includes the road information, the crash information, and the catalog of safety countermeasures, a mixed non-linear programming was established.

The objective is to maximize the safety benefit.

$$Max \sum_{i \in \{I\}} \sum_{h \in \{H\}} \sum_{j \in \{J\}} D_{ih} CCRF_{jh} CC_h \times R_{ij} X_{ij} \quad (D.1)$$

s.t.:

Total budget constraint

$$\sum_{i \in \{I\}} \sum_{j \in \{J\}} C_{ij} R_{ij} X_{ij} \leq TB \quad (D.2)$$

Countermeasure budget constraint

$$\sum_{i \in \{I\}} C_{ij} R_{ij} X_{ij} \geq B_j \quad (D.3)$$

Regional budget constraint

$$L_k \leq \sum_{i \in \{I_k\}} \sum_{j \in \{J\}} C_{ij} R_{ij} X_{ij} \leq U_k \quad (D.4)$$

Mutually exclusive constraint

$$\sum_{j \in \{J_m\}} X_{ij} R_{ij} \leq 1 \quad (D.5)$$

Where:

D_{ih} = crash frequency at location i at level of severity h ,
 $CCRF_{jh}$ = combined crash reduction factor for countermeasure j at severity h ,
 CC_h = crash cost for h severity (KA, BC, or PDO),
 R_{ij} = relevant countermeasure j at location i ,
 C_{ij} = cost at location i for project j ,
 TB = total budget,
 B_j = minimum budget for projects of type j ,
 U_k = upper bound budget for regional k ,
 L_k = lower bound budget for regional k ,
 I_k = location I belongs to regional k ,
 J_m = mutually exclusive countermeasures matrix, and
 X_{ij} = binary variable.

The formulation shown in Equations 2.1 through 2.5 is consistent with mixed non-linear programming. If several countermeasures were implemented together to enhance the safety benefit at some locations, it is intended to use combined crash reduction factors, and the equivalent is as follows (Lacy, 2001):

$$\begin{aligned} CCRF_{jh} &= 1 - [(1 - CRF_{1h}) * (1 - CRF_{2h}) * \dots * (1 - CRF_{jh})] \\ &= 1 - \prod_{j=1}^j (1 - CRF_{jh}) \end{aligned} \quad (D.6)$$

Equations 1 through 6 constitute a general formation that includes most of the practical situations but may change depending on the different circumstances at a given site.

ALGORITHM

The above formulations of the optimization problem were approximately solved by using the greedy heuristic search.

The optimization algorithm defines the principles of the heuristic method. The first step of the algorithm is to select the hazardous crash locations and then to create a list of potential safety countermeasures. Some countermeasures can be applied together, but some of them are mutually exclusive; therefore, the

list is presented in a 0–1 matrix. Next, the benefit for each countermeasure is calculated as shown in Appendix E. Then, the benefit/cost ratio is calculated for the safety countermeasures, and the highest B/C ratio is selected.

If a safety countermeasure satisfies all the constraints, then it will be implemented; however, if it violates some constraint, then it is deleted from the list. The algorithm continues running until no further countermeasures can be selected. The algorithm flowchart in Figure D.1 shows the steps of the optimization algorithm.

Using the developed tool, the safety countermeasures for the crash location can be identified. In accordance with the assigned constraints for different conditions, the optimizer estimates the expected total cost and the total benefit and calculates the benefit/cost ratio. In order to pictorially explain the process, a figure also appears in the interface. The optimal results are saved as a csv file.

ECONOMIC CALCULATIONS

Step 1. The required input: safety countermeasures; number of PD, NI, and IF crashes during the period with crash data; average AADT in the period with crash data; unit capital costs in Y_2 year dollars; service life of the countermeasures; interest rate; inflation rate; and unit costs of PD, NI, and IF crashes in Y_1 year dollars. The traffic is assumed fixed over the years (zero growth rate).

Step 2. Determine the crash reduction factors (CRF). In the case of multiple improvements, the $CRFs$ for the individual improvement are combined into one value:

$$CRF = 100 \cdot \left(1 - \Pi_n \left(1 - \frac{CRF_n}{100} \right) \right),$$

where:

CRF = total percent crash reduction factor for multiple improvements,

CRF_n = crash reduction factor for the n^{th} improvement.

Step 3. Calculate the expected the over-dispersion parameter D .

$$M = \frac{\sum_j A_j}{N}, \quad V = \frac{\sum_j (A_j - M)^2}{N - 1}, \quad R = \frac{\sum_j A_j}{Y \sum_j E_j}, \quad m_j = R \cdot Y \cdot E_j,$$

$$V_p = \frac{\sum_j (A_j - m_j)^2}{N - 1},$$

$$D = \frac{V_p - M}{M^2} \text{ if } D < 0 \text{ then assume } D = 0,$$

where:

M = average annual frequency of crashes in the group of roads,
 A_j = number of PD, NI, or IF crashes on road j during period with crash data,

Y = number of years in the period with crash data,

N = number of roads in the group of roads,

V = variance of crashes in the group of roads,

R = average crash rate in the group of roads,

E_j = exposure (average daily VMT, AADT, length L) that represents exposure on road j ,

m_j = expected number of PD, NI, IF crashes on road j ,

V_p = average squared residual $(A_j - m_j)$ where $m_j = R \cdot Y \cdot E_j$, and

D = over-dispersion parameter.

Step 4. Estimate the crash frequency in the period with crash data. The reported crashes and the expected crashes for the exposure are combined.

$$\alpha = R \cdot E$$

If $D = 0$ then $\hat{a} = \alpha$, otherwise

$$\hat{a} = \left(\frac{1}{D} + \alpha \right) / \left(\frac{1}{D\alpha} + Y \right),$$

where:

\hat{a} = best estimate of the PD, NI, or FI frequency (crashes/year),

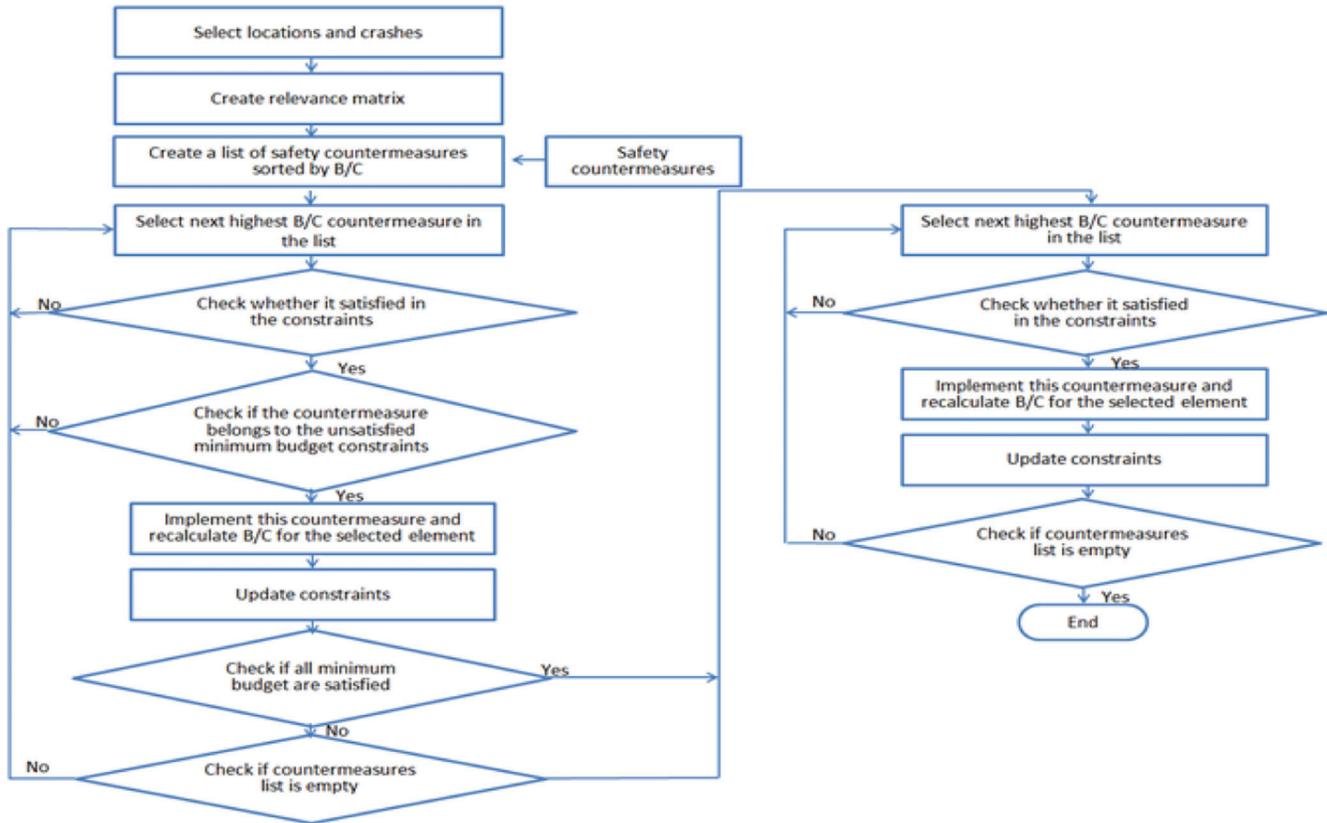


Figure D.1 Heuristic algorithm flowchart.

a = PD, NI, or FI frequency (crashes/year) expected at location,
 E = exposure measure,
 D = over-dispersion parameter,
 A = number of PD, NI, or FI crashes during the period with crash data (crashes in Y years), and
 $Y = (LY - FY + 1)$ number of years in the period with crash data (FY and LY are the first and last years of the period with crash data).

Step 5. Estimate the annual safety benefit, which is the product of the annual frequency a of PD, NI, or IF crashes, the CRF, and the crash cost adjusted for inflation.

$$B = a \cdot \frac{CRF}{100} \cdot C_1 \cdot \left(1 + \frac{F}{100}\right)^{PY - Y1},$$

where:

B = annual crash benefit for reducing crashes PD, NI, or FI,
 a = PD, NI, IF crash frequency,
 CRF = percent crash reduction factor of k severity,

C_1 = average cost of PD, NI, or IF crash in year Y1,
 F = inflation rate, assumed to be 2% unless otherwise specified, and
 PY = present year.

Step 6. Calculate the present worth of the total agency costs, which is the accumulated capital costs of all the improvements. The changes in the maintenance costs and the salvage values are neglected.

$$C = \frac{I}{100} \cdot \left(1 + \frac{F}{100}\right)^{PY - Y2} \cdot \sum_i \left(C_{2i} \cdot \frac{(1 + I/100)^{SL_i - 1}}{(1 + I/100)^{SL_i} - 1} \right),$$

where:

C = annualized countermeasure cost,
 C_{2i} = the capital cost of the i^{th} improvement in year Y2,
 I = interest rate, assume 4% unless specified otherwise, and
 SL_i = service life of the i^{th} improvement.

About the Joint Transportation Research Program (JTRP)

On March 11, 1937, the Indiana Legislature passed an act which authorized the Indiana State Highway Commission to cooperate with and assist Purdue University in developing the best methods of improving and maintaining the highways of the state and the respective counties thereof. That collaborative effort was called the Joint Highway Research Project (JHRP). In 1997 the collaborative venture was renamed as the Joint Transportation Research Program (JTRP) to reflect the state and national efforts to integrate the management and operation of various transportation modes.

The first studies of JHRP were concerned with Test Road No. 1—evaluation of the weathering characteristics of stabilized materials. After World War II, the JHRP program grew substantially and was regularly producing technical reports. Over 1,500 technical reports are now available, published as part of the JHRP and subsequently JTRP collaborative venture between Purdue University and what is now the Indiana Department of Transportation.

Free online access to all reports is provided through a unique collaboration between JTRP and Purdue Libraries. These are available at: <http://docs.lib.purdue.edu/jtrp>

Further information about JTRP and its current research program is available at: <http://www.purdue.edu/jtrp>

About This Report

An open access version of this publication is available online. This can be most easily located using the Digital Object Identifier (doi) listed below. Pre-2011 publications that include color illustrations are available online in color but are printed only in grayscale.

The recommended citation for this publication is:

Tarko, A. P., Li, M., Romero, M., & Thomaz, J. (2014). *A systematic approach to identifying traffic safety needs and intervention programs for Indiana: Volume II—SNIP2 user manual* (Joint Transportation Research Program Publication No. FHWA/IN/JTRP-2014/04). West Lafayette, IN: Purdue University. <http://dx.doi.org/10.5703/1288284315498>