Spring 2019

Data-driven framework for forecasting sedimentation at culverts

Haowen Xu
University of Iowa

Follow this and additional works at: https://ir.uiowa.edu/etd
Part of the Civil and Environmental Engineering Commons

DATA-DRIVEN FRAMEWORK FOR FORECASTING SEDIMENTATION AT CULVERTS

by

Haowen Xu

A thesis submitted in partial fulfillment
of the requirements for the Doctor of Philosophy
degree in Civil and Environmental Engineering in the
Graduate College of
The University of Iowa

May 2019

Thesis Supervisors: Adjunct Professor Marian Muste
Assistant Professor Ibrahim Demir

ACKNOWLEDGEMENTS

It was my great honor to study at the Department of Civil and Environmental Engineering (CEE) at the University of Iowa where I received an in-depth education in environmental engineering, and to work at the IIHR—Hydroscience & Engineering with so many world-renowned researchers and outstanding students. At IIHR, I was able to establish my career goals and receive a top-quality research training with hands-on experiences in water resources engineering. Many people have contributed to the completion of this multidisciplinary research.

First and foremost, I would like to express the deepest appreciation to my advisors: Professor Marian Muste and Professor Ibrahim Demir. I could not have completed my dissertation without their unfailing encouragement, valuable guidance, great patience, and continual support. I would especially like to express my sincere gratitude to my committee members for their assistance and support: Professor Larry Weber, Professor Witold Krajewski, and Professor Caglar Koylu.

Second, I am grateful to the Iowa Highway Research Board and the Iowa Department of Transportation for funding this research (Grant TR-655). Special thanks to our agency colleagues David Claman and Vanessa Goetz from the Iowa Department of Transportation. I would also like to thank the support staff at the Iowa Flood Center (IFC), IIHR, and the UI Department of Geography, without whom I would not have been able to finish my project and research. I especially appreciate the help and guidance from Professor David Bennett, Mr. Raymond Hammond, Ms. Jackie Stolze, and Mr. Radek Goska. Moreover, I would like to thank my family, especially my parents, for the financial support and life advice that sustained me through my Ph.D. research. In addition, thank you to everyone in the lab at IIHR for being such good friends and colleagues.

ABSTRACT

The increasing intensity and frequency of precipitation in recent decades, combined with the human interventions in watersheds, has drastically altered the natural regimes of water and sediment transport in watersheds over the whole contiguous United States. Sediment-transport related concerns include the sustainability of aquatic biology, the stability of the river morphology, and the security and vulnerability of various riverine structures. For the present context, the concerns are related to the acceleration of upland erosion (sediment production) and in-stream sediment-transport processes that eventually lead to sediment accumulation at culverts (structures that pass streams under roadways). This nuisance has become widespread in many transportation agencies in the United States, as it has a direct bearing on maintaining normal culvert operations during extreme flows when these waterway crossings are essential for the communities they serve. Despite the prevalence of culvert sedimentation, current specifications for culvert design do not typically consider aspects of sediment transport and deposition.

The overall study objective is to systematically identify the likelihood of culvert sedimentation as a function of stream and culvert geometry, along with landscape characteristics (process drivers of culvert sedimentation) in the culvert drainage area. The ideal approach for predicting sedimentation is to track sediment sources dislocated from the watershed, their overland movement, and their delivery into the streams using physical-based modeling. However, there are considerable knowledge gaps in addressing the sedimentation at culverts as an end-to-end process, especially in connecting the upland with in-stream processes and simulating the sediment deposition at culverts in non-uniform, unsteady flows, while also taking into account the vegetation growth in culverts' vicinity. It is, therefore, no surprise that existing research, textbooks, and guidelines do not typically provide adequate information on sediment control at culverts.

This dissertation presents a generalizable data-driven framework that integrates various machine-learning and visual analytics techniques with GIS in a web-based geospatial platform to explore the complex environmental processes of culvert sedimentation. The framework offers systematic procedures for (1) classifying the culvert sedimentation degree using a time-series of aerial images; (2) identifying key process-drivers from a variety of environmental and culvert structural characteristics through feature selections and interactive visual interfaces; (3) supporting human interactions to perceive empirical relationships between drivers and the culvert sedimentation degree through multivariate Geovisualization and Self-Organizing Map (SOM); and (4) forecasting culvert sedimentation potential across Iowa using machine learning algorithms. Developed using modular design and atop national datasets, the framework is generalizable and extendable, and therefore can be applied to address similar river management issues, such as habitat deterioration and water pollution, at the Contiguous US scale.

The platform developed through this Ph.D. study offers a web-based problem-solving environment for a) managing inventory and retrieving culvert structural information; b) integrating diverse culvert-related datasets (e.g., culvert inventory, hydrological and land use data, and observations on the degree of sedimentation in the vicinity of culverts) in a digital repository; c) supporting culvert field inspections and real-time data collection through mobile devices; and d) hosting the data-driven framework for exploring culvert sedimentation drivers and forecasting culvert sedimentation potential across Iowa. Insights provided through the data-driven framework can be applied to support decisions for culvert management and sedimentation mitigation, as well as to provide suggestions on parameter selections for the design of these structures.

PUBLIC ABSTRACT

Sediment transport through culvert structures has been recognized as a problem for many years. The variety and complexity of the problem of sediment passing continues to be a challenge at stream crossings provided with culverts. In general, current knowledge on sedimentation processes at culverts is fragmented and the literature on this topic is scarce. More recently, the intensification of land use changes (through agriculture and urbanization) and the impact of climate change make the issue a critical area of research. Since culvert sedimentation entails complex and interlinked environmental processes that are difficult to investigate and solve with conventional approaches (e.g., laboratory-based experimental methods, and erosion and sediment transport models), it is no surprise that systematic strategies and practical suggestion for preventing and mitigating detrimental effects of culvert sedimentation do not exist at the present time.

Developing an effective strategy for mitigating culvert sedimentation requires a thorough understanding of its drivers and the capability to forecast the process under various scenarios. Thus, a more holistic and system-based approach is proposed through this dissertation by taking advantage of the capabilities of data-driven modeling. The main goal of this study is to develop a generalized data-driven framework that combines the machine learning techniques with information visualizations to explore the complex environmental processes of culvert sedimentation. The framework is then embedded into a web-based geospatial platform, labelled "IowaDOT Culverts" aimed at disseminating data-driven insights acquired through this research (e.g., culvert sedimentation potential vs. its key drivers) as well as supporting routine culvert management and design across Iowa.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

ANN	Artificial Neural Network
API	Application Program Interface
DW	Digital Watershed
EPA	United States Environmental Protection Agency
FHWA	Federal Highway Administration
GIS	Geographic information system
GUI	Graphic User Interfaces
HDS-5 Hydraulic	Design Series No.5
HUC	Hydrologic Unit Code
IIHR	Iowa Institute of Hydraulic Research
Iowa DOT	Iowa Department of Transportation
MCDA	Multiple-Criteria Decision Analysis
NHD	National Hydrography Dataset
OGC	Open Geospatial Consortium
PCP	Parallel Coordinates Plot
POI	Point of Interest
PSE	Problem Solving Environment
RUSLE	Revised Universal Soil Loss Equation
SFM	Structure from motion
SIIMS	Structure Inventory and Inspection Management System
SOM	Self-Organizing Map
StreamCAT	Stream-Catchment dataset
UAV	Unmanned aerial vehicle
UDOT	Utah Department of Transportation
WBD	Watershed Boundary Dataset
WFS	Web Feature Services
WMS	Web Map Service
WMTS	Web Map Tile Service
WMTS	Web Map Tile Service

CHAPTER 1 INTRODUCTION

1.1 Problem statement

U.S. Midwestern secondary roads often rely on culverts to pass streams under roadways, therefore playing an important role in statewide transportation infrastructure. There are various types of culverts with different purposes and which to use depends on the culvert site and the characteristics of its drainage area. In general, larger flows and road embankment heights entail the use of multi-barrel (a.k.a. multi-box) culverts. Multi-box culverts require less headwater and are more economical than a larger, single-box culvert. Given that the box culverts are typically designed to handle 50-year return period flows, in many areas of Iowa, and indeed elsewhere, water flow through a typical box culvert is relatively low throughout most of the year. Ideal culvert operation is indicated by stable stream geometry in the culvert vicinity as illustrated in Figure 1.1.

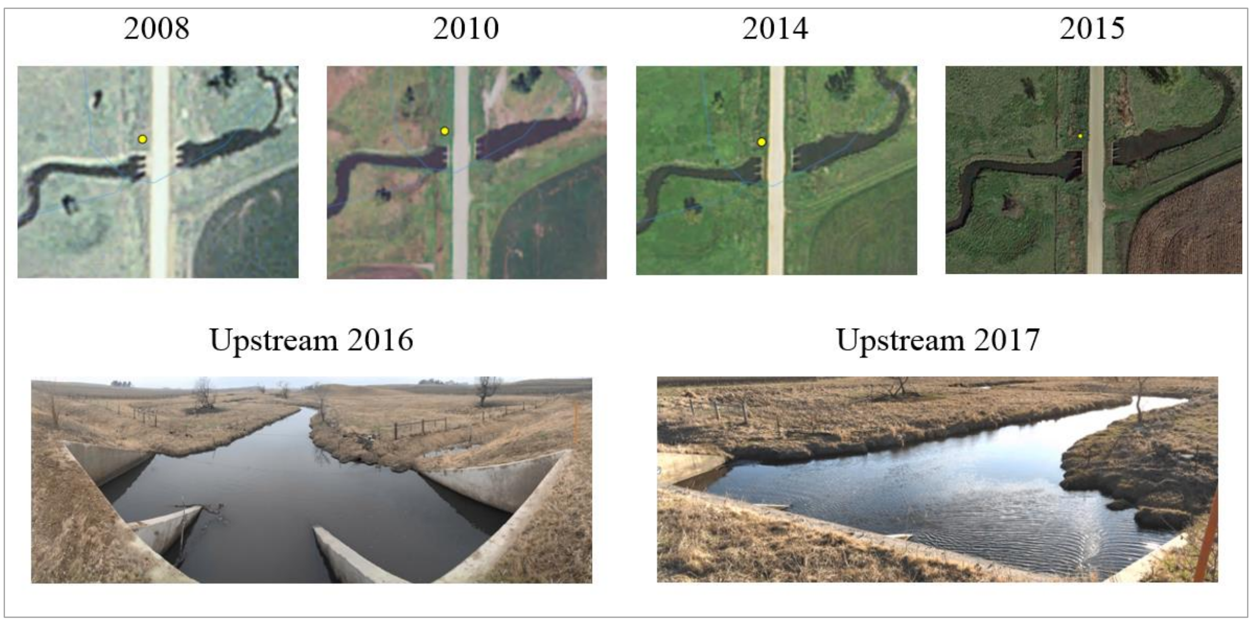


Figure 1.1 Ideal culvert operations

While culverts are commonly sized to accommodate specific return period flows (i.e., 25, 50, or 100 years) depending on the structure type, there is evidence that culvert failures are rarely related to the exceedance of some level of flood flow (Cafferata et al., 2017). Instead, accumulations of debris and sediment at the culvert inlet that partially block the culverts are more often the underlying cause of the failures. In recent decades, the increasing intensity and frequency of precipitation, combined with the human interventions in watersheds, has drastically altered the natural sediment regime in watersheds over the whole contiguous US (Solomon et al. 2007). Sediment transport related concerns include the sustainability of aquatic biology, the stability of river morphology, and the security and vulnerability of various riverine structures (CEI 2018; Tetreault et al., 2018). More specifically, the concerns are related to the acceleration of upland erosion (sediment production) and in-stream sediment-transport processes (i.e., aggradation, scouring, and deposition), which eventually lead to sediment accumulation at culverts.

Multi-box culverts located in flat, erodible watersheds are especially prone to sediment blockage. This is the case for the highly erosive Iowa landscape where sediment accumulation can lead to situations whereby some of the barrels can become partially filled, as illustrated in Figure 1.2. Two surveys were conducted by engineers from the Iowa Department of Transportation (IDOT), counties, and IIHR-Hydroscience & Engineering (IIHR) in 2009 and 2013. Both surveys indicate that more than 95% of Iowa culverts are silted (Ho, 2010; Muste and Xu, 2017). The culvert sedimentation concern is widespread in the nation, from California to Pennsylvania, from Wisconsin to Florida (Rowley, 2014), which has a direct bearing on the ability to maintain normal culvert operation during extreme flows when waterway crossings are essential for the communities they serve.

Figure 1.2 Example of silted culvert (structure code: 165881) in South-West Iowa

Sedimentation (a.k.a., silting) of culverts can considerably reduce their capability to handle larger flow events as the partial blockage of the structure could severely impair the hydraulic capacity to convey design flows. Obstruction of the culvert inlet can cause severe damage to both the transportation structure (due to culvert and road overtopping) and flooding and/or river diversion in the upstream headwater areas, as shown in Figure 1.3. The socio-economic damages associated with culvert sedimentation are unlikely to diminish, as recent studies predict that the frequency and intensity of storms continue to increase throughout the contiguous United States (Villarini et al. 2013).

Damages to culvert structure

(image sources: http://www.roadex.org and http://riversmartvt.umass.edu)

Flooding of the culvert upstream area

(image sources: http://twitter.com/harpersenv and http://www.vicksburgpost.com)

Figure 1.3 Detrimental effects of culvert sedimentation

The information assembled through this research has revealed that the process of sedimentation at culverts can attain a stable form of sediment deposits in no more than four or five years, as illustrated in Figure 1.4. The fast-evolving sedimentation process was observed at many of the Iowa culverts by inspecting the aerial photographs collected over successive years. The severity of the culvert blockage, as well as its fast development in time, requires sustained cleaning operations, as currently, the sedimentation at culverts does not benefit by well-established mitigation solutions. The cleaning of the culverts is one of the costliest maintenance problems for Iowa culverts due to the frequency with which culverts need to be cleaned, the inaccessibility of the sloping terrain, and the man hours required.

Figure 1.4 Sediment deposition at culverts is a fast process, impairing culvert's conveyance capacity in few years from construction completion (structure code: 344691)

1.2 Challenges and motivation

Sediment transport through culvert structures has been recognized as a problem for many years (Haight, 1912). The variety and complexity of the problem of sediment passing continues to be a challenge at stream crossings provided with culverts. In general, current knowledge on sedimentation processes at culverts is fragmented and the literature on this topic is scarce. More recently, however, the intensification of land use changes (through agriculture and urbanization) and the impact of climate change make this a critical area of research. Since culvert sedimentation entails complex and interlinked environmental processes that are difficult to investigate and solve with conventional approaches (e.g., laboratory-based experimental methods, and erosion and sediment transport models), it is no surprise that systematic strategies and guidelines for preventing and mitigating detrimental effects of culvert sedimentation do not exist at the present time.

The work to understand and mitigate sedimentation encounters many challenges due to its process-related complexity (Gregory, 2006; Ho, 2010). At the conceptual level, culvert sedimentation processes can be grouped into three major stages, namely (1) soil particle detachment (erosion and sediment production); (2) movement of sediment toward streams (sediment transport); and (3) sediment deposition at culvert structures (Merritt et al., 2003; Haan, 1994; Muste and Xu, 2017). The complexity of the problem is further increased by fast-evolving anthropogenic activities that are interacting with natural processes occurring in the culverts' drainage area, continuously altering the hydrologic behaviors and erosion-process drivers in watersheds. Additionally, the presence of multi-box culverts modifies the natural stream geometry by introducing a channel transition area in the culverts' vicinity (Ead et al., 2000), which can further complicate the investigation on sediment deposition near culverts. Depending on the characteristics of culvert geometry, either flow expansion or flow contraction can occur near culvert inlet, creating a favorable condition for sediments to deposit near culvert structures. This situation is especially common for the culverts located in watersheds with poor soil protection and intensive agricultural activities.

It is apparent that the complexities of the watershed sediment dynamics continuously shaped by anthropogenic impacts exceed the problem-solving capabilities of the available experimental, analytical, or numerical simulation-based investigations. In order to provide watershed managers and structural engineers in charge of sedimentation at culverts with information needed to make decisions, an alternative end-to-end analysis and associated support tools are proposed. Specifically, a data-driven framework, embedded in a web-based Problem Solving Environment (PSE), is developed to provide the critical information for planning, designing, and maintaining culverts that are also compatible with the natural stream environment. The proposed method maintains the holistic, systems-based approach of the problem investigation but it does so with low-cost and effective means.

1.3 Study objectives

The main goal of this study is to develop a generalized data-driven framework that combines the machine learning techniques with analytical indicators to investigate highly complex transport problems in the natural environment. The framework is modular, scalable to the Contiguous United States, and can be adapted to address other river management issues, such as habitat deterioration and water pollution. The method can only be applied in data-rich watersheds or in areas where surrogates for that data are available. In this study, we test the feasibility of the developed framework in conjunction with the issue of culvert sedimentation in Iowa.

The overall objectives of this research are to test the capabilities of an integrated data-driven framework to: a) investigate the causes of culvert sedimentation and their interdependencies, b) forecast the degree of culvert sedimentation across Iowa, and c) provide data-driven insights for effective culvert sedimentation mitigation. In addition, a geospatial platform, named as the "IowaDOT Culverts", is developed to provide a web-based problem-solving environment for hosting the data-driven framework and disseminating analytical reports. The above-mentioned research objectives are attained through the following research phases:

Phase.1 Review of past studies for justifying the usage of data-driven approach

In this research phase, the physical processes leading to culvert sedimentation are reviewed first to enable the identification of the most efficient strategy for solving the problem. Subsequently, a critical review of the prior studies relevant to culvert sedimentation mitigation is carried out to identify the knowledge gaps and challenges in this area. As most of the available studies are based on erosion and sediment transport numerical simulations, the merits and limitations of a data-driven approach are evaluated to justify the reasoning for proposing the use of the latter for completion of this study. Finally, using the parameterization of the culvert sedimentation mechanism, the environmental and culvert structural characteristics are identified and critically selected to guide the design of the data-driven approach implementation.

Phase.2 Integration of Culvert Information with Watershed Characteristics

In this step, a comprehensive environmental dataset that contains all independent and dependent variables is assembled. The independent variables consist of characteristics of the culvert drainage relevant to sedimentation processes (e.g., hydrologic, geomorphological, and ecological) while the culvert structural attributes control the local aspects of the sediment depositions near the culvert (e.g., geometry, slope, and hydraulic control). The drainage area characteristics are harvested from multiple federal and state datasets while that dependent variable quantifies the sedimentation degrees (severity) at culverts. The sedimentation degree is also determined from third-party sources, i.e., time-series of aerial images of the culvert vicinity. The above variables are combined spatially to create a comprehensive culvert dataset that includes 300 independent variables.

Phase.3 Conduct of Multiple-Criteria Decision Analysis

The concept of Multiple-Criteria Decision Analysis (MCDA) is utilized to support the decisions for mitigating culvert sedimentation by simultaneously evaluating multiple drivers and their effects on the sedimentation. At the technical level, the integration of various machine learning and visual analytics techniques enable to a) reduce the complexity and dimensionality of the dataset, b) improve human perception of the relationships between the drivers and culvert sedimentation degree, and c) forecast culvert sedimentation degree at other sites in real-time.

Phase.4 Optimization of sedimentation forecasting

The forecasting developed in Phase 3 is further optimized by: a) conducting a sensitivity analysis on different spatial units (e.g. catchment, corridors, watersheds, riparian buffers); b) regionalization of the erosion potential through multivariate clustering and spatial partitioning for reducing spatial heterogeneity (e.g. sediment yield, transport process, soil types), and, c) exploring the multivariate relationship between key drivers and the culvert sedimentation degree through geospatial visualization and multivariate clustering analysis using self-organizing maps.

The major contributions of this thesis work include the following:

- Improvement of the current understanding of culvert sedimentation mechanisms and its spatial variabilities across Iowa
- Assembling machine learning and visual-analytics tools to enable forecasting and heuristic exploration of a complex watershed-scale process using data-driven modeling

- Creation of an interactive web-portal for supporting investigations and decision making on mitigating the unfavorable effects of sedimentation at culverts. Currently, the "IDOT Culverts" web-portal is transferred onto the Iowa DOT's servers to support routine culvert management and sedimentation mitigations

1.4 Outline of chapters

The phased research discussed above is mapped in the dissertation through four core chapters succinctly presented below. Chapter 2 provides comprehensive literature review to identify the existing knowledge gaps, evaluate the feasibility of the data-driven approach, and examine detailed physical processes involved in culvert sedimentation, so as to guide the selection of environmental and culvert structural parameters that can be analyzed as independent variables during the data-driven investigations. A summary of individual variables with their associated processes and data sources are provided as tables in this chapter. The chapter consists of three subsections: 2.1) Review of culvert sedimentation processes, 2.2) State of knowledge in sedimentation studies, and 2.3) Data-Driven approaches for erosion and sediment transport.

Chapter 3 presents the methodology of the data-driven framework through four subsections: 3.1) Overview, 3.2) Data compilation & integration, 3.3) Multiple-Criteria Decision Analysis, 3.4) Culvert sedimentation forecasting, and 3.5) Sediment forecast optimization. In Section 3.2, the study describes in detail the creation of a comprehensive environmental dataset that contains all independent and dependent variables. Section 3.3 and 3.4 demonstrate the preliminary analyses and forecasting in this study through multivariate visualization and a random forest algorithm. In Section 3.5, a series of optimization tasks that aim to enhance the investigation of the sedimentation-related drivers, as well as to identify the regional culvert sedimentation potential across Iowa.

Chapter 4 describes the design and development of a web-based geospatial platform, the "Iowa DOT Culverts" management system. The platform assembles and stores culvert-related data, offers an online problem-solving environment served by data-driven workflows and is equipped with visual interfaces that facilitates human-computer interaction for exploring the complex problem of culvert sedimentation. The chapter entails the following subsections: 4.1) Overview, 4.2) Iowa DOT culverts platform architecture, software, and technologies, 4.3) Cyberinfrastructure for data integration, and 4.4) Workflows.

Chapter 5 discusses the results and inferences of this Ph.D. study through two sections. Section 5.1 demonstrates the preliminary results of the MCDA (the early stage of the data-driven analysis based on decision tree and data visualization). Section 5.2 presents the outcomes of the optimized sedimentation forecasting following the implementation of the sensitivity analysis, regionalization, and multivariate clustering analysis. The section ends with illustrations of several culvert sites where the sedimentation forecast is affected by special local features.

Chapter 6 presents the major conclusions of this study and suggests future research directions. Contents in many chapters of this thesis are authorized replications of the IHRB TR-655 report (Muste and Xu, 2017), authored by my advisor Dr. Marian Muste and myself, as the funding for this Ph.D. study was provided by the Iowa Highway Research Board, and Iowa Department of Transportation (Iowa DOT), Grant TR-655.

1.5 References

Cafferata, P., Lindsay, D., Spittler, T., Wopat, M., Bundros, G., Flanagan, S., Coe, D., and Short, W. (2017). "Designing watercourse crossings for passage of 100-year flood flows, wood, and sediment", Department of Forestry and Fire Protection, California Forestry Report No. 1, Pebble Beach, CA.

Comprehensive Environmental Inc. (CEI) (2018). "Flood & culvert management resources". Retrieved on November 11, 2018 from http://ceiengineers.com/innovations/resources-br-innovations/general-br-resources/flood-culvert-management.

Ead, S. A., Rajaratnam, N., Katopodis, C., and Ade, F. (2000). "Turbulent open-channel flow in circular corrugated culverts", Journal of Hydraulic Engineering. 126 (10), 750–757.

Gregory, K. J. (2006). "The human role in changing river channels", Geomorphology 79 (3–4), 172-191.

Haan, C. T., Barfield, B. B., and Hayes, J. C. (1994). "Design hydrology and sedimentology for small catchments", San Diego Academic Press, San Diego, CA, USA.

Haight, W. H. (1992). "Culvert", Google Patents-1912, Retrieved on November 20, 2018 from www.google.sr/patents/US1048153.

Ho, H-C. (2010). "Investigation of unsteady and non-uniform flow and sediment transport characteristics at culvert sites", Ph.D. thesis, Civil & Environmental Engineering, the University of Iowa, Iowa City, Iowa, USA.

Merritt W. S., Letcher R. A., Jakeman A. J. (2003). "A review of erosion and sediment transport models", Environmental Modelling & Software 18 (8): 761-799.

Muste, M. and Xu, H. (2017). "Sedimentation mitigation using streamlined culvert geometry", Report ST-001, Iowa Department of Transportation, Statewide Transportation Innovation Council, Federal Highway Administration, McLean, VA.

Muste, M. and Xu, H. (2017). “Mitigation of sedimentation at multi-box culverts”, IIHR Report No. TR-655, Submitted to the Iowa Highway Research Board, Ames, IA, USA.

Rowley, K. J. (2014). “Sediment transport conditions near culverts”, M.S. Thesis Dissertation, Brigham Young University, Provo, UT.

Solomon, S. D., Qin M., Manning, Z., Chen, M., Marquis, K. B., Averyt, M. T., and Miller, H. L. (2007). “Contribution of working group i to the fourth assessment report of the intergovernmental panel on climate change”, Cambridge United Kingdom and New York NY USA: IPCC.

Tetreault, J., Moore, I. D., Hoult, N. A., Tanzil, D., Maher, M. L. J. (2018). “Development of a sustainability evaluation system for culvert replacement and rehabilitation projects”, Journal of Pipeline Systems Engineering and Practice 9 (2): 1949–1204.

Villarini, G., Scoccimarro, E., Gualdi, S. (2013). “Projections of heavy rainfall over the central United States based on CMIP5 models”, Atmospheric science letters 14 (3): 200–205.

CHAPTER 2 LITERATURE REVIEW

2.1 Review of culvert sedimentation processes

Conducting an effective investigation of culvert sedimentation requires a sound and holistic understanding of the physical processes involved. Conceptually, the processes associated with culvert sedimentation can be grouped into three major categories as illustrated in Figure 2.1: (1) soil detachment (i.e., erosion, sediment supply, and sediment production), (2) sediment transport (overland and in-stream), and (3) sediment deposition (settling at culvert structure and stabilization due to vegetation growth). This grouping tracks the sediment dynamics in the watershed river network continuum as depicted in Figure 2.2, whereby the sedimentation sources (in drainage basins), transport pathways (through watersheds and stream network), and receptors (at basin outlets or downstream culverts) are all connected in an end-to-end system (Haan, 1994; Schumm, 1977; Merritt et al., 2003, and Lord et al., 2009).

a) b) c)

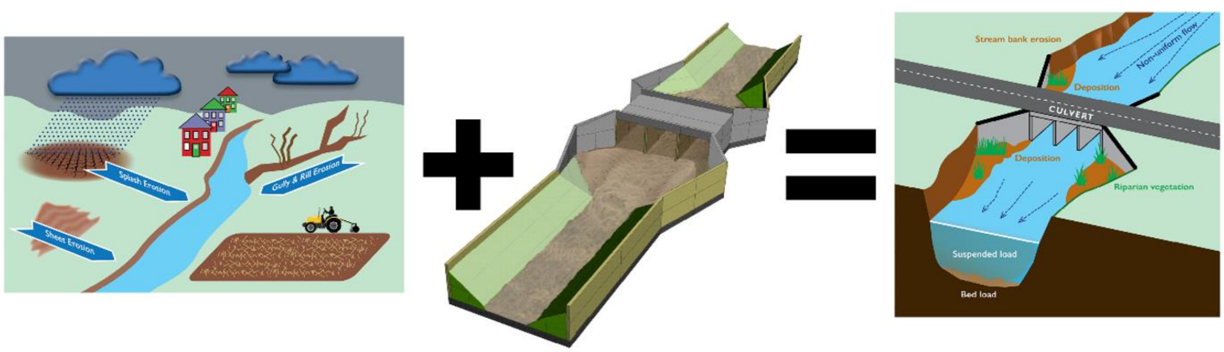


Figure 2.1 End-to-end culvert sedimentation processes: a) soil detachment (sediment production and overland transport); b) in-stream sediment transport to culvert vicinity; and c) deposition processes in the vicinity of the culvert (Muste and Xu, 2017)

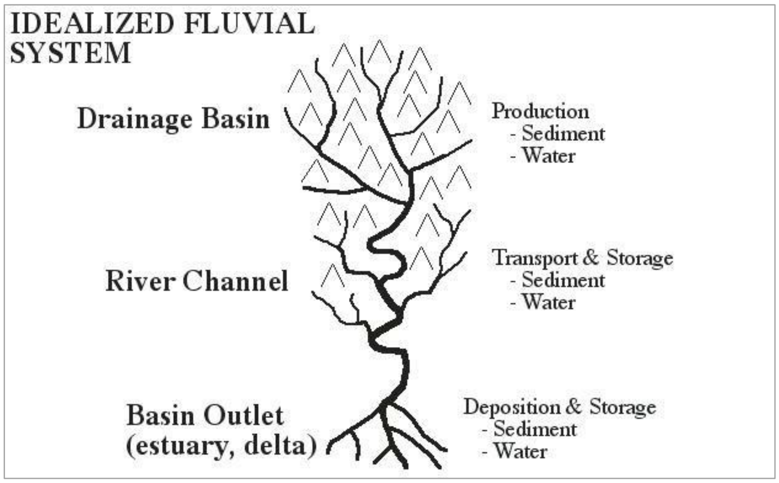


Figure 2.2 Dynamics of an idealized fluvial system (Lord et al., 2009; Schumm, 1977)

Each stage of culvert sedimentation entails a number of physical processes that are independently reviewed in the Subsections 2.1.2 to 2.1.4. Along with their review, the relevant contributing factors (environmental and structural) are identified for each sedimentation stage.

2.1.1 Sediment production

The detachment of soil particles within culvert drainage basins is considered as the source for sediment, which through downstream transport, eventually lead to sediment deposition near culvert structures. This process is also labeled as erosion, sediment generation, or sediment production by different studies. In principle, soil detachment is driven primarily by raindrop impacts and overland flow (Hudson, 1975; Loch and Silburn, 1996) as a two-phase process that entails the detachment of individual soil particles and their transport by erosive agents (Morgan, 2005). Although soil detachment is a natural process that causes mobilization, transport and off-site sedimentation of soil particles in undisturbed landscapes, non-sustainable soil erosion rates are usually results of human activities (Fernández-Raga et al., 2017), such as agriculture, urbanization, and mining. These anthropogenic activities not only alter the pristine hydrologic behaviors of watersheds, but also reshape the landscape surface and natural drainage waterways, causing stream instability and degradation of land. In Iowa, one of the land management challenges is the considerable increase of agricultural land use, which involves the removal and/or alteration of the native ground cover, resulting in changes in the roughness of the landscape surface (through tilling and other practices) (Muste and Xu, 2017). Recent studies reveal that streams in Iowa carried a larger sediment load in the early twentieth century, followed by a drop and stabilization in loads during the recent times that is merely the reflection of the alterations brought in the Iowa natural landscape (Jones and Schilling, 2011).

Depending on their mechanism and location in the drainage basin, the erosion processes can be further categorized into six major types: splash, sheet and rill, gully, in-stream erosion, landslides, and construction site washouts (Emmett, 1978; Foster, 1982; Hairsine & Rose, 1992; Merritt et al., 2003). These processes vary widely both spatially and temporally, may occur in isolation from one another, or can be linked through a splash–sheet–rill–gully erosion sequence (illustrate in Figure 2.3). According to Merritt et al. (2003), the potential of occurrence for different types of erosion is related to both landscape and rainfall characteristics.

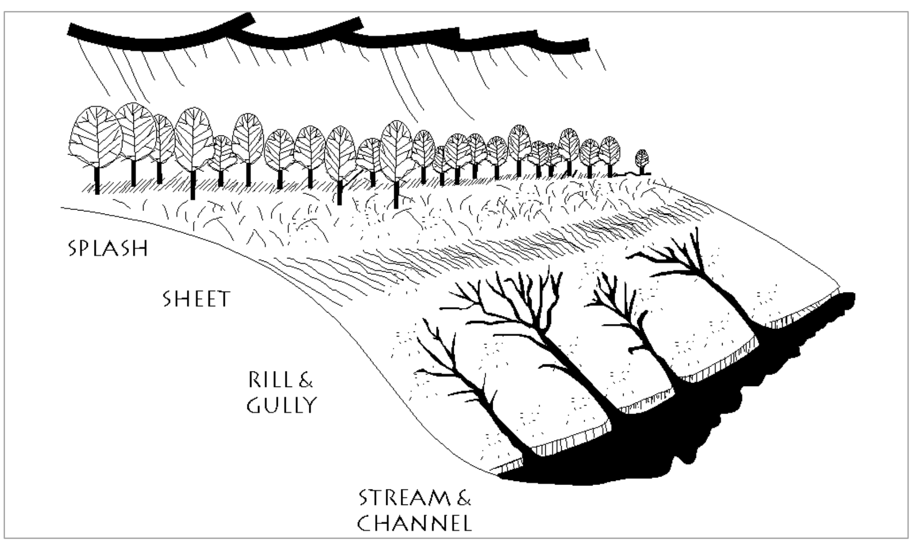


Figure 2.3 Major types of soil erosion (UNEP, 1998)

Given that for Iowa conditions, the culvert sedimentation can be supplied by any of the above-mentioned types of erosion, this study shortly reviews them below.

1. Splash erosion refers to the dislocation of the bare soil surface by raindrops falling on the ground. The raindrop impact destroys soil structure and causes particles to be transported for a short distance (Angulo-Martínez et al., 2012). The process is recognized as the first phase in the soil erosion, causing soil detachment and disintegration so that individual soil particles can be subsequently transported by erosive agents (Fernández-Raga et al., 2017).
2. Sheet and rill erosion is defined as the uniform detachment and removal of sediment particles by overland flow or raindrop impact that are evenly distributed across a slope, followed by the formation of soil surface flows along preferential pathways (Hairsine and Rose, 1992: Rose, 1993). According to Merritt et al. (2003), sheet and rill erosion are often classified as 'overland flow' erosion, and are significant in agricultural watersheds.
3. Gully erosion refers to the formation of channels of concentrated flow that are too deep to be obliterated by cultivation (Rose, 1993; Loch and Silburn, 1996). In general, gully erosion forms deeper channels, and its flow also differs from sheet and rill flows in that raindrop impact is not a critical factor for sediment particle detachment (Bennett, 1974). According to Carey (2006), gully erosion is a significant sediment supplier for culvert and road sedimentation, and its occurrence is affected by land cover (e.g., the presence of vegetation prevent gully formation) and topography (e.g., steepness of channel).
4. In-stream erosion involves the direct removal of sediment from the stream banks (lateral erosion) or the stream bed (Merritt et al., 2003). The process is usually reflected by widening and deepening of the stream channel, and is not recognized as a problem unless the channel stability is compromised due to excessive erosion rate. However, certain

human activities (e.g., stream channelization) that involve straightening of the natural course of a channel, can considerably jeopardize channel stability and lead to excessive rates of streambank erosion. Lord et al. (2009) suggests that the degrees of in-stream erosion can be indirectly evaluated through vital signs of stream channels (longitudinal profile and planform) which can be monitored by examination of historic aerial photographs and maps (Lord et al, 2009).

5. Landslide, and construction sites and road erosion are localized erosion processes that dislocate a large amount of sediment by exposing the disturbed land to both direct raindrop impact and direct transport. These processes are not relevant for the culvert sedimentation in Iowa, as the overall the topography is mild (mean slope below 25 degrees) and the state is not heavily urbanized (Muste and Xu, 2017). Therefore, this study does not consider them as critical sources of sediment sources.

Given that the present study is based on a data-driven investigation, the availability of the data relevant for each process is essential. From this perspective, the literature review presented in these sections includes identification of the relevant parameters associated with each of the erosion process and of the most complete source of information providing the data need for the analysis. A summary of the sediment production processes described above along with the associated parameters and the sources of data documenting these parameters are provided in Table 2.1.

Table 2.1 Parameters that are related to soil detachment processes

Erosion process	Associated parameter	Data source(s)
Splash erosion	Rainfall erosivity	RUSLE
Sheet and rill erosion	Soil erodibility	RUSLE
Gully erosion	Land cover	StreamCAT
Gully & in-stream erosion	Slope-length factor	RUSLE
Gully & in-stream erosion	Channel slopes	NHD Plus
Streambank erosion	Stream channelization	NHD Plus & aerial images
Streambank erosion	Stream sinuosity	NHD Plus & aerial images

2.1.2 Sediment transport

Following the production of free sediment, the dislocated material is transported over the watershed surface and, much more efficiently, through the stream network. Most of the sediment materials generated from soil erosion processes are trapped during overland and in-stream transport, thus are unable to reach far downstream. Conventionally, sediment yield and Sediment Delivery Ratio (SDR) are used to characterize the sediment transport behavior in watersheds. Sediment yield is defined as the quantity of sediment material delivered to a specific stream location, such as a culvert or a bridge. In contrast, the SDR, also known as the sediment transmission coefficient, quantifies the percent of gross soil erosion delivered to a particular point in the drainage system. Large SDR often implies that a watershed can deliver relatively more materials (out of its total sediment supply) to culvert located in the lower part of the watershed (Muste and Xu, 2017). According to the USDA (1972), the SDR is related to several stream and watershed characteristics, including the drainage density, slope-length factors (topography), land use/land cover, and soil textures.

The drainage density (a.k.a. channel density) is a useful indicator for characterizing the propagation of the sediment materials through the watershed. This indicator is defined as the total length of all the streams in a drainage basin divided by the total area of the drainage basin (Carlston, 1963). This indicator reflects the average distance that the dislocated sediment travels within a watershed. A higher drainage density implies that the overland sediment transport distance is relatively short, therefore the dislocated particles are more likely to enter the drainage networks before getting trapped. Closely related with respect to SDR estimation is the channel slope. Da and Bartholic (1997) indicate that drainage areas with steep channels usually have a higher SDR than those with flat and wide valleys as the gravity-driven sediment movements on steep slopes are more active.

Finally, the changes in land use and land cover have considerably impacted the SDR magnitude being a direct reflection of the anthropogenic activities in the watersheds. The changes in the land use/ land cover produce alterations of the natural permeability, surface roughness, and the cover condition of a watershed, therefore, directly influencing both overland and in-stream sediment transport (Muste and Xu, 2017). As an example, the conversion of land into riparian vegetation along a stream can restrain sediment movement towards the stream (Pearce et al., 1998). Relatedly, the changes of the soil texture in the watersheds also affect the SDR, as sediment materials of different size are associated with different types of sediment transport process.

An indirect alternative for the estimation of SDR is proposed by Schumm (1981) and Church (1992) who suggest using stream channel characteristics as SDR indicators. They deem that channel planform reflected by stream sinuosity and width is a result of the balance between stream powers (low to high) and the dominating sediment load fraction carried by the stream (suspended, mixed, and bedload). These planform characteristics are used by Schumm (1981) and Rosgen (1996) to develop a comprehensive channel classification system that offers indicators on the stream stability and in-stream sediment transport within the watershed.

In-stream sediment transport is defined as the movement of solid, non-dissolved particles through the stream network. Conventionally, this process is characterized by two major transport modes, namely suspended load and bedload (Karim, 1981). Only a portion of the soil particle dislocated in the headwaters enter in the drainage network as suspended load and is subsequently delivered to downstream locations (Da and Bartholic, 1997). Bonniwell et al. (1999) and Matisoff et al. (2001) demonstrated that fine-grained sediments (e.g. clay and silt) produced from gully and streambank erosion have a good potential to travel a longer distance along a stream as suspended loads. In principle, coarse-grained sediments, such as sand and gravel particles, normally require a higher stream power to be transported as bedload, and are typically prone to deposition and trapping. Irrespective of the type of transport, the culvert design discharge and sediment transport rate display a positive correlation (Howley, 2004; UDOT, 2017). The underlying mechanism of this correlation is explained by the equilibrium between sediment load and stream power (Lord et al., 2009).

The concept of SDR and its descriptors have been introduced and used in the watershed investigation to overcome the limited capability of the physical models to assess sediment yields. Despite the usefulness of all the above-described parameters, direct measurements or estimations of SDR and sediment yield are not readily available across Iowa. Thereby this study utilizes its relevant process-related drivers (i.e., stressors, contributing factors) to indirectly evaluate the SDR for supporting a data-driven approach. Table 2.2 lists the factors that could provide inferences on sediment transport processes that are responsible for culvert sedimentation.

Table 2.2 Parameters related to sediment transport processes

Processes/drivers	Associated parameter	Data source(s)
Sediment particle size & suspended load	Soil type	SSURGO & StreamCAT
Sediment delivery ratio	Channel density	NHD Plus
Sediment delivery ratio, channel flow	Watershed slopes	NHD Plus
Sediment delivery ratio	Land cover	StreamCAT
Sediment load/stream power ratio	Design discharge	USGS (Eash method)
Sediment load/stream power ratio	Stream planform	Aerial images

2.1.3 Sediment deposition at culverts

The descriptions in Sections 2.3.2 and 2.3.3 review essential processes that are responsible for watershed erosions and sediment transport to the culvert locations. In addition to these processes, there are local hydrodynamic processes defining the type of structure-stream interaction whereby the culvert structural characteristics play an important role in settling and trapping the sediment yield in the culvert vicinity. Given the complexity of the local processes and the nature of the transported materials, there are only a handful of studies that have focused on the stream and structural attributes that can affect sediment deposition at culverts (Cafferata et al., 2017; Ho et al., 2013; Howley, 2004).

Most streams carry a sediment load and tend to deposit this load when their velocities decrease. A stable channel is expected to balance erosion and sedimentation over time (i.e., self-cleaning regime). Similarly, culverts which are located on, and aligned with the natural channel, and are non-vegetated are not expected to develop a sedimentation problem. However, 95% of the three-box culverts in Iowa are severely silted, as illustrated in Figure 1.2. This section focuses on the local processes involved in the structure-stream interaction whereby a culvert's structural characteristics play an important role in settling and trapping the sediment yield generated upstream from the culvert location. Among the critical factors leading to sedimentation the following factors were found critical (Ho et al., 2013; Ho et al., 2013): the Stream-to-Culvert Width (SCW) ratio, the type of culvert flow control (i.e., inlet or outlet), and stream ecology.

Stream-to-Culvert Width (SCW) ratio is one of the most important culvert attributes that can directly affect sediment deposition in the culvert vicinity. The origin of this variable stems from the need for multi-box culverts to pass the design discharges associated with flood events. Most of the time the culverts, typically located on small streams, are conveying only a fraction of the design flow. Sized for much larger than normal flow discharges, the geometry of these culverts requires to introduce a transition area in the stream channel that comprises an expansion upstream from the culvert and a contraction downstream from the culvert (Ho, 2010; Charbeneau et al., 2006), as illustrated in Figure 2.4a. The modification of channel geometry caused by the culvert transition produces flow non-uniformity in this area characterized by the divergence/converges of the streamlines for all flow regimes and the creation of areas of flow recirculation, as illustrated in Figures 2.4b (conceptual sketch) and 2c1 (laboratory experiments conducted by Ho et al., 2013).

In addition to flow non-uniformity induced by SCW, more subtle effects are introduced in the vicinity of the culvert due to flow unsteadiness occurring during the passing of a storm flow through the culvert. Notable, the flow during the propagation of a hydrograph has different behavior on the rising and falling limbs of the hydrograph, a.k.a. hysteresis (see Figure 2.4c2). Hysteresis is currently unaccounted for in flow monitoring for the streams with culverts as it is rarely measured with the needed detail. Typically, hysteresis is usually neglected in the design and analyses of river structures. The unsteady flow associated with the storms occurring in the culvert drainage area entrains the sediment in suspension and activate the bedload transport. The rising phase of the hydrograph is characterized by a dynamic sediment transport phase followed by the falling stage whereby the sediment rates return to normal transport regimes. The sedimentation at culvert occurs on the falling stage of the hydrographs, when the stream power diminishes.

Another effect related to flow unsteadiness is that the flow-sediment interaction is more complex than in steady flow interactions. Specifically, the maximum total sediment load passing through a section during a storm event is uncoupled and precedes the peak flow (irrespective of the discharge magnitude), as shown in Figure 2.4c3. The flow non-uniformity combined with the flow dynamics during the transitions decides the location and time period of the occurrences of sediment deposition in the culvert vicinity. Currently, there are considerable gaps regarding the non-uniform, unsteady sediment-laden flows developing in three-dimensional culvert geometry. Lack of understanding of the complexities of this combination of local processes precludes making predictions with analytical tools.

Similarly to the SCW, the type of hydraulic control sections is another culvert structural attribute that can significantly affect the culvert flow, hence the patterns of the sediment depositions. Depending on which end controls the discharge capacity, culverts can be classified either as inlet or outlet control (HDS-5, 1985). A culvert under inlet control has its flow capacity controlled at the entrance by the headwater depth. The culvert serves as a weir when the inlet is unsubmerged and as an orifice when the inlet is submerged (Ho, 2010), and will always flow partially full in a state of shallow, high velocity known as "Supercritical flow". Conversely, a culvert under outlet control has its control section occurred at the outlet section of the culvert, or further downstream, which can flow either partially full or full. Flows passing outlet control culverts are relatively deep and slower, known as "Subcritical flow" (HDS-5, 1985). The type of control determines the flow velocity at the culvert outlet (Howley, 2004). Comparing this velocity with the erosion potential of the channel could provide engineers with a rough estimation of potential for sediment deposition and scouring.

Stream ecology is another critical factor that is considerably involved in sediment accumulation at culverts (Ho et al., 2013). Depending on their location, the vegetated areas could either increase the size of the culvert sediment deposits or reduce the sediment material supplied from the watershed (Muste and Xu, 2017). The reduction of the sediment material is beneficial for the present context as the riparian vegetation or forested stream vicinities are impeding the sediment to reach the stream (Pearce et al., 1998). However, it is the former aspect that is relevant for the present study, so we will focus only on the role of vegetation in the formation of sediment deposits. The vegetation growth over the sediment deposits occurs between storms when the sediment islands reach a height that exceeds the water surface at low flow conditions. Terrestrial vegetation, such as cattail and weed, grows quickly in the fertile soil provided by the sediment deposits. The grown vegetation acts as additional roughness for the high flows loaded with sediment that further exacerbates the rates of sediment deposition. Field observations conducted by Muste and Xu (2017) revealed that after a 4-5 year deposition cycle the vegetated sediment deposits stabilize and consolidate so that subsequent storms cannot wash away the deposited material. Where the vegetation growth is limited by the absence of light, such as is the area within the culvert boxes, the deposited material is much less than that of outside the culvert (both at the inlet and outlet).

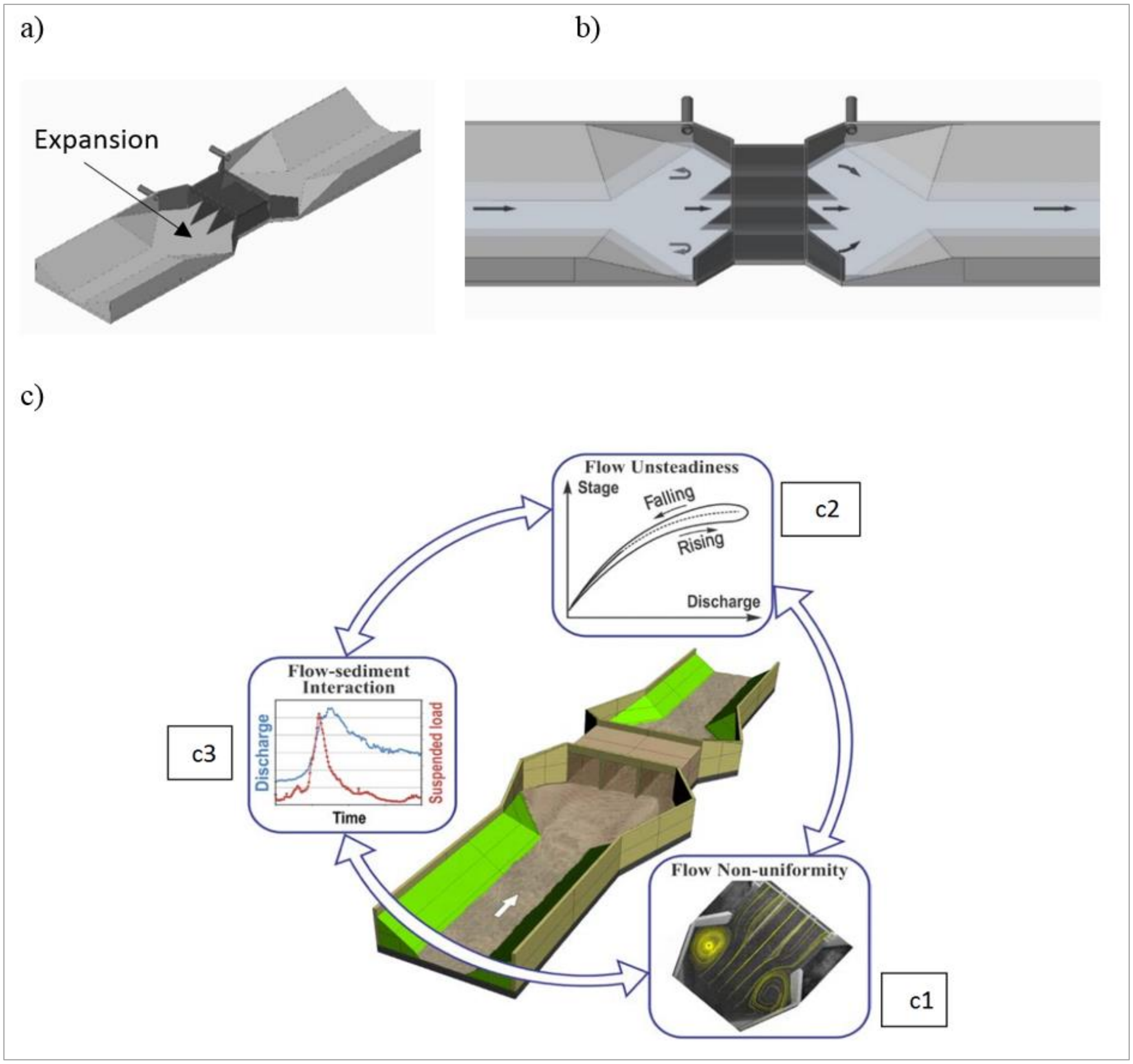


Figure 2.4 Flow configuration at culverts: a) stream-to-culvert transition; b) flow patterns in the vicinity of the culvert (patterns vary with the streamflow magnitude); c) flow complexities due to flow unsteadiness (Ho et al., 2013)

Field inspections are carried out at 257 culverts located on Iowa stream. These inspections reveal that most of the sediment deposits developed near culverts are made of fine sediments delivered as the suspended load (illustrated in Figure 2.5). The preferential location of the depositional areas coincides with the location of the area of flow recirculation in the culvert vicinity (see Figure 2.5a). Analysis of the soils obtained by coring support these statements (Ho et al., 2013). These deposits develop quickly, sometimes within the yearly hydro-meteorological cycle. Carey (2006) also confirms that the suspended sediments, made of fine and colloidal particles, are major causes of siltation at fence lines, waterways, and road culverts.

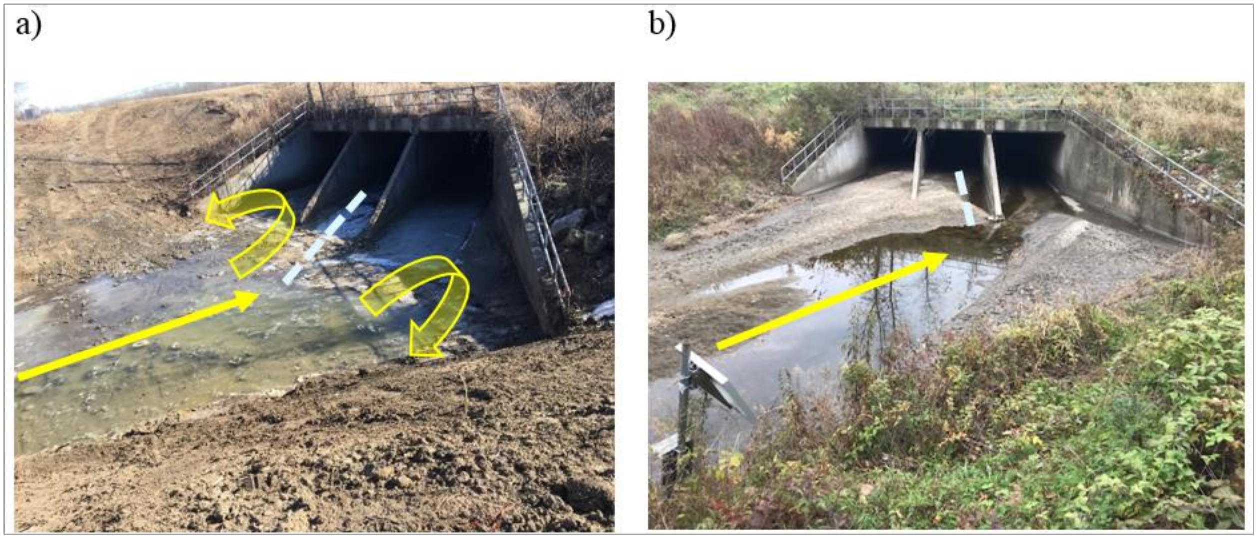


Figure 2.5 Accumulation of sedimentation during eight months of culvert operation from a total cleanup: a) culvert after cleanup in February, 2018; b) sediment deposits in October, 2018. The fine suspension deposits become quickly vegetated retaining sediment at accelerated rates. (Yellow arrow: streamflow dominant direction; dotted white line: culvert direction)

Only a few sites in Iowa located in small or steep drainage basins (e.g., headwaters and plateau areas) have been observed to have depositions consist of bedload (Muste and Xu, 2017). Images of such a culvert are shown in Figure 2.6. This sites can also develop vegetation if the material in the deposits is prone fertile and the site is well illuminated by natural light. As previous and current experimental evidence concurs in their findings, the present study primarily focuses on the contributing factors that control the transport of suspended sediment, meanwhile only consider the effect of bedload transport on culverts in certain areas of Iowa.

Figure 2.6 Fine-grained sediment deposits at culvert (structure code: 75050): a) the sediment deposited at the culvert was accumulated over 5 years; b) the absence of abundant vegetation development due to the nature of the substrate made of well-sorted quartz

As can be observed from the above considerations, there are considerable challenges in thoroughly understanding and modeling the non-uniform, unsteady sediment laden flows, developing in the three-dimensional geometry of the culverts vicinity. The most prominent factor that can be used as a global indicator for evaluating the potential occurrences and magnitude of these complex flow behaviors at a culvert is the Stream-to-Culvert Width (SCW) proportion, i.e., the ratio between B and W. A sketch of this important variable is provided in Figure 2.7. It is this ratio that defines the degree of disturbance of the flow in the vicinity of the culvert from the natural flow conditions (undisturbed) passing through the stream geometry before the culvert construction. These parameters along with the others involved in the sediment deposit formation are listed in Table 2.3.

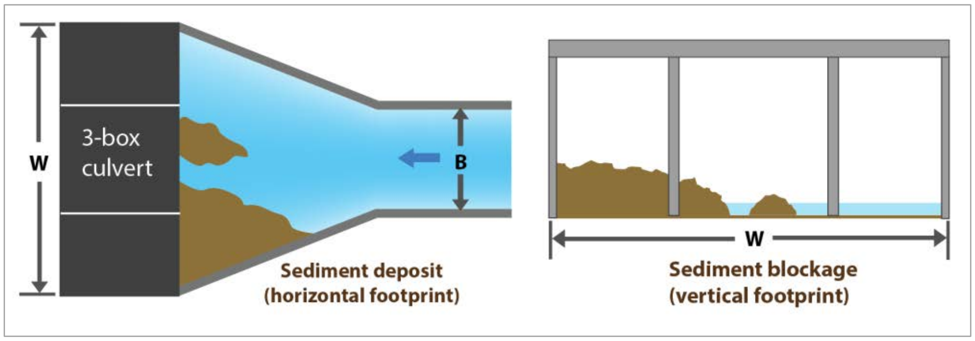


Figure 2.7 Definition of the stream-to-culvert width (SCW) ratio (i.e., B/w)

Table 2.3 Independent variables involved in the local accumulation of sediment at culverts

Parameter	Associated process	Data source(s)
Stream-to-culvert width ratio	Culvert morphodynamics	SIIMS, aerial imagery
Vegetation presence	Sediment deposit growth	Field survey, aerial imagery
Upstream riparian corridor	Streambank erosion	StreamCAT
Angle of incidence	Culvert morphodynamics	Aerial imagery
Culvert flow controls	Culvert morphodynamics	SIIMS

2.2 State of knowledge in sedimentation studies

2.2.1 Past studies on sedimentation at culverts

In general, current knowledge and understanding of culvert sedimentation processes are fragmented and the literature on this topic is scarce (Muste and Xu, 2017). Major knowledge gaps exist in (1) analyzing the interactive chain of the culvert sedimentation processes (sediment production, transport, and deposition) occurring at different spatial scales (drainage area, stream, and culvert structure) as an end-to-end process, and (2) describing and simulating the dynamics of sediment deposition at culvert locations (Ettema and Mutel, 2004). However, the raising concerns of culvert sedimentation and its detrimental impacts have increased the research interests for this subject in recent years. Most of these studies adopt a piece-meal approach, whereby they only focus on certain aspects of culvert sedimentation working in isolation. These type of studies are not able to provide comprehensive guidelines and rigorous design techniques (culvert types and sizing) for preventing sediment deposition at culverts and quantitative assessing potential for sediment deposit formation. There is a need for studies that consider culvert sedimentation as an end-to-end process, which includes soil detachments, sediment transport, and sediment deposition, from a holistic and integrated perspective. Presented next are selected studies to set the stage for the research conducted in the present work.

One of the recent studies was conducted by Howley (2004) who explores the relationships between a culvert structural characteristics and the magnitude sediment depositions at culvert structures with direct reference to the stormwater regulations. The study uses simple graphical depictions and statistical comparisons to analyze site-specific field data collected at 39 culvert locations scattered across the City of Knoxville. Despite its usefulness for culvert maintenance, the study only provides very rough estimates (at correlation level) of the impact of structural factors (e.g., culvert hydraulic controls and slopes) on the formation of sediment deposits at culverts, without considering the aspect of sediment production and transport in the culvert drainage basins.

The study conducted by Goodridge (2009) investigates the behavior of bedload transport (alluvial material in sand and gravel sizes) occurring in pipe culverts. The study develops semi-empirical bedload transport equations for predicting sediment yields at various culvert locations. These equations are derived through the modifications of the existing water conveyance formula by adding factors responsible for the entrainment and transport of sediment as bedload. Since the study is specifically developed to examine the changes in culvert hydraulics in the presence of bedload passing through pipe culverts, its relevance for the present study is limited.

A more general investigative effort is carried out by Bledsoe et al., (2016) who develop scientifically-sound methods for the description of hydraulics and sediment transport at stream crossings (e.g., bridges and culverts) using a watershed approach. Specifically, the analysis explores through theoretical and empirical approaches relationships between design flows and sediment loads for diverse rivers across the U.S and Puerto Rico (Bledsoe et al., 2016). The study's outcomes offers a useful general framework based on physical reasoning for the design hydrology at stream crossings and a series of tools for: (1) generating design hydrology metrics under existing and future land use scenarios, (2) relating channel response potential to selected culvert and bridge designs, and (3) guiding selection of analog reaches (also referred to as reference reaches) and performing rapid geomorphic assessments of channel instability in the field (Bledsoe et al., 2016). Notable is their technological perspective to the design hydrology analysis whereby the use of the tools is described step-by-step in a logical order (e.g., decision trees and a spreadsheet-based application). The scope of Bledsoe et al., (2016) study is relevant to this research as their study aims to optimize hydrology metrics for bridge and culvert design with consideration of stream stability. However, the approach taken in this study does not include site-specific attributes related to culvert operations (hydrodynamic behavior of the culvert during unsteady flows) and the impact of vegetation growth in the culvert vicinity) hence providing an overly simplified tool for predicting the deposition/erosion potential of culverts.

The closest in scope to our study is Rowley's (2014) work which aims to understand how coarse sediments behave near culverts. Rowley investigated embedded-type culverts (bottomless), a culvert-type that is promoted for its ability to enable migration of aquatic organisms. He collected data at multiple sites and compiled them to inform a hydraulic numerical model for predicting the deposition of sediments at the entrance of the culverts, sediment replenishment inside the culverts, and lateral fining within the culvert barrel. This is the first time, according to Rowley, that deposition of sediments upstream of a culvert and lateral fining within a culvert barrel have been successfully modeled. The main distinctions between this study and Rowley's are that the type of the sediment (i.e., suspended vs bedload, respectively) and the simplicity of the culvert geometry (multi-box vs pipe, respectively). These differences were sufficient to produce a completely different sedimentation pattern upstream from culverts: central deposition in Rowley's study vs predominantly lateral deposition in the present study.

Given the complexity of sedimentation at culverts investigations, there are no rigorous design techniques available to size culverts for sediment passage and predict the loading and accumulations of sediment. There are, however, strategies based on engineering judgment that can guide the design of culverts with consideration of sedimentation. Most of the available guidelines are developed for pipe culverts that are easier to deal with due to the simpler flows that they convey. For example, Cafferata et al. (2004) recommends that engineers should:

- Choose a culvert width as close as possible to the width of the natural channel.

- Keep the headwater depth at the culvert inlet at half-full (no more than two-thirds) of the culvert height for the design flow.
- Install the culverts with a slope close to the natural channel.
- Avoid oblique stream-to-culvert angles by setting the culvert along the channel direction.

Excepting the first recommendation, the goals of the above design guidelines are not different from the considerations required for the hydraulic sizing of the multi-box culverts. However, the quantitative aspect of these recommendations does not guarantee better sediment conveyance or the reduction of the probability of structure operational failure. Few additional practical considerations for culvert design are available in UDOT (2017) where the changes induced by deposited sediments within the culverts are estimated as reflected in the modification of the hydraulic gradient and friction factor. The UDOT (2017) guidelines warn users that these assessments are not thoroughly scientific, therefore engineering judgment is essential in their implementation.

2.2.2 Overview of erosion and sediment transport models

Given the complexity of each of the processes involved and the relatively large number of the parameters involved the process description, the erosion and sediment transport physics-based simulation models are preferred over other approaches for the investigation of sediment-related problems (Nearing et al., 2005; Merritt et al., 2003). A particular advantage of the simulations is that they can be easily adapted to different physical domains than the physical models or data-driven approaches which are typically driven by the conditions at the investigation site. The recent advancements in model skills and computer technologies have triggered a growing exploration of catchment erosion and overland and in-stream sediment transport via numerical simulations with computer models (Merritt et al., 2003). These models are designed to address a wide variety of problems such as those related to excessive erosion, water pollution, and river stability, sustainability of watershed water resources (Merritt et al., 2003; Sorokine et al., 2006; Jha and Paudel, 2010; and Papanicolaou et al. 2008).

One of the most recent comprehensive reviews of the erosion and sediment transport simulation models for catchment-scale applications is offered by Merritt et al. (2003). Given that the catchment-scale prediction of sediment generation and transport requires consideration of both land surface processes and in-stream processes, this review also includes a limited number of models for in-stream processes. The range of reviewed models covers all the types of process representations, namely empirical (lumped parameter), conceptual (comprehensive, partly empirical/mixed) and physically-based (spatially distributed). These representations employ algorithms associated with various degrees of process simplifications. The list of reviewed models and their main specifications are provided in Table 2.4.

Selection of the appropriate models for a specific application requires a thorough understanding of the problem under investigation along with the evaluation of the data requirements, model complexity, the accuracy and validity of the model, model assumptions, the spatial and temporal for various components of the model (Merritt et al., 2003). In using physics-based models for the sedimentation at culvert processes involved in this study, the most difficult obstacle to overcome is a thorough understanding of the spatial distribution of the individual processes. Especially difficult to assess are the natural complexity of the sediment pathways (Jakeman et al.,1999; Prosser et al. 2001), the links between runoff and sediment sources within and outside the river channel, and determining the 'best' parameter combinations for simulating sediment generation and transport (Beck, 1987; Merritt et al., 2003). These knowledge gaps necessary requires to estimate the uncertainties associated with the model dimensionality and structure, especially as these complex models normally suffer from problems with error accumulation and model identifiability, due to over-parameterization (Beven, 1989, 1991, 1996).

Table 2.4 Comprehensive review of the erosion and sediment transport simulation models by Merritt et al. (2003)

Model	Type	Scale	Algorithm
ANSWERS	Physical-based	Small catchment	Yalins' (1963) transport equation
CREAMS	Physical-based	Field 40–400 ha	USLE Conceptualization and Yalins' (1963) transport equation
EMSS	Conceptual-based	Catchment	SIMHYD model (Chiew et al., 2002)
HSPF	Conceptual-based	Catchment	Integrated conceptual models (Cheung and Fisher, 1995)
IHACRES-WQ	Empirical/ Conceptual	Catchment	The STARS model (Merritt et al. 2013)
IQQM	Conceptual-based	Catchment	Conceptual Sacramento model (Merritt et al. 2013)
LASCAM	Conceptual-based	Catchment	USLE Conceptualization (Viney and Sivapalan, 1999)
SWRRB	Conceptual-based	Catchment	Bagnold's stream power concept and the EPIC model (Williams et al., 1984)
Erosion GUEST	Physical-based	Plot	Rose concept (Misra and Rose, 1996)
LISEM	Physical-based	Small catchment	SWATRE soil water model (Belmans et al., 1983)
PERFECT	Physical-based	Field	MUSLE Conceptualization (Merritt et al. 2003)
SEDNET	Empirical/Conceptual-based	Field Catchment	USLE Conceptualization and reservoir/lake trap-efficiency sub-model Brune (1953)
TOPOG	Physical-based	Hillslope	Rose concept (Hairsine and Rose, 1992).
USLE	Empirical	Hillslope	USLE Conceptualization (Wischmeier and Smith, 1978)
WEPP	Physical-based	Hillslope/catchment	GAML equation and Foster's equation (Laflen et al., 1991; Lane et al., 1995)
MIKE-11	Physical-based	Physical	1D advection–dispersion equation Hanley et al. (1998)

Another common feature of the erosion and sediment transport models, especially the physics-based ones, is that due to the process complexity, large-scale for their implementation (from lot level to stream network and river crossway), and resolution of the models they need a data-rich environment. In other words, models often exceed the data availability in the area being modeled (Merritt et al., 2003). While significant progress has been made by the increased availability of high-resolution topographic data through DEM obtained with various technologies, the resolution of the soil properties, vegetation cover, and other parameters of the process components are not matched over the entire range of scales.

Another comprehensive review on the sediment transport models focused solely on the in-stream processes is published by Papanicolaou et al. (2008). This review scrutinizes models based on their analytical model formulation, their spatial and temporal characteristics, the coupling/linkage of the hydrodynamic and sediment components, and the model's predictive capabilities. The review includes models that encompass a wide spectrum of practical applications (e.g., suspended load versus bed-load) and spatio-temporal formulations (e.g., one-dimensional model (1D); two-dimensional model (2D); or three-dimensional model (3D); and steady versus unsteady). Caution is advised on the selection of a certain model for solving a specific problem in close correlation with the nature and complexity of the problem itself, the chosen model capabilities to simulate the problem adequately, data availability for model calibration, data availability for model verification, and overall available time and budget for solving the problem (Papanicolaou et al., 2008). Table 2.5 lists useful information about the model capabilities to handle unsteady flows, bed load and suspended load, sediment exchange processes, type of sediment-cohesive versus cohesionless, and multi-fractional sediment transport.

Table 2.5 Comprehensive review of sediment transport simulation models by Papanicolaou et al. (2008)

Model	Type	Flow	In-stream sediment transport				Sediment exchange processes
			Bed sediment	Suspended sediment	Sediment mixtures	Cohesive sediment	
HEC-6	1D	Steady	Y	Y	Y	N	Entrainment deposition
MOBED	1D	Unsteady	Y	Y	Y	N	Entrainment deposition
IALLUVIAL	1D	Quasi-steady	Y	Y	Y	N	Entrainment deposition
FLUVIAL	1D	Unsteady	Y	Y	Y	N	Entrainment deposition
GSTARS	1D	Unsteady	Y	Y	Y	N	Entrainment deposition
SERATRA	2D	Unsteady	Y	Y	N	Y	Advection diffusion
SUTRENCH-2D	2D	Quasi-steady	Y	Y	N	N	Advection diffusion
TABS-2	2D	Unsteady	Y	Y	N	Y	Entrainment deposition
MOBED2	2D	Unsteady	Y	Y	Y	N	Entrainment deposition
USTARS (alluvial river only)	2D	Unsteady	Y	Y	Y	N	Entrainment deposition
ECOMSED	3D	Unsteady	Y	Y	N	Y	Entrainment deposition
RMA-10	3D	Unsteady	Y	Y	N	Y	Entrainment deposition
GBTOXe	3D	Unsteady	N	Y	N	Y	Entrainment deposition
EFDC3D	3D	Unsteady	Y	Y	Y	Y	Entrainment deposition
ROMS	3D	Unsteady	Y	Y	Y	N	Entrainment deposition
CH3D-SED	3D	Unsteady	Y	Y	Y	Y	Entrainment & deposition

Similar to the review on the catchment-scale modeling provided by Merritt et al. (2003), the Papanicolaou's paper point out that a mismatch exists in the theoretical foundations and performance of the hydrodynamic and sediment components of models. The disparity that still exists between the hydrodynamic and sediment transport components is attributed to the difficulties to model turbulent flow interaction with sediment transported as suspended and bed loads in addition to technical problems involved in the development and discretization of the models. The research on sediment-flow interaction is considerably lagging our knowledge on clear-water turbulent flows and continues to remain an open scientific issue.

The review carried out in Section 2.1 illustrates that sedimentation at culverts is driven by the interaction among geomorphological (different types of erosions in watersheds), hydrological (instream sediment transport), structural hydrodynamics (sediment deposition at culvert), and ecological (growth of the riparian vegetation that stabilizes sediment deposition) processes (Cafferata, 2004; Haan et al, 1994; Ho, 2010; Merritt et al., 2003). Some of these processes are further controlled by other external drivers (e.g., climate change and socio-economic aspects of land and river use) and hidden variables (e.g., local structural retrofitting and channel modifications). This natural landscape complexity is difficult to be captured with the available models critically reviewed in Section 2.2.

The role of the review simulation models presented in Section 2.2 is two-fold: a) to evaluate the capabilities of the models to investigate the sedimentation at culverts using a holistic, end-to-end approach, and b) to inform with insights the data-driven approach adopted for this study. Related to the first aspect, it can be concluded that the model performance and accuracy remain a major difficulty in using simulation models, particularly for physics-based spatially distributed models, which are the only types of models that can reliably tackle the problem of sedimentation at culverts from the sediment source to the location of permanent deposition at the culvert location. A summary of the most common limitations ensuing from the literature review is provided below:

a) The quality of modeling prediction is ultimately dependent on the data that are used to support the models. Physical-based erosion models (e.g., ANSWERS, WEPP, and CREAMS) require a considerable amount of data in model-specific formats and a prescription of initial and boundary conditions which are not available over large areas (Abaci and Papanicolaou, 2009; Merritt et al., 2003).

b) Physical-based models are time- and resource-intensive demanding high computational requirements when applied to large catchment scales. Thus, it is not practical and economical to apply these models to simulate the sedimentation processes at hundreds of culverts across the state (Papanicolaou and Abaci, 2008);

c) Conceptual models, such as USLE and RUSLE, have a limited ability to estimate soil losses in drainage basins with heterogeneous land use and topography (Jakeman et al.,1999; Merritt et al., 2003). More sophisticated conceptual models, such as IQQM and HSPF, rely heavily on calibration against field data for parameterization (Walton and Hunter, 1996).

d) Very few erosion models can address multiple types of erosion and sediment transport simultaneously (see Table 2.6) as most of these models are developed to simulate very specific erosion scenarios, such as sheet and rill erosion working independently in agricultural watersheds (Ganasri and Ramesh, 2016; Nearing et al., 2005; Wischmeier and Smith, 1978);

e) There are very few models that consider gully erosion (Merritt, et al 2003). However, gully erosion is considered to be an important contributor to culvert sedimentation for supplying fine-grained sediment materials (Carey, 2006). Lacking the ability to consider gully erosion is also a major limitation of WEPP (which is currently used by the Iowa Daily Erosion Project) (Merritt, et al 2003).

f) Most erosion and sediment transport models do not simulate sediment deposition at the man-made structures (Rowley, 2014), while the ins-stream models that account for hydrodynamic aspects of sediment deposition (e.g., SMS/SRH-2d) do not account for sediment production at watershed scale.

g) Most of the in-stream erosion and sediment transport models are developed to simulate natural channel flows that are uniform, under steady conditions, without disturbances produced by man-made structures. There are still few examples of models that are capable of simulating process representation on an event basis (Merritt et al., 2003), while flow unsteadiness is a key process in the formation of sediment deposits at culverts.

It is quite obvious from the review provided in this section that despite the usefulness of the simulation models, culvert sedimentation cannot be well addressed only through soil erosion modeling alone. Insights from models applied to appropriate erosion and sediment transport scales can be beneficial for developing a sound data-driven approach by providing guidance in selecting relevant variables and offering importance ranks, as well as for providing benchmark data for various process components for training and validation of the algorithms used in the data-driven approach.

Table 2.6 Merritt et al. (2003) review of the erosion and sedimentation models

Model	Rainfall-runoff	Land surface sediment			Gully	In-stream sediment			Sediment associated water quality	
		Generation	Transport	Deposition		Generate	Transport	Deliver	Land	In-stream
AGNPS	Y	Y	N	N	Y	Y	Y	Y	Y	Y
ANSWERS	Y	Y	Y	Y	N	N	N	N	N	N
CREAMS	Y	Y	Y	Y	Y	N	N	N	Y	N
EMSS	Y	N	N	N	N	Y	Y	Y	N	N
GUEST	Y	Y	Y	Y	N	N	N	N	N	N
HSPF	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
IHACRES-WQ	Y	N	N	N	N	Y	Y	Y	Y	Y
IQQM	Y	N	N	N	N	N	N	N	N	N
LASCAM	Y	Y	N	N	N	Y	Y	Y	Y	Y
LISEM	Y	Y	N	N	N	Y	Y	Y	N	N
MIKE-11	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
PERFECT	Y	Y	N	N	N	N	N	N	Y	N
SEDNET	Y	Y	N	N	Y	Y	Y	Y	Y	Y
SWRRB	Y	N	N	N	N	Y	Y	Y	Y	Y
TOPOG	Y	Y	Y	Y	N	N	N	N	N	N
USLE	N	Y	N	N	N	N	N	N	N	N
WEPP	Y	Y	Y	Y	N	Y	Y	Y	N	N

2.3 Data-Driven approaches for erosion and sediment transport

2.3.1 Overview

Data-driven approaches can effectively help environmental scientists in analyzing and solving large-scale complex problems that involve handling of a large number of dependent and independent variables, pattern recognition in the variable distributions, handling relationships among variables (even if they are of non-linear nature), and classification based on supervised or non-supervised machine learning (Tayfur 2002). Different from theory-driven modeling and experimental methods, data-driven approaches are able to map the empirical relationships between the dependent variable (i.e., the response variable) and independent variables (i.e., predictors) without detailed consideration of the internal structure of the physical process (Dibike and Solomatine, 2001).

Ideally, data-driven investigations of sediment-related issues should rely on the direct measurements of sediment-related variables (i.e., bedload and suspended load) in streams (Marteau et al., 2018). However, direct observations of sediment-related parameters are relatively scarce across Iowa (shown in Table 2.7). Even within certain areas that have data coverage, these observations only cover short time intervals (Wohl et al., 2015). Given the scarcity of sedimentation-related data in Iowa, the present study utilizes deductive reasoning to investigate the complex processes of culvert sedimentation through environmental variables that may reflect the main drivers of sedimentation processes (Olden et al., 2012; Rymszewicz et al., 2018; Wagener et al., 2007). The hypothesis of this deductive approach, along with other premises required by the data-driven approach, are presented below.

Table 2.7 Availability of sedimentation observations by State or Region (Hinton and Rollin, 2017)

Location	Number of data sets	Number of observations
Outside the United States	15	505
Alaska	23	489
Arizona	1	34
California	161	4270
Colorado	130	3051
Idaho	51	4240
Indiana	6	61
Iowa	**1**	**13**
Michigan	1	7
Mississippi	1	358
Nebraska	3	30
Nevada	10	254
New Mexico	3	7
North Carolina	3	15
North Dakota	2	12
Oregon	9	279
South Dakota	4	17
Utah	4	108
Washington	4	184
Wisconsin	1	4
Wyoming	51	1143
Total	484	15081

Use of the data-driven approaches for investigating the culvert sedimentation depends on the following premises (Olden et al., 2012; Rymszewicz et al., 2018; Wagener et al., 2007; Wohl et al., 2015). First, as final recipients of the water-borne material transported from local and upstream basins, streams are a good reflection of the hydrologic characteristics of upstream watershed areas (Allan, 2004; Auerbach et al., 2016). From this perspective, in-stream sediment conditions can be indirectly estimated using various extents of the culvert drainage areas (watershed, catchment, riparian corridors). According to Wohl et al., (2015), sediment regime can be estimated based on changes in river planform, substrate characteristics, and floodplain characteristics through time or with respect to reference channels. These characteristics are increasingly available in recent years from analyses of remote-sensing imagery. Additionally, Lord et al. (2009) list environmental variables that are considered as signs of fluvial systems vitality, hence enabling inferences on the sediment budget in watersheds.

Secondly, there are abundant data about watersheds and streams from various public sources that are relevant to erosion and sedimentation processes despite they are not collected for this purpose. This ad-hoc situation creates a data-rich environment that is freely available for both Iowa and the contiguous United States. Consequently, a data-driven approach developed for Iowa can be readily generalizable and extended to address similar environmental issues in other US regions. Thirdly, it is possible to conveniently characterize the sedimentation-at-culvert problem by taking advantage of the aerial photographs (e.g. National Digital Orthophoto Program, National Agriculture Imagery Program, Google Maps, and EagleView) that continuously cover the development of the sedimentation process in time and space through repeated surveys. These aerial imagery surveys cover the entire US territories but the resolutions of the maps might vary across regions. In this research, the reliability of the quantified the culvert sedimentation degree obtained through aerial photography has been successfully tested with multiple in-situ site surveys conducted with conventional and Unmanned aerial vehicle (UAV) -based imaging (Muste and Xu, 2017).

With these premises satisfied, the data-driven approach is able to identify and map the relationships between the degree of culvert sedimentation and its contributing factors (drivers). Based on the review of culvert sedimentation mechanisms in Section 2.1, this study identifies a wide variety of watershed and culvert-related characteristics that may reflect the drivers of culvert sedimentation processes (see Tables 2.1 to 2.3). Many of these characteristics are variables (see the highlighted variables in Table 2.8) that describe the vitality of a fluvial system as proposed by Lord et al (2009).

In terms of the type of variables used for the data-driven approach, the process-drivers of culvert sedimentation are considered as the independent variables (i.e., predictors), while the degree (severity) of sedimentation at each culvert is considered the dependent variable (i.e., response variable). By quantifying the dependencies between the process-drivers and the culvert sedimentation degree, this study intends to shed light on the mechanisms behind the complex processes involved in culvert sedimentation. The relationships between variables provide useful insights for developing effective strategies for mitigating culvert sedimentation by addressing specific drivers through culvert structural retrofitting and upstream management practices.

Table 2.8 Summary of environmental variables that may affect fluvial processes based on information compiled from Allen (2004), Gregory and Walling (1987), and Lord et al. (2009)

Scale	**Driver source area**	**Driver category**	**Specific driver examples**
Basin	Outside watershed	Altered hydrology and altered sediment budget	climatic change
			air pollution
	Inside watershed	Altered hydrology	urbanization
			roads and parking lots
			storm drains
			vegetation changes
			forestry practices
			consumptive use of groundwater
		Altered sediment budget	construction
			forestry practices
			mining
			dirt and gravel roads
			trails
			agriculture (crops and grazing)
			water impoundments
			vegetation changes
Channel	Inside watershed	Altered hydrology	water impoundments (dams)
			water diversions
			consumptive use of surface water
		Altered sediment budget	dredging
			road and trail crossings
		Altered channel	channelization
			dredging
			bridges
			bank stabilization structures
			grazing
			removal or change in bank or riparian vegetation landslides

2.3.2 Relevant data-driven techniques

Depending on their analytical reasoning, data-driven techniques can be further categorized as either computational data models or visual-analytics. In this study, both of these technique categories have been used. Computational methods, such as machine-learning algorithms and statistical models, analyze complex system and patterns in datasets relying on the machine or mathematical reasoning (Roderick, 2015). Typically, their implementations consume extensive computational resources. Techniques such as artificial neural networks (Tayfur, 2002), random forests (Francke et al. 2008; López-Tarazón et al. 2011), and M5 model trees (Onderka et al., 2012) pertain to this category and have been used to estimate sediment transport and budgets in the past. Despite their capability of handling large environmental datasets and making predictions, these computational models are deemed as “black-boxes” as their derivation processes are often hidden and their results are difficult for human understanding and interpretation (Keim, 1997). Given that this study aims to support the decision-making process for culvert sedimentation mitigation along with improved understanding of the problem’s nature (e.g., data, dynamics, and the critical relationships involved), the computational techniques alone are not sufficient for fulfilling the targeted goals.

Visual analytics integrates both computational and theory-based tools with interactive visual interfaces to facilitate high-quality human reasoning and enable human-information discourse (Thomas and Cook, 2005; Qiu, 2007). Since visual analytics combines the power of computational techniques with human reasoning, it is able to foster better human interpretation of abstract patterns and correlations within a dataset and support decision-making through human-computer collaboration (Keim, 1997). For the present context, the outcomes of the visual-based methods are deemed more intuitive for data scientists and other professionals (e.g., culvert managers and structural engineers), rendering these methods effective for providing easy-to-understand suggestions and insights for improving culvert design.

The design and implementation of visual analytical tools for environmental decision-making are challenging due to the complexity, size, and high-dimensionality of geospatial environmental data (Andrienko et al., 2007; Demir et al., 2015). To address these limitations, efforts are made in Geographic Information Science (GIS) to integrate visual and computational methods for conducting exploratory data analysis separately from supporting decisions (Anselin, 2000; Bertolotto et al., 2007; Demir and Beck, 2009; Guo et al., 2005). Visual analytics, is still in early implementation stages for investigation of water and sediment management-related studies except for a few studies (Kollat et al., 2011; Leonard et al., 2016; Matrosov et al., 2015).

Both Kollat et al. (2011) and Matrosov et al. (2015) incorporate visual analytics techniques into water management frameworks for supporting decision-making. Kollat et al. (2011) combines visual analytics techniques with contaminant flow-and-transport modeling, bias-aware ensemble Kalman filtering (EnKF), and many-objective evolutionary optimization, to improve long-term groundwater monitoring decisions across space and time while accounting for the influences of systematic model errors. Matrosov et al. (2015) demonstrate a water resource management simulator that utilizes many-objective visual analytics to reveal the key trade-offs inherent in planning a real-world water supply system. Differently, Leonard et al. (2016) applies visual analytics techniques to big data cleaning and integration that involves very stream networks (graphs) to support national-scale hydrological modeling.

2.4 References

Abaci, O. and Papanicolaou, A. N. T. (2009). "Long-term effects of management practices on water-driven soil erosion in an intense agricultural sub-watershed: monitoring and modelling" , Hydrological Processes. 23 (19), 2818-2837.

Allanm, J. D. (2004). "Landscapes and riverscapes: the influence of land use on stream ecosystems", Annual Review of Ecology, Evolution, and Systematics. 35(1): 257-284.

Andrienko, G., Andrienko, N., Jankowski, P., Keim, D., Kraak, M-J., MacEachren, A. and Wrobel, S. (2007). "Geovisual analytics for spatial decision support: Setting the research agenda", International Journal of Geographical Information Science. 21 (8): 839-857, DOI: 10.1080/13658810701349011.

Angulo-Martínez, M., BeguerÃ-a, S., Navas, A. and MachÃ-n, J. (2012). "Splash erosion under natural rainfall on three soil types in NE Spain", Geomorphology. 175-176: 38-44.

Anselin, L. (2000). "Computing environments for spatial data analysis", Journal of Geographical Systems. 2 (3), 201–220.

Auerbach, D. A., Buchanan, B. P., Alexiades, A. V., Anderson, E. P., Encalada, A. C., Larson, E. I., and Flecker, A. S. (2016). "Towards catchment classification in data-scarce regions", Ecohydrology. 9 (7): 1235-1247.

Beck, M. B. (1987). "Water quality modelling: a review of uncertainty", Water Resources Research. 23 (8): 1393–1442.

Belmans, C., Wesseling, J.G., and Feddes, R.A. (1983). "Simulation model of the water balance of a cropped soil: SWATRE", Journal of Hydrology. 63: 271–286.

Bennett, J. P. (1974). "Concepts of mathematical modelling of sediment yield", Water Resources Research. 10: 485–492.

Bertolotto, M., Martino S. Di., Ferrucci, F., and Kechadi, T. (2007). “Towards a framework for mining and analysing spatio-temporal datasets”, International Journal of Geographical Information Science. 21 (8), 895-906.

Beven, K. (1989). “Changing ideas in hydrology—the case of physically based models”, Journal of Hydrology. 105: 157–172.

Beven, K. (1991). “Spatially distributed modelling: conceptual approach to runoff prediction”. In: Bowles, D.S., O’Connell, P.E. (Eds.), Recent Advances in the Modelling of Hydrological Systems: 373–387, Kluwer Academic, Boston.

Beven, K. (1996). “A discussion of distributed hydrological modelling”. In: Abbott, M.B., Refgaard, J.C. (Eds.), Distributed Hydrological Modelling: 255–278, Kluwer Academic.

Bledsoe, B., Baker, D., Nelson, P., Rosburg, T., Sholtes, J., and Stroth, T. (2016). “Design hydrology for stream restoration and channel stability at stream crossings”, National Cooperative Highway Research Program (NCHRP) Project 24-40 Transportation Research Board of the National Academies of Sciences, Engineering, and Medicine, Department of Civil and Environmental Engineering Colorado State University, Fort Collins, Colorado.

Bonniwell, E. C., Matisoff, G., and Whiting, P. J. (1999). “Determining the times and distances of particle transit in a mountain stream using fallout radionuclides”, Geomorphology. 27 (1): 75-92.

Brune, G.M. (1953). “Trap efficiency of reservoirs”, Trans American Geophysical Union. 22: 649–655.

Cafferata, P. H. (2004). “Designing watercourse crossings for passage of 100-year flood flows, wood, and sediment”, California Dept. of Forestry and Fire Protection, Sacramento, CA (Updated 2017).

Carey, B. (2006). "Gully erosion, facts natural resources and water", Managing Queensland natural resources for today and tomorrow, Natural Resource Science, Retrieved on November 7, 2017 from www.nrw.qld.gov.au.

Carlston, C. W. (1963). "Physiographic and hydraulic studies of rivers: drainage density and streamflow", geological survey professional paper 422-C, U.S. Geological Survey, Reston, VA.

Charbeneau, R. J., Henderson, A. D., and Sherman, L. C. (2006). Hydraulic Performance curves for Highway Culverts. Journal of Hydraulic Engineering. 132 (5): 474-481. doi:10.1061/(ASCE)0733-9429(2006)132:5(474).

Cheung, A. S. and Fisher, I. H. (1995). "The use of HSPF program in total catchment management". In: Proceedings of the 16th Federal Convention, AWWA, vol. 2, pp. 747–753.

Chiew, F. H. S., Peel, M. C., and Western, A. W. (2002). "Application and testing of the simple rainfall-runoff model SIMHYD". In: Singh, P. (Ed.), Mathematical Models of Small Watershed Hydrology and Applications. Water Resources Publication, Littleton, Colorado Cited in Vertessey et al., 2001.

Church, M. (1992). "Channel morphology and typology", in Calow, P., and Petts, G.E., eds., The Rivers handbook: hydrological and ecological principles: Malden, Blackwell Scientific Publications, 1: 526.

Da, Q. and Bartholic, J. (1997). "Predicting sediment delivery ratio in Saginaw Bay watershed", Proceedings of 2nd National Association of Environmental Professionals Conference, Orlando, FL.

Demir, I., Conover, H., Krajewski, W. F. , Seo, B. , Goska, R., He, Y., McEniry, M. F., Graves, S. J., and Peterson, W. (2015). "Data enabled field experiment planning management and research using Cyberinfrastructure", Journal of Hydrometeorology. 3: 1155-1170.

Demir, I., Beck, M. B. (2009). "GWIS: a prototype information system for Georgia Watersheds", Paper 6.6.4. In: Proceedings Georgia Water Resources Conference: Regional Water Management Opportunities, April 27-29, 2009. UGA, Athens, GA, USA.

Dibike, Y. B. and Solomatine, D. P. (2001). "River flow forecasting using artificial neural networks", Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere. 26 (1), 1-7.

Emmett W. W. (1978). "Overland flow". Kirkby M.J. (Ed.), Hillslope Hydrology, Wiley, Chichester: 145-176.

Ettema, R. and Mutel, C. (2004). "Hans Albert Einstein: innovation and compromise in formulating sediment transport by rivers", Journal of Hydraulic Engineering 130(6): 477-486.

Fernández-Raga, M., Palencia, C., Keesstra, S., JordÃ¡n, A., Fraile, Roberto., Angulo-MartÃ-nez, M., and CerdÃ, A. (2017). "Splash erosion: A review with unanswered questions", Earth-Science Reviews. 171: 463-477.

Foster, G. R. (1982). "Modelling the erosion process". In Hydraulic Modelling of Small Watersheds. ASAE Monograph, 5.

Francke, T., López-Tarazón, J. A., and Schröder, B. (2008). "Estimation of suspended sediment concentration and yield using linear models, random forests and quantile regression forests". Hydrol Process. 22 (25), 4892–4904.

Ganasri, B. P. and Ramesh, H. (2016). “Assessment of soil erosion by RUSLE model using remote sensing and GIS - A case study of Nethravathi Basin”, Geoscience Frontiers. 7 (6), 953-961.

Goodridge, W. H. (2009). “Sediment transport impacts upon culvert hydraulics”, Master Thesis, Utah State University, Logan Utah.

Gregory, K. J. and Walling, D. E. (1987). “Human activity and environmental processes: Chichester”, John Wiley and Sons: 466.

Guo, D., Gahega, M., MacEachren A., and Zhou, B. (2005). “Multivariate analysis and geovisualization with an integrated geographic knowledge discovery approach”, Cartography and Geographic Information Science. 32 (2): 113–132.

Haan, C.T., Barfield, B. B., Hayes, J. C. (1994). “Design hydrology and sedimentology for small catchments”. San Diego Academic Press, San Diego CA.

Hairsine, P. B. and Rose, C. W. (1992). “Modeling water erosion due to overland flow using physical principles: 1. Sheet flow”, Water Resources Research. 28 (1): 237-243. doi:10.1029/91WR02380

Hairsine, P. and Rose, C. (1992). “Modelling water erosion due to overland flow using physical principles: 2. Rill flow”, Water Resources Research. 28 (1): 245–250.

Hanley, N., Faichney, R., Munro, A., Shortle, J. S. (1998). “Economic and environmental modelling for pollution control in an estuary”, Journal of Environmental Management. 52: 211–225.

Hinton, D. D. and Rollin, H. (2011). “Comprehensive and quality-controlled bedload transport database”. AGU Fall Meeting Abstracts. 143. 0710-. 10.1061/(ASCE)HY.1943-7900.0001221.

Ho, H. C. (2010). "Investigation of unsteady and non-uniform flow and sediment transport characteristics at culvert sites". Ph.D Dissertation, University of Iowa, Iowa City, IA.

Ho, H. C., Muste, M., and Ettema, R. (2013). "Sediment self-cleaning multi-box culverts", Journal of Hydraulic Research. 51 (1): 92-101.

Howley, C. S. (2004). "The relationships among culvert characteristics and culvert sedimentation". Master Dissertation, University of Tennessee Knoxville TN.

Hudson, N. W. (1975). "The factors determining the extent of soil erosion". In: Greemland, R. (Ed.), Soil Conservation and Management in the Humid Tropics. John Wiley and Sons.

Jakeman, A. J., Green, T. R., Beavis, S. G., Zhang, L., Dietrich, C. R., and Crapper, P.F. (1999). "Modelling upland and in-stream erosion, sediment and phosphorus transport in a large catchment", Hydrological Processes. 13 (5): 745–752.

Jakeman, A. J., Green, T. R., Beavis, S. G., Zhang, L., Dietrich, C. R., and Crapper, P. F. (1999). "Modelling upland and in-stream erosion, sediment and phosphorus transport in a large catchment", Hydrological Processes. 13 (5): 745–752.

Jha, M. K. and Paudel, R. C. (2010). "Erosion predictions by empirical models in a mountainous watershed in Nepal", Journal of Spatial Hydrology. 10 (1): 89-102.

Jones, C. S. and Schilling, K. E. (2011). "From agricultural intensification to conservation: Sediment transport in the Raccoon River, Iowa, 1916–2009". J. Environmen. Qual. 40 (6): 1911–1923.

Karim, M. F. (1981). "Computer-based predictors for sediment discharge and friction factor of alluvial streams". Ph.D. Dissertation, University of Iowa, Iowa City, IA.

Keim, D. A., (1997). "Visual techniques for exploring data-bases", Invited tutorial, Int. Conference on Knowledge Discovery in Databases (KDD'97), Newport Beach, CA, USA.

Kollat, J. B., Reed, P. M., and Maxwell, R. M. (2011). "Many - objective groundwater monitoring network design using bias - aware ensemble Kalman filtering, evolutionary optimization, and visual analytics", Water Resources Research. 47 (2).

Lane, L., Nichols, M., and Paige, G. (1995). "Modeling erosion on hillslopes: Concepts, theory and data". In: International Congress on Modelling and Simulation Proceedings (Agriculture, Catchment Hydrology and Industry), 1: 1–17.

Laflen, J. M., Lane, L. J., Foster, G. R. (1991). "WEPP: A new generation of erosion prediction technology", Journal of Soil and Water Conservation. 46: 34–38.

Leonard, L., MacEachren, A. M., and Madduri, K. (2016). "Graph-based visual analysis for large-scale hydrological modeling", Information Visualization. 16 (3): 205 – 21.

Loch, R. J. and Silburn, D. M. (1996). "Constraints to sustainability—soil erosion". In: Clarke, L., Wylie, P.B. (Eds.), Sustainable Crop Production in the Sub-tropics: an Australian Perspective. QDPI.

Lord, M. L., Germanoski, D. and Allmendinger, N. E. (2009). "Fluvial geomorphology: Monitoring stream systems in response to a changing environment, in Young, R., and Norby, L., Geological Monitoring: Boulder, Colorado", Geological Society of America: 69–103, doi: 10.1130/2009.monitoring (04).

López-Tarazón, J. A., Batalla, R. J., Vericat, D., and Francke, T. (2011). "The sediment budget of a highly dynamic mesoscale catchment: the River Isábena", Geomorphology. 138 (1): 15–28.

Marteau, B., Batalla R. J., Vericat, D., Gibbins, C. (2018). "Asynchronicity of fine sediment supply and its effects on transport and storage in a regulated river", Journal of Soils and Sediments. 18 (7): 2614–2633. https://doi.org/10.1007/s11368-017-1911-1.

Matisoff, G., Ketterer, M. E., Wilson, C. G., Layman, R., and Whiting, P. J. (2001). "Transport of rare earth element-tagged soil particles in response to thunderstorm runoff", Environmental Science and Technology. 35 (16): 3356-3362. doi:10.1021/es001693m

Matrosov, E. S., Huskova, I., Kasprzyk, J. R., Harou, J. J., Lambert, C., and Reed, P. M. (2015). "Many-objective optimization and visual analytics reveal key trade-offs for London's water supply", Journal of Hydrology. 531 (3): 1040-1053.

Merritt, W. S., Letcher, R. A., and Jakeman, A. J. (2003). "A review of erosion and sediment transport models". Environmental Modelling & Software. 18 (8): 761-799.

Misra, R. K. and Rose, C. W. (1996). "Application and sensitivity analysis of process-based erosion model GUEST", European Journal of Soil Science. 47: 593–604.

Morgan, R. P. C. (2005). "Soil Erosion and Conservation". Blackwell Publishing, Oxford.

Muste, M. and Xu, H. (2017). "Sedimentation mitigation using streamlined culvert geometry", Report ST-001, Iowa Department of Transportation, Statewide Transportation Innovation Council, Federal Highway Administration, McLean, VA.

Muste, M. and Xu, H. (2017). "Mitigation of sedimentation at multi-box culverts", IIHR Report No. TR-655, Submitted to the Iowa Highway Research Board, Ames, IA, USA.

Nearing, M. A., Jetten, V., Baffaut, C., Cerdan, O., Couturier, A., Hernandez, M., Le Bissonnais Y., Nichols, M. H., Nunes J. P., Renschler C. S., SouchÃ¨re, V., and Van Oost, K. (2005). "Modeling response of soil erosion and runoff to changes in precipitation and cover", CATENA. 61 (3): 131-154.

Normann, J. M., Houghtalen, R. J., and Johnson, W. J. (1985). "Hydraulic design of highway culverts" Hydraulic Design Series No. 5, 2nd Ed., Federal Highway Administration, Washington, D.C.

Olden, J. D., Kennard, M. J. and Pusey, B. J. (2012). "A framework for hydrologic classification with a review of methodologies and applications in ecohydrology", Ecohydrology. 5 (4): 503-518.

Olden, J. D., Lawler, J. J. and Poff, N. L. (2008). "Machine learning methods without tears: a primer for ecologists", The Quarterly Review of Biology. 83(2): 171-193. doi:10.1086/587826

Onderka, M., Krein, A., Wrede, S., Martinez-Carreras, N., and Hoffmann, L. (2012). "Dynamics of storm-driven suspended sediments in a headwater catchment described by multivariable modeling", Journal of Soils and Sediment. 12:620-635.

Papanicolaou, A., Elhakeem, M. Prakash, S. and Edinger, J. (2008). "Sediment transport modeling Review – Current and future developments", J. Hydraulic Engineering. 134 (1): 1-14.

Pearce, R. A., Trlica, M. J., Leininger, W. C., Mergen, D. E., and Frasier, G. (1998). "Sediment movement through riparian vegetation under simulated rainfall and overland flow", Journal of Range Management. 51 (3): 301-308.

Prosser, I. P., Rutherford, I. D., Olley, J. M., Young, W. J., Wallbrink, P. J., and Moran, C. J. (2001). "Large-scale patterns of erosion and sediment transport in river networks, with examples from Australia", Marine and Freshwater Research. 52: 81–99.

Qiu, G. P. (2007). "Advances in visual information systems". 9th International Conference (Visual).

Roderick, M. (2015). "Mathematical and computational modeling: with applications in natural and social sciences, engineering, and the arts". Wiley. ISBN 978-1-118-85398-6.

Rose, C. W. (1993). “Erosion and sedimentation”. In: Bonell, M., Hufschmidt, M.M., Gladwell, J.S. (Eds.), Hydrology and Water Management in the Humid Tropics: Hydrological Research Issues and Strategies for Water Management. Cambridge niversity Press: 301–343.

Rose, C. W. (1993). “Erosion and sedimentation”. In: Bonell, M., Hufschmidt, M. M., Gladwell, J. S. (Eds.), Hydrology and Water Management in the Humid Tropics: Hydrological Research Issues and Strategies for Water Management. Cambridge University Press: 301–343.

Rosgen, D. L. (1996). “Applied river morphology” (second edition): Pagosa Springs, Colorado, Wildland Hydrology: 390.

Rowley, K. J. (2014). “Sediment transport conditions near culverts”. MS Thesis Dissertation, Brigham Young University, Provo, UT.

Rymszewicz, A., Bruen, M., O'Sullivan, J. J., Turner, N. J., Lawler, D. M., Harrington, J. R., Conroy, E., and Kelly-Quinn, M. (2018). “Modelling spatial and temporal variations of annual suspended sediment yields from small agricultural catchments”, Science of The Total Environment. 619-620: 672-684.

Schumm, S. A. (1977). “The fluvial system”: New York, John Wiley and Sons: 338.

Schumm, S.A. (1981), “Evolution and response of the fluvial system”, sedimentologic implications: SEPM Special Publication, No. 31: 19–29.

Sorokine, A., Bittner, T., and Renschler, C. (2006). “Ontological investigation of a multiscale ecosystem classification using the “national hierarchical framework of ecological units” as an example”, Geoinformatica. 10 (3): 313-335. doi:10.1007/s10707-006-9830-0.

Tayfur, G. (2002). "Artificial neural networks for sheet sediment transport", Hydrological Sciences Journal. 47:879–892.

Thomas, J. J. and Cook, K. A. (Eds). (2005). "Illuminating the path. the research and development agenda for visual analytics": IEEE Computer Society, Los Alamitos, CA.

UDOT (2017). "Drainage manual, appendix 9.C: sedimentation at culverts", Utah Department of Transportation (UDOT), Salt Lake City, UT (https://www.udot.utah.gov/main/f?p=100:pg:0:::1:T,V:826).

United Nations Environment Programme (UNEP). (1998), "Best management practices for agricultural non-point sources of pollution", Caribbean Environment Programme (CEP) Technical Report NO. 41.

USDA (1972). "Sediment sources, yields, and delivery ratios", US Department of Agriculture, Soil Conservation Service, National Engineering Handbook, Section 3, Sedimentation, Chapter 6, Washington, D.C.

Viney, N. R. and Sivapalan, M. (1999). "A conceptual model of sediment transport: application to the Avon River Basin in Western Australia", Hydrological Processes. 13: 727–743.

Wagener, T., Sivapalan, M., Troch, P. A. and Ross, W. A. (2007). "Catchment classification and hydrologic similarity", Geography Compass. 1 (4): 901–931, 10.1111/j.1749-8198.2007.00039.

Walton, R. and Hunter, H. (1996). "Modelling water quality and nutrient fluxes in the Johnstone River Catchment, North Queensland". In: 23rd Hydrology and Resources Symposium, Sydney.

Williams, J. R., Jones, C. A., and Dyke, P. T. (1984). "A modelling approach to determining the relationship between erosion and soil productivity". Transactions of the ASAE. 27: 129–144.

Wischmeier, W. H. and Smith, D. D. (1978). "Predicting soil erosion losses: a guide to conservation planning. usda agricultural handbook". No. 537: 58.

Wohl, E., Bledsoe, B. P., Jacobson, R. B., Poff, N. L., Rathburn, S. L., Walters, D. M., and Wilcox, A. C. (2015). "The natural sediment regime in rivers: broadening the foundation for ecosystem management", BioScience. 65 (4): 358–371.

Yalin, M.S. (1963). "An expression for bed load transportation". Journal of Hydraulics Division, American Society of Civil Engineers. 98 (HY3): 221–250.

CHAPTER 3 DATA DRIVEN FRAMEWORK DESIGN

3.1 Overview

This chapter discusses the design considerations and methods of the data-driven framework developed to meet the research objectives. The key components of the framework include (see Figure 3.1): (1) Data Compilation & Integration, (2) Multiple-Criteria Decision Analysis, (3) Culvert Sedimentation Forecasting, and (4) Forecasting Optimization.

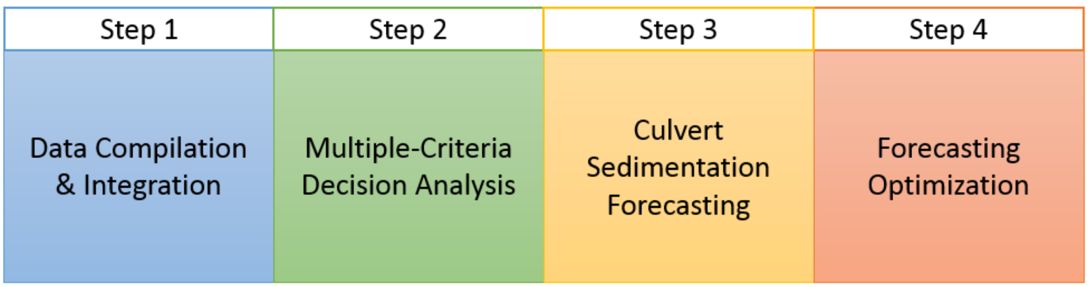


Figure 3.1 Major components of the data-driven framework

Details of the four components are provided through the following sections: Section 3.2 elaborates on the preparation of a comprehensive culvert sedimentation dataset to be used in the data-driven analysis. The dataset is assembled through the integration of third-party environmental data with proprietary data regarding the culvert sedimentation degree, which is the targeted variable for the present study. The latter data is also based on processing third-party data, i.e., time series of aerial images collected by various agents. Section 3.3 describes the core element of the data-driven investigation, the Multiple-Criteria Decision Analysis (MCDA). This component uses quantitative and qualitative data along with human judgment to derive empirical relationships between the degree of culvert sedimentation and key process drivers within culverts' drainage area. The analysis is backed with machine-learning and information visualization techniques, therefore is able to facilitate human exploration and understanding of the culvert sedimentation drivers. Section 3.4 presents the use of the quantified relationships, derived in the previous section as decision trees, to forecast culvert sedimentation potential across Iowa in real-time for the first time. Despite the usefulness of the early forecasting, it is, however, relatively coarse and has limited ability to enable further exploration of the multivariate relationships (patterns) between the culvert sedimentation degree and process-drivers. Consequently, optimization of forecasting is carried out to improve the outcomes of the predictions. These analyses include: (a) conducting sensitivity analysis on different spatial extents for characterizing process-drivers, (b) partitioning the state of Iowa into multiple regions with more homogeneous erosion potential and soil mobility through multivariate clustering analysis and regionalization, and (c) performing multivariate clustering analysis (Self-Organizing Map) to further explore multivariate relationships between the culvert sedimentation degree and process-drivers within each region.

Later in the study, the data-driven framework and its optimizations are embedded into a web-based geospatial environment, named as the "IowaDOT Culverts" platform. The platform provides users with critical information for planning, designing, and maintaining sustainable culverts through user-friendly visual interfaces. Details regarding the actual development of the "IowaDOT Culverts" platform and the usage of the associated workflows are provided in Chapter 4.

Many contents in this chapter are authorized replications of the IHRB TR-655 report (Muste and Xu, 2017), authored by my advisor Dr. Marian Muste and myself, as the funding for this Ph.D. study was provided by the Iowa Highway Research Board, and Iowa Department of Transportation (Iowa DOT), Grant TR-655.

3.2 Data compilation & integration

The data for multiple variables identified in the previous chapter to drive the erosion and sedimentation processes at culverts are widely available for all U.S. territories. However, they are stored in various formats and multiple repositories which makes their extraction and handling difficult. This study's primary goal is to integrate the independent and dependent variables in one digital repository hosted in a web-GIS environment that facilitates the tasks to be conducted in a subsequent study phase. The sources and the associated data generated by third-party producers are presented next.

Structure Inventory and Inspection Management System (SIIMS). The SIIMS data repository contains culvert structural and maintenance information. Specifically, data regarding culvert geometry, materials along with results of periodic site surveys are all organized and made available in a uniform format. The Iowa DOT SIIMS is the single source location for entering and reviewing condition information on all Iowa bridges, both local and state-owned (https://siims.iowadot.gov). The system offers web customized interfaces to store in a uniform, structured format data and information regarding culvert structural information, ownership and location, maintenance and inspections results, and observations regarding the hydraulics and sedimentation aspects in the vicinity of and within the culvert. The system can generate formal PDF reports using the data stored in the SIIMS. These reports may include photos, sketches, and inspection data, along with a variety of additional information.

National Hydrography Dataset (NHD) and Watershed Boundary Datasets (WBD). These data are used in this study to map the streams and hydrographic units (e.g. watersheds, catchments, and corridors). Tools to characterize the longitudinal and lateral stream network hydrologic connectivity are added to the IowaDOT Culverts platform to define the topology of the in-stream transport processes leading to culvert sedimentation. The National Hydrography Dataset (NHD) Plus Version 2 (NHDPlusV2) is used to develop the Digital River Network (DRN) and ancillary watersheds (McKay et al., 2012; NHDPlus, 2006). The original NHD is a publicly available repository that depicts the network of streams and rivers within the conterminous U.S. based on the digitized lines of U.S. Geological Survey (USGS) topographic quadrangle maps (Hill et al., 2016; McKay et al., 2012).

The NHDPlusV2 is a value-added product to the original NHD that integrates ridgelines from the Watershed Boundary Dataset (USGS & USDA, 2013) and the National Elevation Dataset (USGS, 2006) with the original USGS digital stream networks (USGS, 2001). The NHDPlusV2 fulfills the same role as the original NHD but has a number of considerable refinements and data extensions. The new version of the database integrates the NHD hydrologic connections (i.e. stream topology) with flow descriptors (i.e. flow directions and accumulations) across the 30-m digital elevation models (DEM) from the National Elevation Dataset (NED) improving the overall accuracy and resolution of the dataset. Meanwhile, through the integration, flow rasters derived from the NED are subsequently used to delineate catchments for each stream segment in vector format.

The new hydrographic dataset also provides a geospatial context for storing, indexing, and organizing water-related information in a relational data structure. Each hydrographic unit within the datasets is indexed with unique IDs (e.g. stream segments are indexed "COMID" in NHDPlusV2). The indexing provides the datasets with the flexibility to extend stream-associated attributes and spatial entities (e.g. stream connectivity, stream-associated catchments). In terms of data organization, the NHDPlusV2 divides the original data structure into "core" components and "extended" components. The set of "core" data components are mostly the spatial description of hydrographic units (e.g. stream networks, catchments) from the original NHD, WDB, and NED. The "extended" components are numeric descriptions that are indexed with the spatial "core" components containing additional information, such as stream topology, flow estimation, and catchment attributes (e.g. VAA table). Most of "extended" components are stored in value-added tables that can be joined with hydrographic units in the "core" components through the NHD indexing.

This study uses the spatial and numeric descriptions from the NHDPlusV2 to create digital river network (DRN) and ancillary watersheds for the state of Iowa. The study adopts the conventions (e.g. indexing, data structure) of the NHDPlusV2 geospatial framework to develop river corridor datasets (both spatial and numeric).

Stream-Catchment (StreamCat) Dataset. This dataset is the primary source for characterizing watersheds. It contains extensive synthetic summary statistics for ~2.65 million stream segments and their associated catchments within the U.S. including information on the anthropogenic impacts (Auerbach et al., 2016; Hill et al., 2016). The dataset is publicly available on EPA website (http://www2.epa.gov/national-aquatic-resource-surveys/streamcat), and provides landscape summary statistics for both local catchments and full upstream watersheds of any stream reach. Developed by the U.S. Environmental Protection Agency following the NHD data convention, the StreamCAT can be considered as an NHDPlusV2 extension to characterize the nation's rivers and streams. When connected with an existing geospatial framework of the nation's rivers and streams (NHDPlusV2), the spatial distribution of catchment and watershed characteristics can be visualized and analyzed for various watershed and stream management applications at the contiguous (a.k.a. conterminous) U.S scale (Hill et al., 2016).

The StreamCAT dataset contains a wide range of landscape metrics, characterizing stream segments and their associated catchments from a holistic view. Based on the domain and nature, metrics in the dataset can be classified into either nature layers that emphasize the pristine (i.e. normative) behaviors of stream environments, or anthropogenic layers that describe the degree of human activities within catchments. Natural layers consist of information describing land cover, soils, lithology, runoff, and topography. Anthropogenic layers include roads, dams, mines, U.S. Census data on population and housing unit densities, land use (urbanization and agriculture), imperviousness of man-made surfaces, and EPA Facilities Registry Service locations (e.g., Superfund sites). Most of the landscape metrics in StreamCAT exist in the form of statistical summaries, created for hydrographic units at different spatial scales, including local catchments and buffer areas within 100-m of stream segments.

The underlying methodologies for creating these statistical summaries are zonal statistics and tabulated areas. These summaries are accessible through ArcGIS toolboxes and ArcPy library and can be stored in many types of spatial databases (e.g. PostGIS). The StreamCat produce summaries for local catchments and 100-m buffer areas that are associated with stream segments by cross-tabulating areas of the landscape data layers (raster datasets) with the boundary of hydrographic units that define statistical zones. Through the stream topology (i.e. connectivity that defines the upstream-downstream relationship), the StreamCat accumulates local statistical summaries upstream of a river segment and translates them into watershed characteristics at various lateral scales.

Revised Universal Soil Loss Equation (RUSLE). These are modeling results used to characterize erosion produced mainly by rill- and sheet-based erosion that are easier captured than other types of upland erosion (Ganasri & Ramesh, 2016; Merritt et al., 2003). The RUSLE model was developed primarily to guide conservation planning, inventory erosion rates, and estimate sediment delivery based on additional analysis and knowledge that were unavailable when USLE was developed. RUSLE uses the same formula as USLE with improvements in determining factors. These include some new and revised isoerodent maps, a time-varying approach for soil erodibility factor; a sub-factor approach for evaluating the cover-management factor, a new equation to reflect slope length and steepness, and new conservation-practice values. Input parameters required by RUSLE include the following information; the rainfall erosivity factor (R), the soil erodibility factor (K), the slope length (L), gradient (S), the cropping management factor (C), and the erosion control practice factor (C).

While the capabilities to accurately predict erosion and sedimentation are limited, the RUSLE model represents a good indicator of the severity of soil loss in landscapes such as Iowa. This study uses annual soil loss produced from RUSLE as an independent variable to estimate the sediment potential of headwater areas where culverts are located. As culvert sedimentation involves other aspects, such as gully erosion, sediment transport, and sediment deposition, which are not well addressed by RUSLE alone, the model will be used only as a reference for the severity of soil losses in conjunction with the MCDA secondary analysis. The primary analysis of culvert sedimentation will not rely on the simulation provided by RUSLE and its siblings (e.g. USLE, RUSLE v2, MUSLE).

Annual Stream Exceedance-Probability Statistics (Eash Method). Currently, Iowa DOT uses flood-frequency analyses to estimate the design discharge for culverts. Two major techniques that are adopted by the department's Iowa Bridge Backwater Software are the USGS Lara method and the USGS Eash method. Both methods use regional regression equations to calculate the magnitude and frequency of floods (annual exceedance-probability discharges) at ungagged sites on unregulated rural streams in Iowa. The USGS guidelines distinguish several hydrologic regions for the state of Iowa (5 regions in the Lara method and 3 regions in the original Eash method). Each hydrologic region has local flood-frequency equations and specific regression equation coefficients. As an example, the local flood-frequency equations appear in the form: $Q_t = cAb$;

where Q_t is the discharge for a selected recurrence interval; A is the drainage area upstream of the culvert; and b and c are regression equation coefficients (Jones, 2013, p. 28). The newer version of the USGS Eash method includes a three-parameter version which takes into consideration the main channel slope (MCS) and DML (Des Moines Lobe) parameters to improve the estimation.

Aerial Imagery. The present study relies extensively on aerial images for quantifying the degree of sedimentation at culverts and to establish the stream-to-culvert width ratio. The degree of culvert sedimentation is defined as the area of the original inlet and outlet clearance occupied by sediment deposits. Aerial images are retrieved from various sources (e.g. NAIP, NDOP, Google, Bingmap, and ESRI). These are popular cyber-GIS technologies available online that efficiently support the MCDA on culvert sedimentation developed in the next phase of the study. Culvert identification using the aerial photographs can be conveniently conducted using the Iowa Geographic Map Server – IGMS (http://ortho.gis.iastate.edu). This resource site contains historical and current images with spatial resolution up to 2ft for surveys acquired from 2007-2010. In addition, the study explores information embedded in Google Earth (www.google.com/earth) for complementing missing information in IGMS. The two resources include easy-to-use features and imagery, and specific tools that enable the location of targeted areas, map areas of interest using topographic or imagery layers, compute distances and areas, and capture and annotate images. In the later stage of this study, aerial images provided from EagleView Technologies are ingested into the analysis to increase the sample size of culverts, as well as to validate the quantification of the culvert sedimentation degree that are developed using other sources. Unlike IGMS and Google, EagleView Technologies is a commercial provider of aerial imagery, data analytics, and geographic information system. The company offers high-quality time-series aerial images (resolution up to 3 ft.) that cover most areas in Iowa.

In-situ terrestrial surveys. During this study, the survey of more than 250 three-box culvert sites was conducted in-situ for garnering the project-specific data and information needed for the MCDA analysis. The map of the in-situ surveyed culverts during the 2016 and 2017 field campaigns is shown in Figure 5.The surveys were conducted in early March and April, before the spring vegetation growth began. The data was uploaded in real time using an interface of the IDOT Culvert portal.

The culvert surveys goals included documentation of the following aspects:

- the degree of sedimentation at the culvert inlet (photo-documentation).
- the degree of blockage at the culvert entrance cross-section (survey).
- critical features characterizing relationships between culvert structures and the associated stream, as well as specifications on sediment deposits (notes).

The field inspections were conducted using a rigorous experimental protocol. Figure 3.3 documents the organization of the protocols used in conjunction with the photo documentation and tracing of the sediment blockage at the culvert entrance. The instrumentation used in conjunction with these protocols were computer-embedded cameras, laser-based rangers, and portable communication hubs. This type of survey is conducted at more than 257 culverts sites, displayed on the map in Figure 3.2.

The information compiled from the from the photo-documentation, along with the field survey notes, can be uploaded in real time using a customized interface of the "IowaDOT Culverts" portal, as illustrated in Figure 3.4. This field data uploader is an efficient tool for supporting periodic inspections conducted by the culverts' maintenance personnel. The rigor of the data acquisition protocol and the quickness of the field data acquisition represents a significant improvement over the piece-meal approach and laborious procedures currently used by county and IDOT district engineers for conducting the routine culvert inspections.

Drone-based surveys. For selected sites, the photo documentation was carried out with drone-based surveys subsequently processed with Structure from Motion (SFM) software. This contemporary type of survey is increasingly popular due to its low cost and simple deployment. As most civil drones are equipped with digital cameras (with continuously increased resolution and positioning accuracy), they become reliable substitutes for aerial photo surveys. Drone-based surveys are compared to in-situ photography in Figure 3.5. Intermediary steps leading to the orthorectified aerial photo using the SFM software are shown in Figure 3.6.

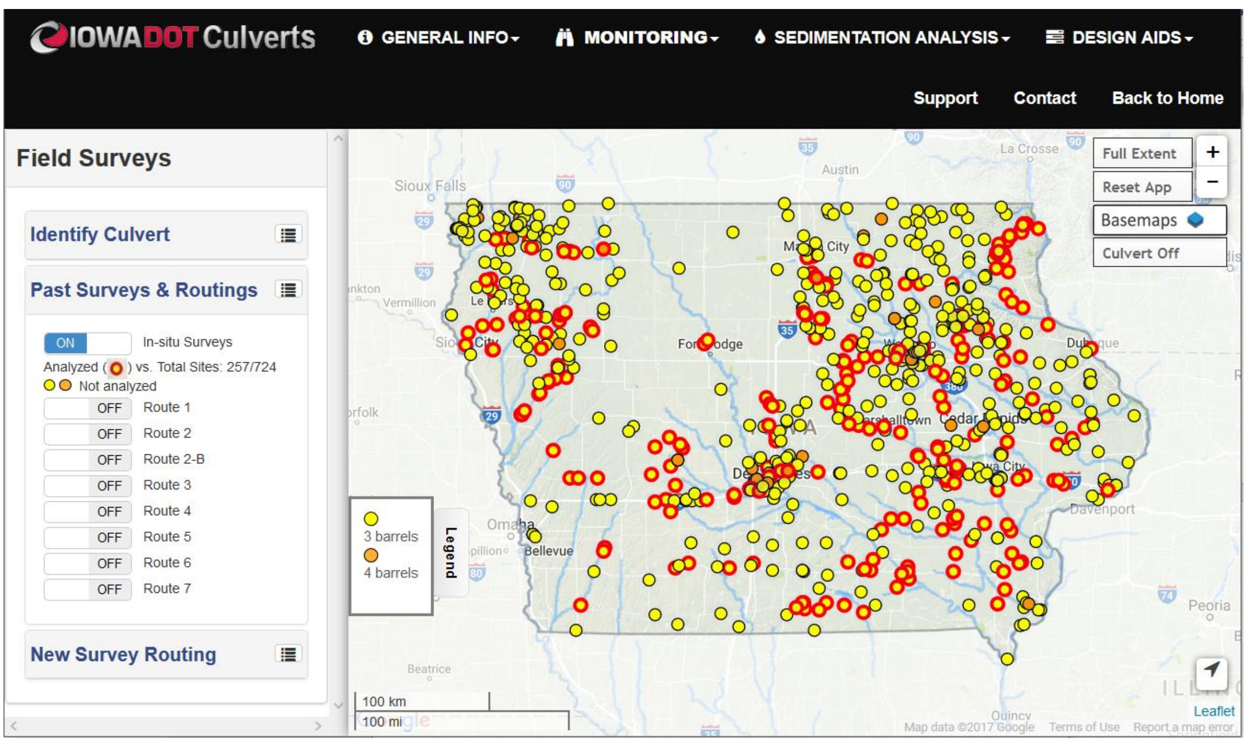


Figure 3.2 In-situ culvert surveys conducted in 2016 and 2017 (red symbols)

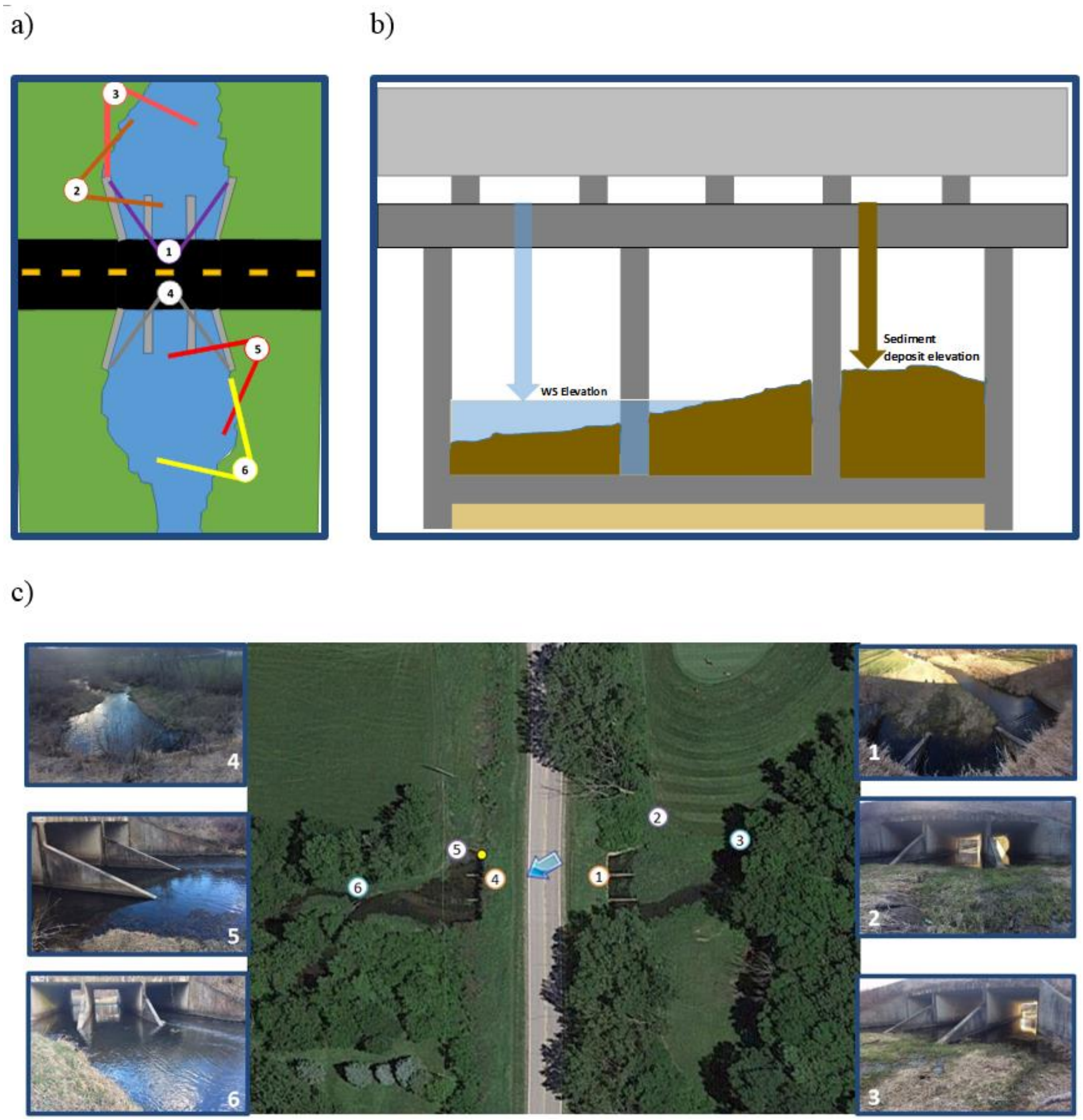


Figure 3.3 Field measurement protocols: a) positioning of the photo camera for the survey; b) illustration of the measurement acquired for the degree of sedimentation blockage; and c) illustration of the photo-documentation acquired at the culvert site

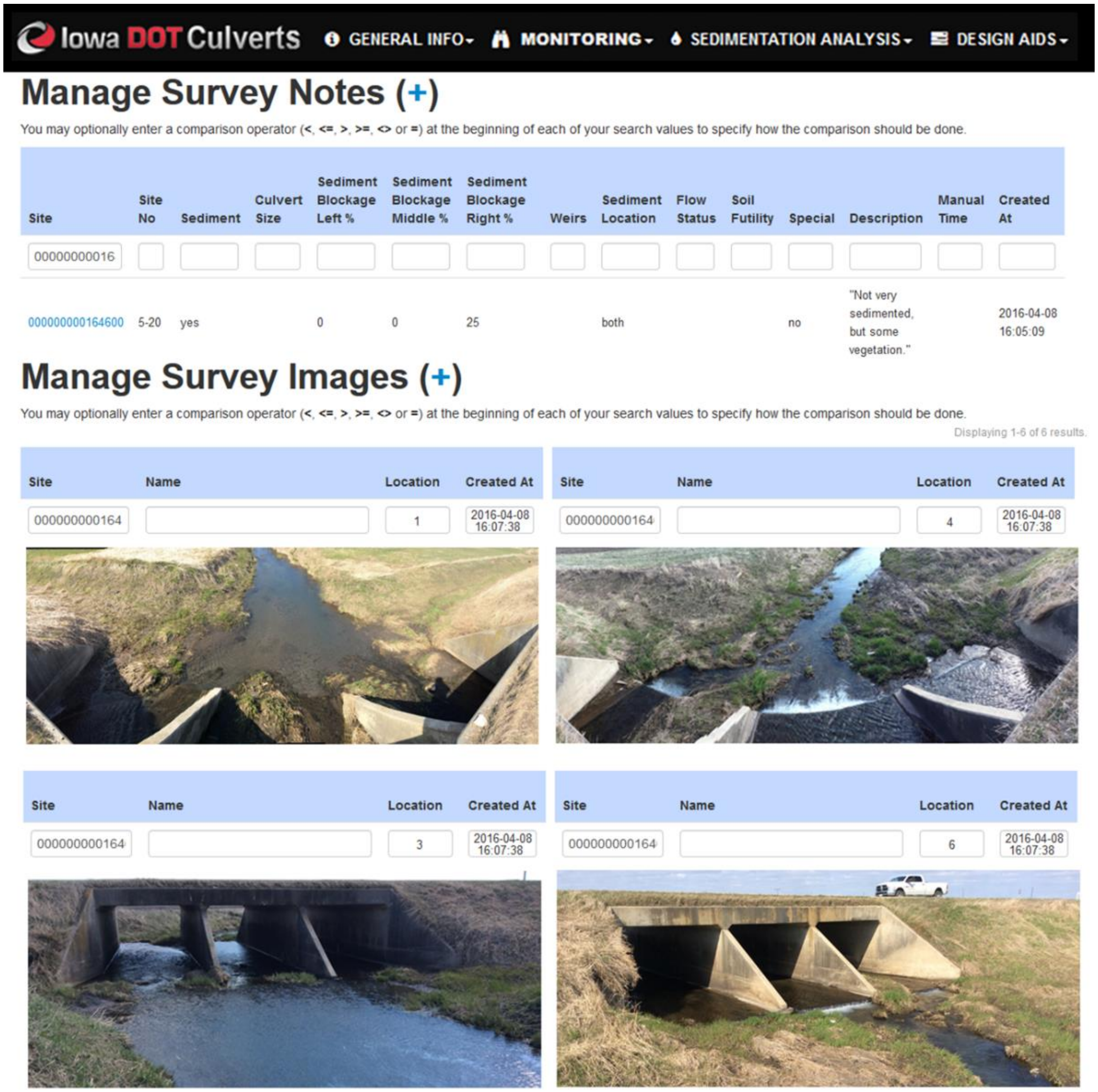


Figure 3.4 Illustration of the web-database storing the field surveys

The in-situ surveys with Real Time Kinetic (RTK) GPS instrumentation as well as the drone-based surveys of the sediment deposits allow to efficiently quantify both the degree of culvert sedimentation and the blockage at the culvert inlet and outlet. These indicators define the most important parameters of the investigated functional relationship, hence accurately estimating them from field data is critical. The role of the in-situ surveys was to compare the such-obtained estimates with those based on aerial photographs obtained, as described next.

a) b)

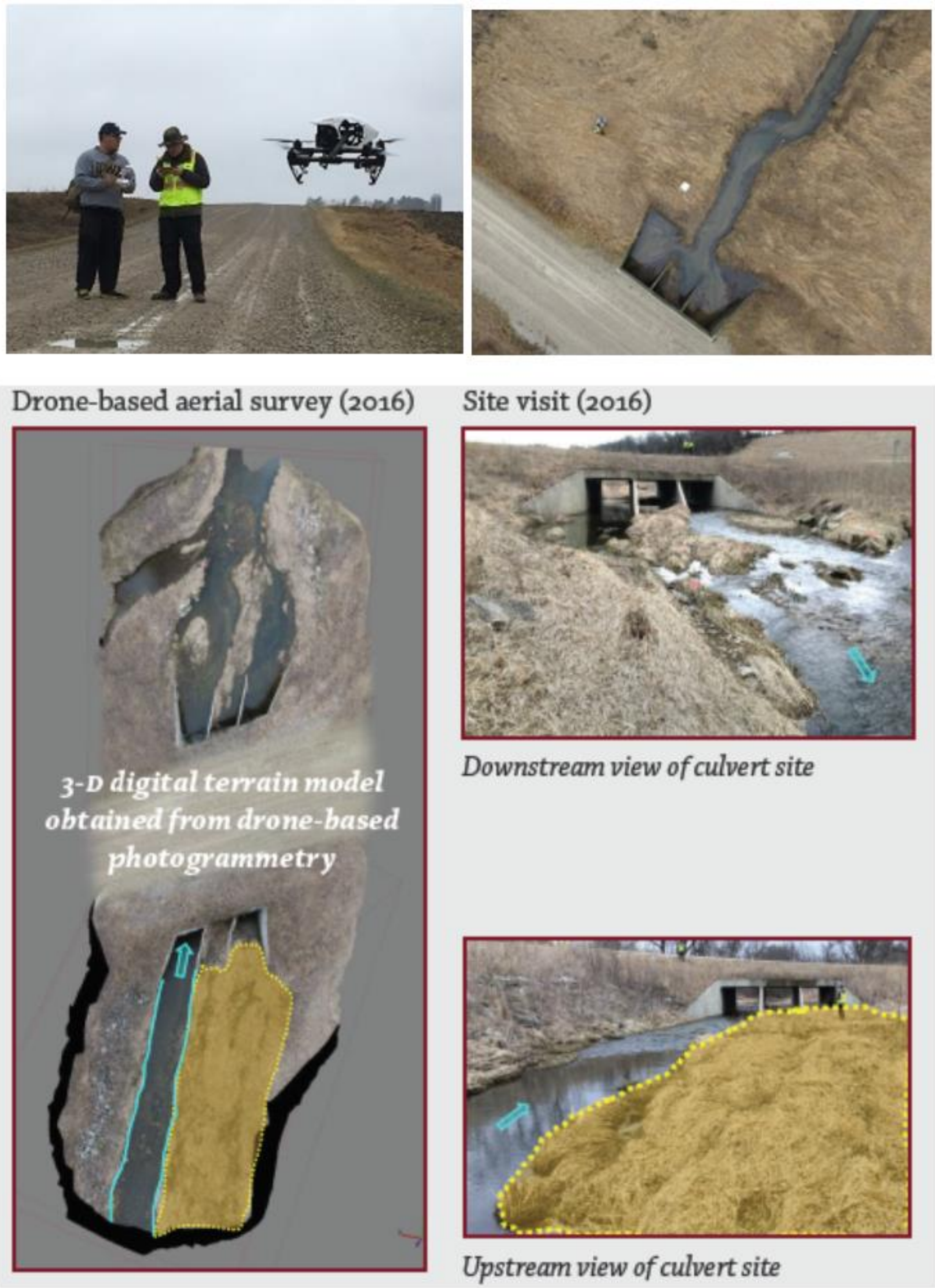
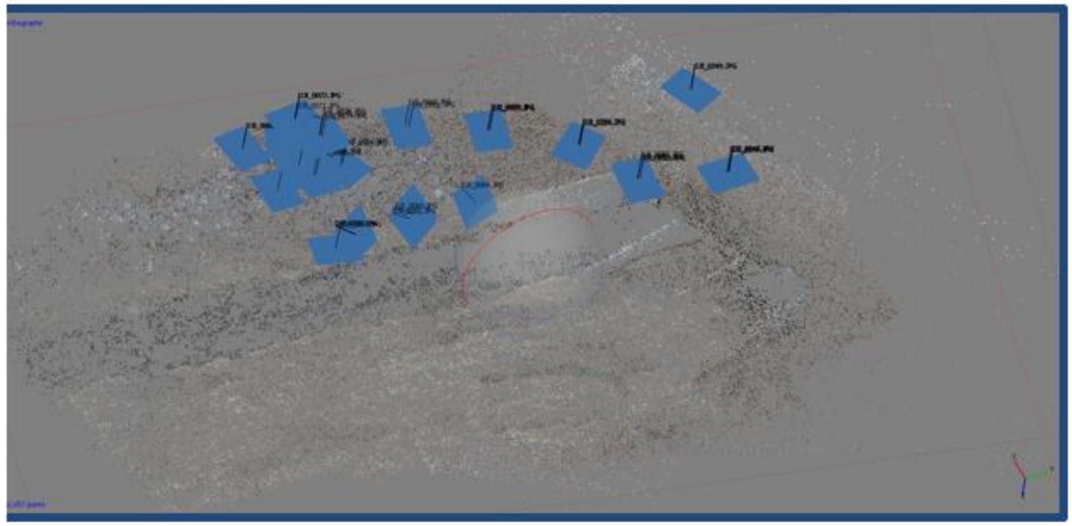


Figure 3.5 Drone-based culvert surveys: a) data acquisition; b) processed data and comparison between aerial- and ground-based photo-documentation

a)

b)



c)

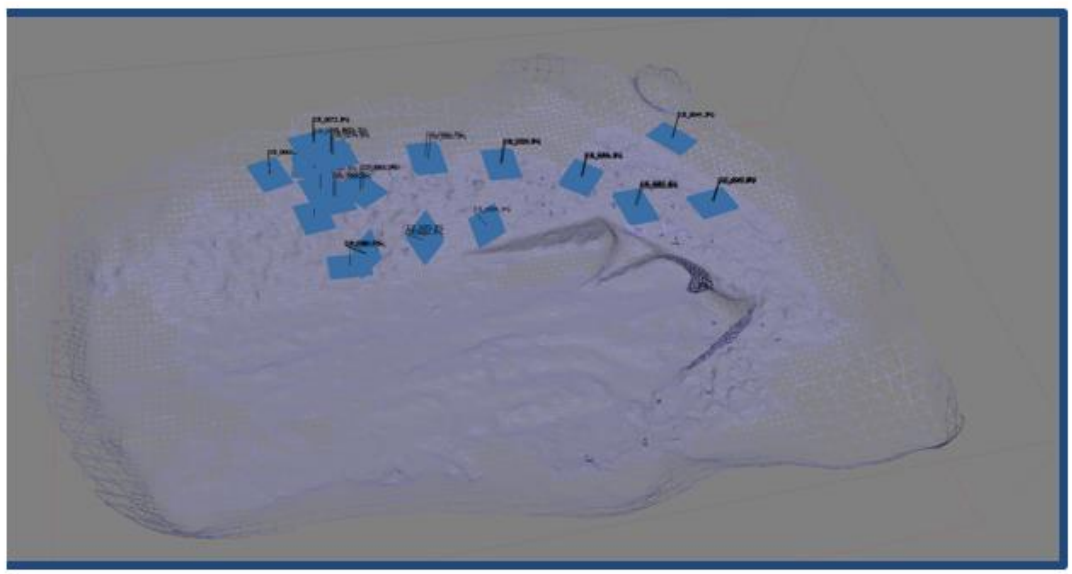


Figure 3.6 Structure from Motion (SFM) image reconstruction: a) raw images of the culvert inlet; b) image stitching; and c) identification of the tie points for image reconstruction

Aerial imagery surveys. As a surrogate for the in-situ surveys, processing of the aerial photographs was considered as an alternative approach to estimate the degree of sedimentation. For this purpose, the study screens the whole 3-box culvert database made available by IDOT for assessing the spatial extent of the sedimentation. The screening involved first a classification of the culvert "visibility" in a series of recent base maps. Geo-processing tools designed to work on the top of the maps enabled to identify and map the area that was originally cleaned at the time of culvert construction. The sediment deposit, as found in the most recent aerial photograph, was subsequently traced on the photograph to establish the current degree of sedimentation upstream and downstream the culvert, as illustrated in Figure 3.7. The geo-referenced mapping is made in real-time on any type of maps stored in the portals' map database. The outcomes of the on-screen measurements of the segment lengths and polygon areas are displayed on the screen as the measurement is made. The mapping data is saved in the "Iowa Culvert Sedimentation Observation" database.

The visual inspection of the aerial photographs also allows for the identification of the culverts with problems, tracing the patterns of sediment development, rates of evolution over time (at the culvert locations where this information is available), and the formulation of the correlations between stream geometry, culvert geometry, and the nature of the drainage area upstream from the culvert site. Simple and intuitive web-mapping tools were developed to allow users of the platform to extract spatial attributes of sedimentation and their evolution in time (the best quality for the images are for the 2004-2016 time interval).

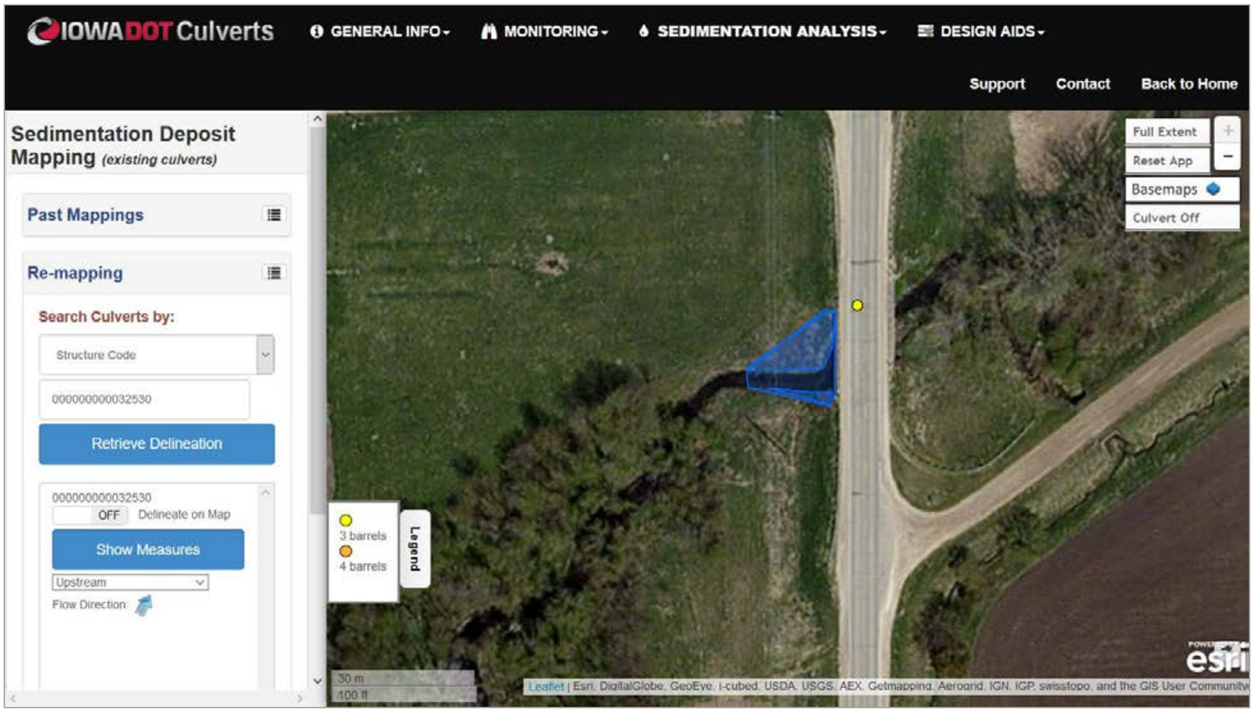


Figure 3.7 Geo-portal interface for the geo-processing tools associated with the estimation of the degree of sedimentation at culverts

3.3 Multiple-Criteria Decision Analysis (MCDA)

This study utilizes Multiple-Criteria Decision Analysis (MCDA) to inform on the culvert sedimentation potential, thereby directly aiding in the design and operation of the multi-box culverts. The MCDA is a visual analytics approach that uses quantitative and qualitative data, along with expert judgment, to support decisions even if the data used in the analysis are not abundant (Vulevic & Dragovic, 2017). The MCDA pertains to the cross-disciplinary research that looks for ways to provide computer support to solving space-related decision problems through enhancing human capabilities to analyze, envision, reason, and deliberate (Andrienko et al., 2007). The MCDA enables decision making when the number of alternatives or actions is evaluated in terms of more than one, usually conflicting, criteria. It is important to emphasize upfront that the outcomes of the data-driven investigative approach are only as good qualitatively as the data used in the analysis.

According to Belton & Stewart (2002), MCDA methods can be classified as measurement models, outranking methods, and goal/aspiration models. Irrespective of the method, special attention is needed in the initial structuring of the decision problem, which involves: a) selection of criteria, (b) selection of decision options, (c) weighting the criteria, and (d) obtaining performance measures to populate the evaluation matrix. Many researchers have found that MCDA provides an effective tool for water management by adding structure, auditability, transparency, and rigor to decisions (Hajkowicz and Higgins, 2008). The MCDA usage in the present context has the following purposes:

1. Ranking environmental and structural drivers behind culvert sedimentation using feature selection techniques.
2. Developing quantified relationships between culvert sediment degrees and key drivers within each region using deductive hydrological classification.

Development of functional relationships. From a technical perspective, the use of the decision tree enables reliable quantifications of the cause-effect relationships between culvert sedimentation and its drivers. The strength of the method resides in the fact that there is no need to elucidate complex, inter-related processes and their evolution over large spatio-temporal scales as it is done in physical-based modeling approaches. Instead, the dependence among the variables in the relationships is developed by training smart algorithms on a set of data to develop the predictive capabilities needed in forecasting. The analysis outcome is dynamically visualized with parallel coordinate plots and principal component charts. This interaction facilitates the integration of human judgment in the final stage of the decision support loop.

Developing quantitative relationships between the culvert sedimentation degree and the key drivers within the culvert drainage area is accomplished in this study using the hydrologic classification (Olden et al., 2012). The hydrologic classification was originally defined as the process of systematically arranging streams or rivers into groups that are similar with respect to characteristics or determinants of their flow regime (Kampichler et al., 2010; Olden et al., 2008). The classification can be categorized either as (Auerbach et al., 2016; Olden et al., 2012; Wagener et al., 2007): a) inductive (whereby classifications are made based on statistical similarity in the hydrologic data directly), or b) deductive (whereby environmental variables such as watershed characteristics are analyzed as key drivers of hydrology to create classifications). As the inductive approach requires abundant hydrologic data over a very long period of time, the data available for sedimentation at culverts in Iowa is not sufficient enough to produce such empirical relationships. Thus, this study uses the deductive approach. The drivers selected for classification include the watershed and stream variables identified in Section 2.2. This hydrologic classification coupled with machine learning techniques (Auerbach et al., 2016; Lins, 1985) is used to establish cause-effect relationships between culvert sedimentation and its drivers within each culvert drainage area.

The learning technique proposed for conducting the criteria analysis is the tree ensembles. Tree ensemble method uses multiple decision tree algorithms to obtain better predictive performance than those obtained from any of the constituent trees alone. The tree-ensemble method derives the empirical relationship using a tree-like model of the key drivers and their possible consequences regarding the culvert sedimentation degree. The method is a popular supervised machine-learning technique applied in hydrologic research (Galelli & Castelletti, 2013; Schnier & Cai, 2014; Schnier, 2016). In terms of training and testing, all culvert sediment observations are used as training datasets. The analysis outcomes are the results of decision-trees empirical relationships such as the one illustrated in Figure 3.8.

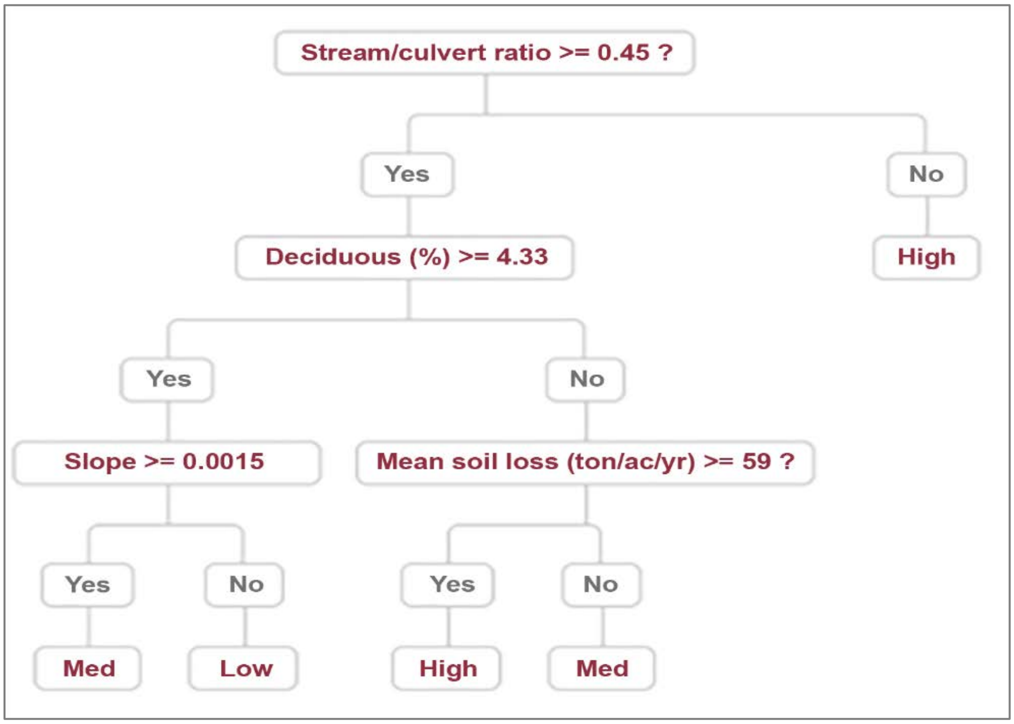


Figure 3.8 Decision tree representation of the cause-effect relationship between culvert sedimentation and its drivers

3.4 Culvert sedimentation forecasting

By assembling the outcomes of the data-driven activities described in the previous section in a web system, this study obtains a useful forecasting tool for sedimentation forecasting. To provide data-driven decision supports for mitigating culvert sedimentation, this study uses derived relationships that are capable to predict the severity of culvert sedimentation at any location and stream in Iowa.

Since environmental and ecological factors, and human activities vary across different areas in Iowa, it would be beneficial to integrate the capabilities of GIS and MCDA for tackling spatial decision-making problems (Jelokhani-Niaraki et al., 2018). This combined approach is defined as the GIS-based Multi-criteria Decision Analysis (GIS-MCDA) and have been adopted by many past studies for solving a wide variety of environmental decision and assessment problems while considering the multiple criteria and decision makers' preferences (Chen et al., 2011; Crossman et al., 2011; Labiosa et al., 2013; Sadeghi-Niaraki et al., 2011). In the context of this study, the GIS component provides powerful capabilities for storing, managing, and visualizing geographically the culvert-related environmental data (drivers) for the MCDA. The MCDA provides the algorithms for structuring the problem and evaluating the contributions of different drivers to the culvert sedimentation degree. In terms of the sedimentation forecasting, the GIS component retrieves the environmental attributes that are associated with the user-defined point of interests, and import these data into the MCDA (based on the decision trees). The MCDA then evaluates these environmental attributes as different criteria and provides the corresponding forecast on the degree of culvert sedimentation at the user-defined location.

All the features described in Section 3.2-3.4 are hosted in a customized web-GIS portal developed for Iowa DOT. The workflows hosted by the portal allows users to: (1) integrate, access, store, and manage culvert related information; (2) provide a web-based problem-solving platform for conducting data-driven investigations; and (3) access, analyze and visualize the data and information with user-friendly interactive interfaces. The previously mentioned workflows play the role of identifying, accessing and visualization various type of data and information about the culverts of interest. The forecasting features are built atop and it is complemented by the above workflows with the role to assist designers and culvert maintenance personnel on the potential for sedimentation at an existing of new culvert site location.

All workflows and the associated tools and functions, along with the associated user interfaces, are hosted in the IowaDOT Culverts platform residing at http://iowawatersheds.org/idotculverts. The users of the platform can predict the degree of culvert sediment for any stream location in Iowa. Given that all the variables, excepting the evaluation of the degree of sedimentation, are available for the contiguous US, the platform can be easily extended to other U.S. territories. In the initial stages of its development, the culvert sedimentation forecasting has been relatively coarse. Further improvements can be obtained by conducting sensitivity analysis on different spatial units (i.e., consideration of watershed characteristics within sub-areas of the culvert drainage area or imposed buffer regions along the streams) and regionalization of the MCDA and forecasting of the erosion potential through clustering and partitioning. These optimization features are described next.

3.5 Sediment forecast optimization

3.5.1 Background

While the MCDA provides such benefits, there are several considerations that need to be accounted for when using the analysis outcomes. A first caution in the result interpretation is that most of the key drivers analyzed through the MCDA are characterizing the whole drainage basin (watershed) associated with the culverts. While dealing with transport through the natural landscape can potentially be directly linked to the entire drainage basins, the movement of the sediment over the landscape is susceptible to local retentions and other discontinuities of the terrain that are not typically revealed by working at the whole-basin scale. Given the potential discontinuities, it is useful to conduct a sensitivity analysis that explores the degree of dependence of a specific variable with alternative spatial extents. The most appropriate spatial analysis for our context is the River Continuum Concept developed by Blankenship et al. (2000). The spatial extents associated with this concept are derived based on the principle of multi-dimensional hydrologic connectivity within the culvert drainage basin which includes the natural connection among: (a) the whole drainage area, (b) river corridor over the entire drainage basin, and (c) immediate corridor upstream of a culvert.

The second cautionary item related to MCDA outcome is that the initial efforts for establishing relationships between the process-drivers and the culvert sedimentation degree carried out for this study used as sample population all the three-box culverts available in the Iowa SIIMS database. The obtained forecast is relatively coarse due to overfitting, the situation that occurs when using a single model for culverts located over the whole state that however can considerably differ from region to region with respect to erosion and sediment transport potential. More specifically, the tree-based representation of the drivers-sedimentation relationships of the state might be too generic and complex for informing culvert designers and managers on the potential of culvert sedimentation in another region of the state (e.g., a county or an Iowa DOT maintenance garage). In this regard, it seems rational to partition Iowa into sub-regions with homogeneous erosion potential and soil mobility, and analyze the culvert sedimentation potential and response separately for each region.

The third cautionary note is related to the fact that the use of decision trees for quantifying the relationships between the culvert sedimentation degree and the process drivers has limited ability to capture the multivariate relationships (patterns) in the culvert sedimentation dataset, as the relationship can potentially represent the combined effects of multiple process drivers (of different value ranges) on the degree of culvert sedimentation. In other words, there are potentially multiple combinations of process drivers of well-defined ranges that can lead to the same degree of sedimentation. For example, the combination of certain proportions of agricultural areas, terrain, slope, and design discharges can lead to the same degree of sedimentation as another combination for the same specific geographic location and spatial extent.

To address the above-mentioned concerns in the interpretation of the MCDA-based forecasting three optimization procedures are developed and tested through this study: (1) sensitivity analysis of the spatial extent for the key drivers, (2) regionalization of the analysis, and (3) multivariate clustering analysis for culverts in each region.

3.5.2 Spatial extent sensitivity analysis

The forecasting procedure introduced Section 3.4 utilizes watershed characteristics provided in the StreamCAT dataset to relate erosion and watershed sediment transport characteristics with culvert sedimentation. The forecasting is conducted under the assumption that the stream network integrates a large portion of the hydrologic characteristics in the upstream watershed (Allan, 2004). However, it may not be appropriate to directly relate watershed characteristics to local river features, such as sediment loads and nutrient loads (Hill et al., 2016), because these features entail environmental transport processes that are affected by the multi-dimensional stream connectivity in fluvial systems and the presence of discontinuity on overland transport. The concept of multi-dimensional stream connectivity, originating from the River Continuum Concept (RCC), specifies that rivers have interactive pathways along three river spatial dimensions (Demir and Szczepanek, 2017; Tockner and Stanford, 2002). According to Ward and Stanford (1989) and Ward (1989), the three representative river dimensions are longitudinal (upstream-downstream), lateral (channel-bank / floodplain / corridors), and vertical (atmosphere channel – subsurface). According to the dynamic interactions along the longitudinal and lateral dimensions are proven to have strong influences on both sediment budgets and sediment transport processes (Tockner and Stanford, 2002; and Wohl et al., 2015). In this regard, other spatial extents that are derived from the stream connectivity might play a more important role in sediment-related processes. To improve the characterizations of drivers, the present study conducts a sensitivity analysis of different spatial extents to identify the most relevant spatial extent for each environmental driver.

Before proceeding with the actual analysis, a definition of the fours spatial extents associated with the concept of multi-dimensional stream connectivity as used in the present study is provided in Figure 3.9. The spatial extents are defined around the convention of National Hydrography Dataset (NHD). A widely used hydrographic concept adopted by many government agencies, research, and watershed communities.

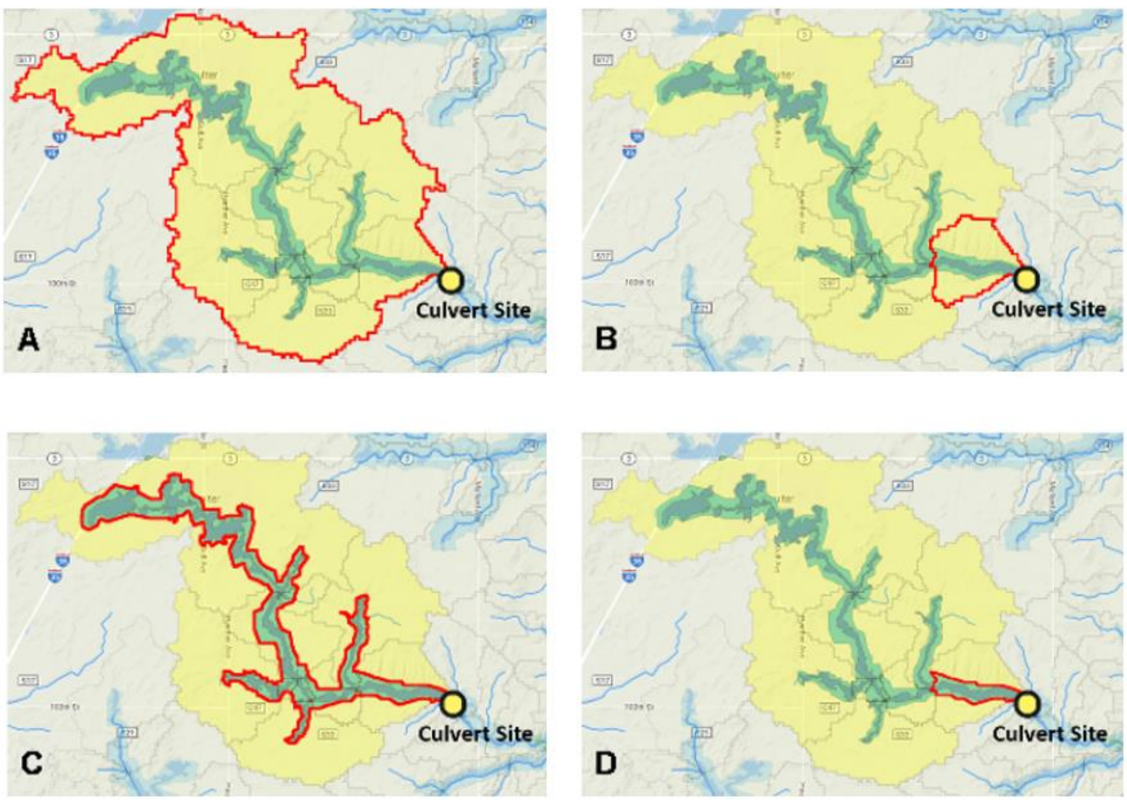


Figure 3.9 Spatial extents for characterizing culvert sedimentation drivers (highlighted with red lines on the maps): (a) drainage area, (b) catchment, (c) river corridor, and (d) immediate corridor upstream of culvert

The "watershed" and "catchment" illustrated in Figures 3.9a and 3.9b respectively, are the two major components associated with the longitudinal river dimension. According to the NHD-associated Watershed Boundary Dataset (WBD), the catchments represent the fraction of the landscape where surface flow drains directly into a stream segment, excluding any upstream contributions. Stream segments and its associated catchment are considered "headwater" when they do not have an upstream drainage area. A catchment is a local and atomic area associated to a single stream segment, whereas the term "watershed" is a broad unit that refers to the set of hydrologically connected catchments, consisting of all upstream catchments that contribute flow into any catchment (Hill et al., 2016). Along the longitudinal direction of hydrologically connected catchments, the hydrologic and environmental characteristics of individual catchments accumulate into watershed characteristics from downstream to upstream, which is referred to "feature accumulation" in many studies (Hill et al., 2016). The current hydrological terminology the U.S. uses the term "watershed" to refer to the topographically delineated area that is drained by a stream system (Easter & Hufschmidt, 1985). However, watersheds are used interchangeably with "catchment" and "basin" in many research articles (Hill et al., 2016).

The "river corridor" and "immediate corridor upstream of culvert" illustrated in Figures 3.9c and 3.9d, respectively, are spatial extents associated with the lateral river dimension but there are somewhat also related with the longitudinal direction. The concept of "corridor" has been used to describe the active stream-bank interaction area around the wetted river channel, which is formed through the periodic inundation of the floodplain (Freeman et al., 2007; MNDNR, 2016). The spatial extent of a corridor within an individual catchment area is typically narrower than that of the lateral dimension of the catchment for the same length of the river corridor. The corridor is an active area of sediment production and deposition. Soil particles dislocated in the corridor area are most likely to be delivered into the stream compared to those traveling overland as the transport distance is shorter. By contrast, active riparian vegetation growth within the corridor can potentially prevent stream-bank erosion and retain the lateral sediment inputs (delivered from overland transport).

It is obvious that the stream corridors as defined above are important spatial extents for the culvert sedimentation may only occur within the stream corridors, thus this spatial extent is very relevant to the objectives of the study. Building on the experiences of previous studies (e.g., Carlisle et al., 2009; Hill et al., 2016), this study considers 100-m buffer areas around the stream centerline segments to represent the extent of river corridors. To immediate corridor upstream of culverts for the present study is defined as the 100-m buffer area around the stream centerline segments extending 2500 meter upstream from the location of a culvert. This longitudinal distance is determined based on the statistics of culverts' total upstream length across Iowa, as most culverts in headwater areas have their upstream channels less than 2500 meter.

A necessary precursor of the spatial sensitivity analysis is the evaluation of the statistical summaries for the erosion and sediment transport parameters relevant to overland and in-stream processes for the spatial extents shown in Figure 3.9. The statistical summaries for each zone are built with geo-processing tools available from ArcGIS toolboxes, ArcPy library, and spatial databases (e.g. PostGIS). The use of these tools enables to assemble the data presented in Section 3.2 to and create a comprehensive culvert dataset for characterizations of the independent variables for all four different spatial extents.

The optimization forecasting also considers two additional variables compared with the initial dependency datasets. The added variables are stream sinuosity and stream width to account for the geomorphological status of the river in the vicinity of the culvert. In this study, the stream sinuosity (shown in Figure 3.10) is calculated as the ratio of stream length (red line) to valley length (orange line) (Rosgen, 1998). The data needed to estimate stream sinuosity are readily available from the

NHD Plus Dataset, while the width can be obtained from aerial imagery-based delineations using the dedicated Geo-processing tool used to determine the culvert sedimentation degree (see Section 3.4).

Meta-data are created to further categorizes all the independent variables based on their associated spatial extended, as “stream characteristics”, “watershed characteristics” or “stream and watershed characteristics”. This implementation aims to separate the visualizations presented in later analyses, so as to help us explore the contributors that are pertaining to specific spatial extents (watersheds vs stream/corridors).

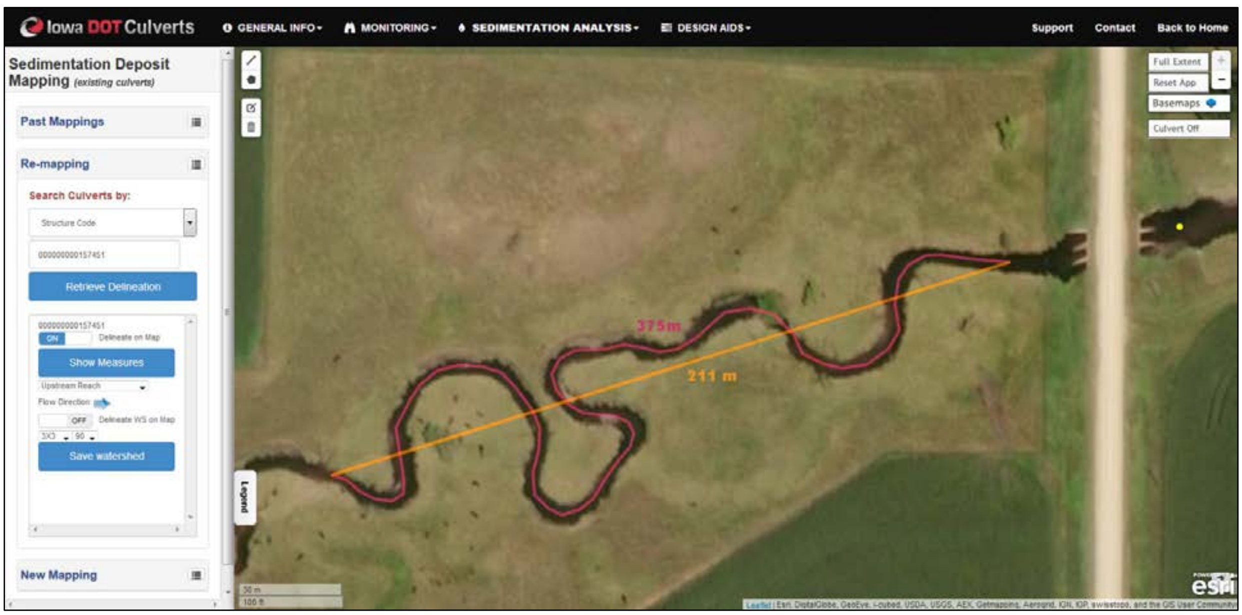


Figure 3.10 Geo-portal interface for illustrating the geo-processing tools associated with the estimation of stream sinuosity

Although all the independent variables, together with their spatial extents, are pre-selected based on hydrologic and sediment transport domain knowledge, it is assumed that not all of them are equally contributing to the culvert sedimentation processes under investigation. The large quantity of the contributing variables involved in natural-scale environmental processes and their possible correlations poses a significant challenge for analysis by hiding or diluting the multifunctional relationships among the most relevant variables. In order to overcome this limitation, the study adopts a feature selection strategy (Liu and Motoda, 1998) with the purpose of identifying the most significant environment variable and spatial extents that contributes to culvert sedimentation processes. A tree-based feature selection algorithm from the Python Scikit-learn library (Buitinck et al., 2013) is adopted herein to rank the independent variables (together with their associated spatial extents) based on their relevance to the classified sedimentation severity. The algorithm utilizes a forest of decision trees to evaluate the importance of each variable based on the average impurity decrease (characterized through the information gain/entropy reduction). While there are several feature selection algorithms available, the justification for selecting this method is based on the following premises: (1) the performance of decision trees is not affected by non-linear relationships, (2) the method requires less data transformation, and (3) the output of decision tree-based ranking is intuitive and easy-to-understand.

After applying the machine learning-based feature selection, visualization of the dependency between the selected variables and the degrees of culvert sedimentation in a Parallel Coordinates Plot (PCP) for human interferences. This visualization technique can present a multidimensional space through a two-dimensional display by using parallel axes to represent variables (Edsall, 2003; Inselberg, 1985) from which the correlations between drivers in different axes can be easily understood.

3.5.3 Regionalization

Regionalization in the environmental research context is defined as the process of grouping spatial units into contiguous or non-contiguous regions that are considered homogeneous in regard to certain environmental characteristics at a particular scale (e.g., Bryce and Clarke, 1996; Loveland and Merchant, 2004; Olden et al., 2012). This process is widely used by past studies to identify groups of hydrographic units (e.g., watersheds and catchments) which have similar characteristics (Yadav et al., 2007; Zhang et al., 2008), with the objectives of improving hydrologic prediction, model calibration, and uncertainty reduction.

Due to the heterogeneous distribution of environmental variables and the spatial variability of sedimentation potential, the empirical relationship between drivers and culvert sedimentation may vary in different regions of Iowa. Through regionalization, the present study aims to distinguish areas/watersheds in Iowa based on the similarities of environmental drivers that are responsible for culvert sedimentation, as well as to generate multiple sub-regions with explicit and uniform sedimentation potential (e.g. runoff behaviors, channel erosion potential, and geomorphological features).

In order to identify regions with similar erosion potential and culvert sedimentation response, this study utilizes a Self-Organizing Map (SOM) to group culverts (with their associated drainage area) into clusters based on their multivariate similarities in drivers. The Self-Organizing Map (SOM) is a multivariate clustering technique that utilizes an artificial neural network (ANN) to preserve and abstract patterns in the dataset (Kohonen, 2001; Skupin and Fabrikant 2003; Koua and Kraak 2004; Koylu 2014; and Guo, 2015). In this study, the SOM is preferred over other clustering algorithms (e.g., K-Means Clustering) because it can provide an effective data visualization to convert complex, nonlinear relationships between high-dimensional data items into simple geometric relationships through a 2-D map (i.e., a SOM map created with color schema) (Kohonen, 2001).

During the SOM analysis, this study normalizes the state-wide culvert sedimentation data using the minimum and maximum values (a.k.a. feature scaling techniques) and defined a SOM network of 15x15 nodes (size). Through a series of preliminary testing and experiments, the study discovers that the size of 15x15 allows the SOM to produce clusters with a minimum value range of the culvert sedimentation degree, indicating that each cluster is fully separated and has relatively low entropy. If the SOM network is undersized, there won't be enough neurons in the network to separated culverts with complex characteristics, and the SOM is likely to put all the culvert sites into a single cluster. It is also important to avoid oversizing the SOM, as large networks are likely to overclassify the input data with unnecessary details and produce a number of clusters that only contain 1 or 2 data items. Additionally, constructing a 2D color scheme to visualize large SOM networks is very difficult, as it requires a large number of colors (Guo et al., 2005). Other SOM parameters (e.g., learning rates and initial radius) are determined based on the range and scale of the input data, following the instructions provided in Kohonen (2001).

To bring human intervention into the analysis, a web-based visual interface is developed that allows users to test the outcomes of different SOM configurations (e.g., colors, size, and learning rates) by providing evaluative feedbacks of the SOM performance (clusters size, average cluster range, and entropy) in Figure 3.11. Figures 3.11a and 3.11b illustrates different SOM configurations and evaluative feedback, respectively. The new interface (illustrated as Figure 3.13) is an integral part of the "IowaDOT Culverts" platform as an extended component of the MCDA.

a)

b)

Cluster Size	1-4	5-9	10-19	20-99	>100
Counts	27	0	1	1	0
Max Cluster Range: 0.27			Number of Cluster: 29		

Figure 3.11 Visual interface for (a) modifying the SOM configuration, (b) displaying the evaluative feedbacks of the SOM outputs

The interface also provides a range of visualization techniques to display and explore the SOM outputs from different perspectives. The SOM clusters are visualized through an interactive 2D map that adopts a hexagonal layout of SOM nodes (illustrated as Figure 3.12a). Each hexagon in the map represents a node, containing a cluster of culverts that are similar to each other and are different from those in other nodes (shown in Figure 3.12b). The orientation of the SOM map also indicates that neighboring nodes (hexagon) are similar to each other in multivariate space. A 2-D color scheme is them applied to the SOM map rendering each node with a discrete color (cluster-color). Afterward, the study connects the 2D map to a modified PCP that bundles edges (Palmas, 2014) by the cluster-average value of each variable and renders them using their associated cluster-color (depicted through Figure 3.12c).

In the context of regionalization, this study synchronizes the SOM map with a geographic map to display the spatial distribution of different culverts clusters produced by the SOM. In the geographic map, the location of a culvert is displayed as dots and are color-coded using the color that is assigned to the culvert's pertaining cluster (as shown in Figure 3.12b). The degree of culvert sedimentation at each site is represented as a normalized radius of the dot symbology (as illustrated in Figure 3.12c).

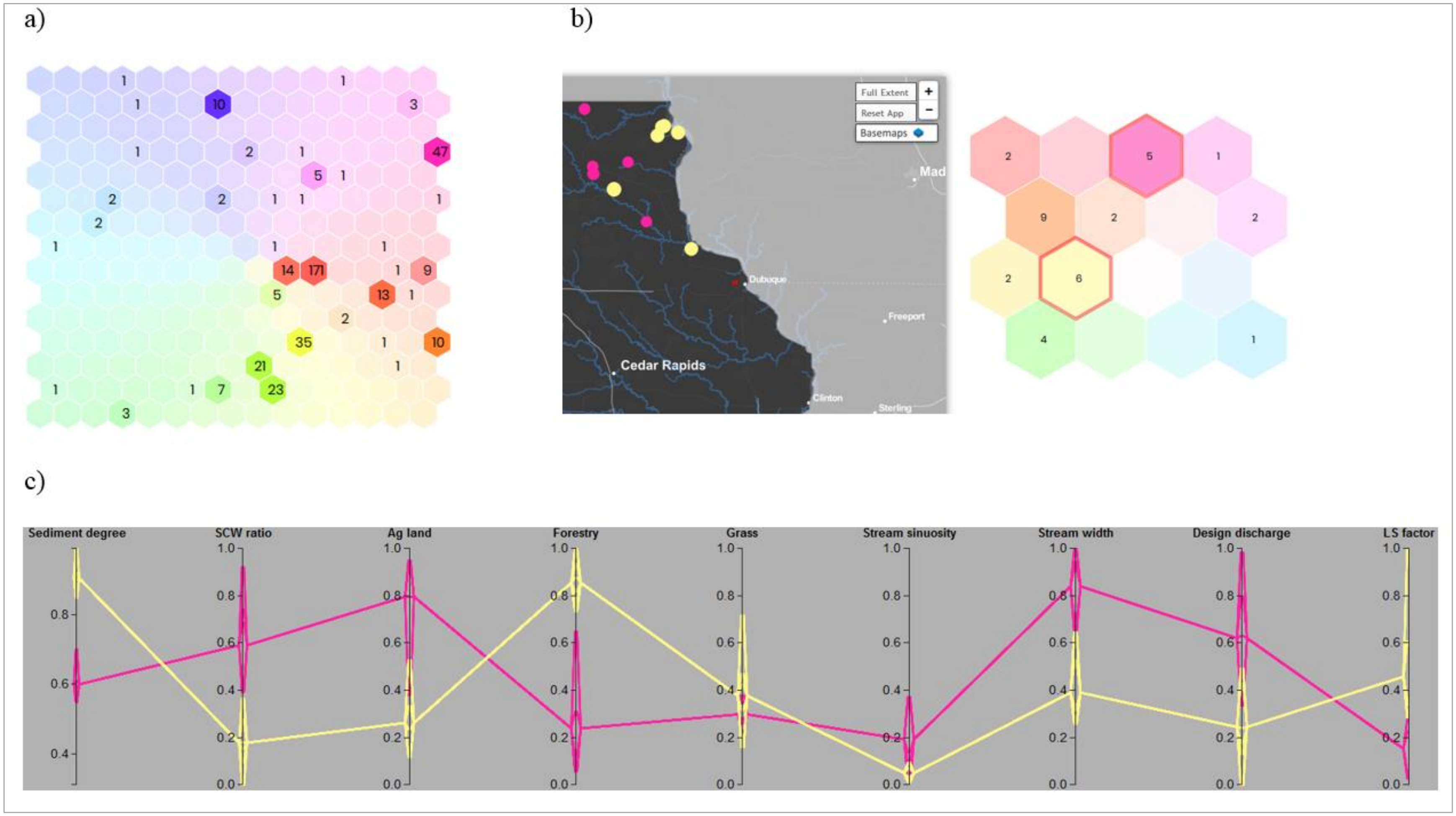


Figure 3.12 Visualizations of the SOM outputs: (a) hexagonal layout of SOM nodes, (b) spatial locations of culverts that are contained in a single cluster (node/hexagon), and (c) Edge-bundling PCP that displays the multivariate patterns and overall trends of each culvert clusters

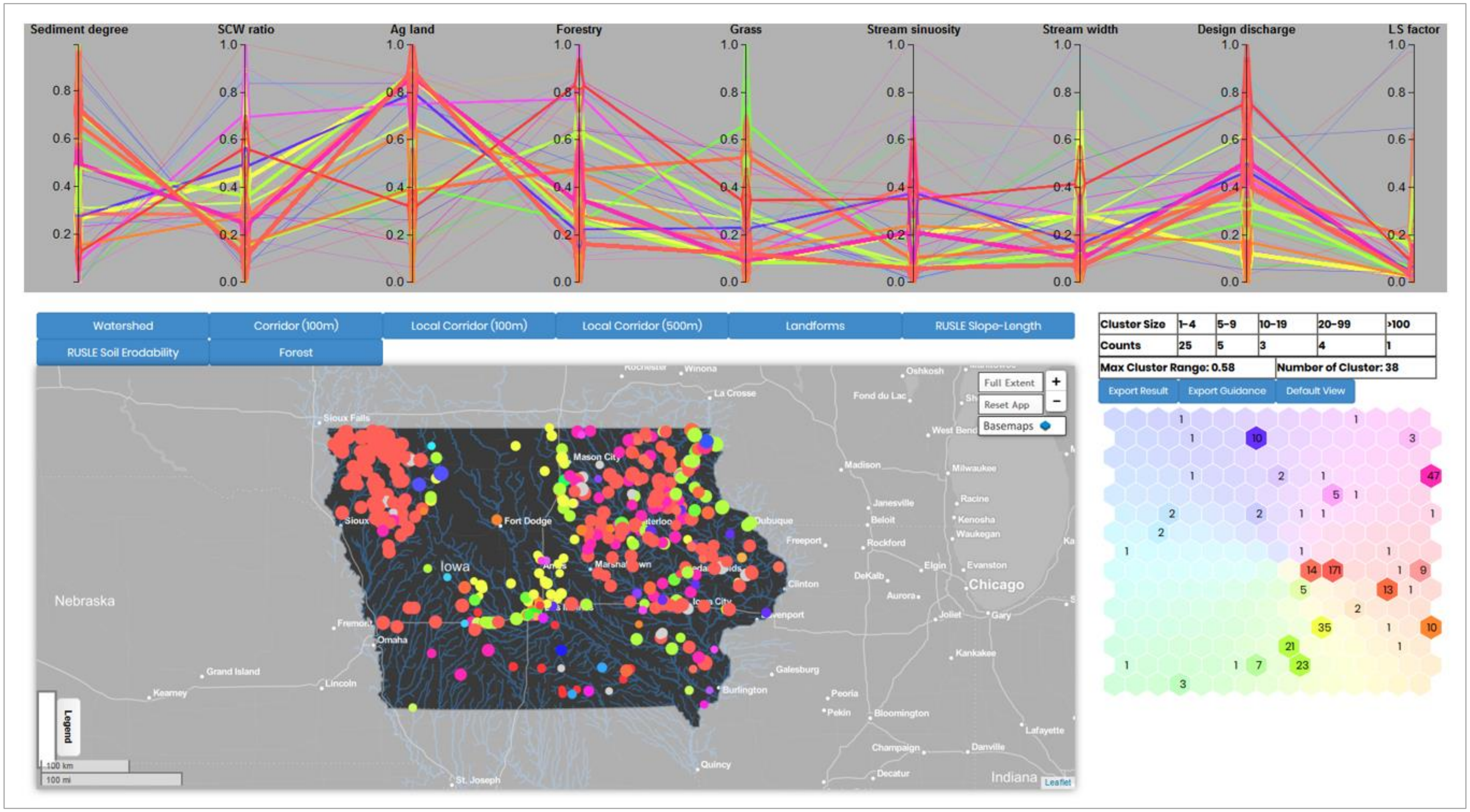


Figure 3.13 Combined visual interface with different visualization techniques, including SOM, edge-bundling PCP, and web map

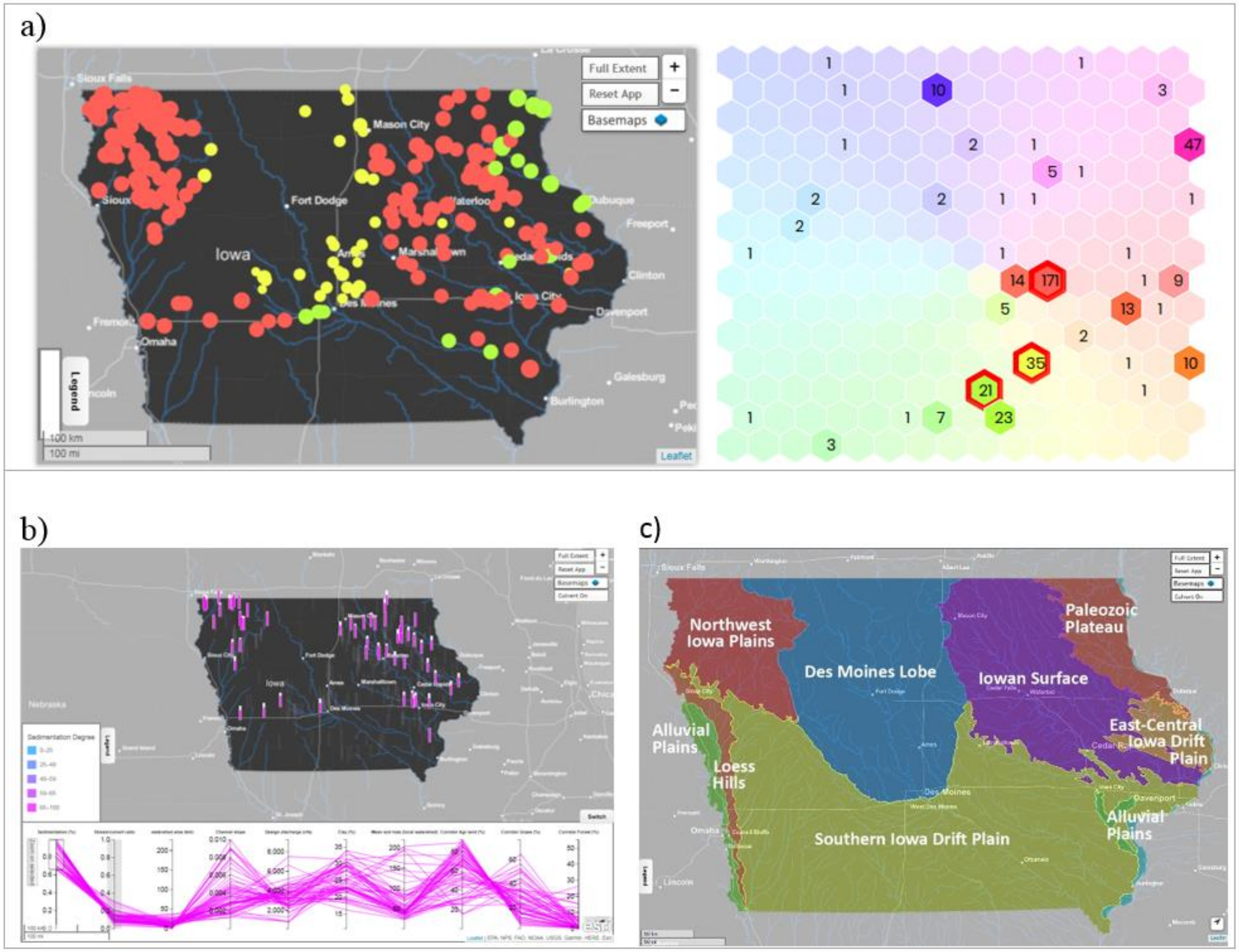


Figure 3.14 Regionalization interface: a) SOM clusters that are spatially correlated to the Northwest Iowa Plains and Iowan Surface region (orange dots), the Des Moines Lobe (yellow dots), and the Paleozoic Plateau (green dots), (b) regions of high culvert sedimentation potential determined through the MCDA, and (c) Landform regions of Iowa (Prior, 1991)

At the overview level, this study discovers that the spatial patterns of SOM clusters are strongly correlated with the landform regions covering the state. Past studies also indicate that landform regions are good reflections of the landscapes' erosional history, as well as other synthetic factors characterizing topography, soil erodibility, and soil mobility (Scheidegger, 1973; Prior, 1991). This discovery can also be validated through the previous MCDA, which allows users to easily identify regions with high culvert sedimentation potential by brushing the PCP as illustrated in Figure 3.14b. As the previous literature and the SOM outcomes concur in their findings, the study utilizes the geological landform regions to spatially partition the State of Iowa into six sub-regions (shown in Figure 3.14c). Each of these regions is considered homogeneous with respect to erosion potential and dominant process-drivers. Chapter 5 presents both the justification of the regionalization and the culvert sedimentation potential within each region.

3.5.4 Multivariate clustering analysis for culverts in each region

Despite the appeal and the advantages of estimating the similarity of erosion potential on the basis of landform regions, culverts within the same region are not guaranteed to have homogenous sedimentation degrees or be affected by same process-drivers. Due to natural complexity, cause-effect relationships between stream responses and their drivers (stressors) can be either divergent or convergent in a fluvial system (Schumm, 1991). Schumm further explains that this complexity is due to the fact that different processes (hence their associated drivers) can result in a certain response in some watersheds, while the same stressors can yield different responses in other similar size areas. Moreover, it has been long recognized that even if individual sedimentation process drivers may have a predictable effect on sedimentation response acting alone, when they are combined the resultant (multivariate) pattern of sedimentation response can be complex and unpredictable (Merritt et. al., 2013). For example, the effect of topography on sediment delivery patterns is quite predictable, with higher erosivity rates on steeper slopes. However, complications in the prediction of erosion on low slope topography occur due to the fact that topography is also related to land use choices (Rustomji and Prosser, 2001). Consequently, the balance between these two factors (depending on their impact ratio over the analyzed area) can lead to unexpected results, e.g., low slope lands are used for intensive agriculture (hence an enhancement of sedimentation) whereas these areas are producing typically low erosion rates due to the landscape slope.

These discrepancies can occur even after the regionalization process described above is implemented due to the fact that the degree of culvert sedimentation in some regions may depend on the combined effect of multiple drivers (natural and external). More clarity in forecasting can be brought in by deriving the multivariate patterns (relationships) between the culvert sedimentation and the drivers for the same region. In order to extract the multivariate patterns, this study applies the Self-Organizing Map (SOM) within each of the area defined after regionalization. The SOM, rather than other clustering algorithms, proved to be adequate for the present application due to the relatively small number of samples (e.g., 50-100) of the analyzed populations (Kiang et al., 2005). The outcomes of the SOM for multivariate clustering are visualized using the approach presented in Section 3.5.3 to enable human perception and judgment of the pattern through human-computer interaction. The results and data-driven insights derived from the analysis are presented in Chapter 5.

3.6 References

Allan, J. D. (2004). "Landscapes and riverscapes: the influence of land use on stream ecosystems". Annual Review of Ecology, Evolution, and Systematics. 35 (1): 257-284.

Andrienko. G., Andrienko, N., Jankowski, P., Keim, D., Kraak, M-J., MacEachren, A., and Wrobel, S. (2007). "Geovisual analytics for spatial decision support: setting the research agenda", International Journal of Geographical Information Science. 21(8): 839-857.

Auerbach, D. A., Buchanan, B. P., Alexiades, A. V., Anderson, E. P., Encalada, A. C., Larson, E. I., and Flecker, A. S. (2016). "Towards catchment classification in data-scarce regions", Ecohydrology. 9 (7): 1235-1247.

Belton, V. and Stewart, T. J. (2002). "Multiple criteria decision analysis: an integrated approach", Kluwer Academic Publishers, Boston, MA.

Bryce, S. A. and Clarke, S. E. (1996). "Landscape-level ecological regions: linking state-level ecoregion frameworks with stream habitat classifications", Environmental Management. 20: 297–311.

Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., et al. (2013). "API design for machine learning software: experiences from the scikit-learn project". arXiv Preprint arXiv:1309.0238.

Carlisle, D. M., Falcone, J., and Meador, M. R. (2009). "Predicting the biological condition of streams: use of geospatial indicators of natural and anthropogenic characteristics of watersheds". Environmental Monitoring and Assessment. 151 (1): 143-160. doi:10.1007/s10661-008-0256-z.

Chen, H., Wood, M. D., Linstead, C., and Maltby, E. (2011). "Uncertainty analysis in a GIS-based multi-criteria analysis tool for river catchment management", Environmental Modelling & Software. 26 (4): 395-405.

Crossman, N. D., Bryan, B. A., and King, D. (2011). “Contribution of site assessment toward prioritising investment in natural capital”, Environmental Modelling & Software. 26 (1): 30-37.

Demir, I and Szczepanek, R. (2017). “Optimization of river network representation data models for web-based systems”. Earth and Space Science. 4 (6): 336-347.

Easter, W. K., & Hufschmidt, M. M. (1985). “Integrated Watershed Management Research”. Honolulu, Hawaii: Environment and Policy Institute, EastWest.

Edsall, R. M. (2003). “An enhanced geographic information system for exploration of multivariate health statistics”, The Professional Geographer. 55 (2): 146-60.

Freeman, M. C., Pringle, C. M., and Jackson, C. R. (2007). “Hydrologic connectivity and the contribution of stream headwaters to ecological integrity at regional scales”. JAWRA Journal of the American Water Resources Association. 43 (1): 5-14. doi:10.1111/j.1752-1688.2007.00002.x.

Galelli, S. and Castelletti, A. (2013). “Tree-based iterative input variable selection for hydrological modeling”, Water Resources Research. 49 (7): 4295-4310. doi:10.1002/wrcr.20339.

Guo, D., Gahega, M., MacEachren A., and Zhou, B. (2005). “Multivariate analysis and geovisualization with an integrated geographic knowledge discovery approach”. Cartography and Geographic Information Science. 32 (2): 113–132.

Hajkowicz, S. and Higgins, A. (2008). “A comparison of multiple criteria analysis techniques for water resource management”, European Journal of Operational Research. 184: 255–265.

Hill, R. A., Weber, M. H., Leibowitz, S. G., Olsen, A. R., and Thornbrugh, D. J. (2016). "The Stream-Catchment (StreamCat) dataset: a database of watershed metrics for the Conterminous United States", Journal of the American Water Resources Association. 52(1): 120-128.

Inselberg, A. (1985). "The plane with parallel coordinates". The Visual Computer. 1 (2): 69-97.

Jelokhani-Niaraki, M., Sadeghi-Niaraki, A. and Choi, S-M. (2018). "Semantic interoperability of GIS and MCDA tools for environmental assessment and decision making", Environmental Modelling & Software. 100: 104-122.

Jones, S. (2013). "Selecting a culvert design option". Ames, IA: 67th Iowa County Engineers Conference.

Kampichler, C., Wieland, R., Calmé, S., Weissenberger, H. and Arriaga-Weiss, S. (2010). "Classification in conservation biology: A comparison of five machine-learning methods", Ecological Informatics. 5(6): 441-450.

Kiang, M. Y., Hu, M. Y., Fisher, D. M., and Chi, R. T. (2005). "The effect of sample size on the extended self-organizing map network for market segmentation", Proceedings of the 38th Annual Hawaii International Conference on System Sciences. 9, January 03 - 06, 2005.

Kohonen, T. (2001). "Self-organizing maps". Berlin, Germany, New York: Springer, 501.

Koua, E. L. and Kraak, M. J. (2004). "Geovisualization to support the exploration of large health and demographic survey data". International Journal of Health Geographics. 3 (13).

Koylu, C., 2014. Understanding geo-social network patterns: computation, visualization, and usability. Ph.D thesis, University of South Carolina, Columbia, SC, USA.

Labiosa, W. B., Forney, W. M., Esnard, A. M., Mitsova-Boneva, D., Bernknopf, R., and Hearn, P., et al. (2013). “An integrated multi-criteria scenario evaluation web tool for participatory land-use planning in urbanized areas: the Ecosystem Portfolio Model”, Environmental Modelling & Software. 41 (2013): 210-222.

Lins, H. F. (1985). “Streamflow variability in the united states: 1931–78”, Journal of Climate and Applied Meteorology. 24 (5), 463-471. doi:10.1175/1520-0450(1985)024<0463:SVITUS>2.0.

Liu, H. and Motoda, H. (1998). “Feature transformation and subset selection”, IEEE Expert. 13(2): 26-28.

Loveland, T. R. and Merchant, J. M. (2004). “Ecoregions and Eco regionalization: geographical and ecological perspectives”. Environmental Management. 34: 1–13.

McKay, L., Bondelid, T., Dewald, T., Johnston, C., Moore, R., and Rea, A. (2012). “NHDPlus version 2: user guide”. Retrieved on November 7, 2017 from: ftp://ftp.horizonsystems.com/NHDPlus/NHDPlusV21/Documentation/NHDPlus-V2_User_Guide.pdf.

Merritt, W. S., Letcher, R. A., and Jakeman, A. J. (2003). “A review of erosion and sediment transport models”. Environmental Modelling & Software, 18 (8), 761-799. doi:https://doi.org/10.1016/S1364-8152(03)00078-1.

MNDNR. (2016). “Watershed Health Assessment Framework (WHAF)”. Retrieved from St. Paul, MN: http://www.dnr.state.mn.us/aboutdnr/mission.html.

Muste, M. and Xu, H. (2017). “Sedimentation mitigation using streamlined culvert geometry”, Report ST-001, Iowa Department of Transportation, Statewide Transportation Innovation Council, Federal Highway Administration, McLean, VA.

Muste, M. and Xu, H. (2017). "Mitigation of sedimentation at multi-box culverts", IIHR Report No. TR-655, Submitted to the Iowa Highway Research Board, Ames, IA, USA.

NHDPlus (2006). "NHDPlus Applications", Horizon systems - NHDPlus website. Retrieved on November 7, 2017 from http://www.horizon-systems.com/NHDPlus/applications.php.

Olden, J. D., Kennard, M. J., and Pusey, B. J. (2012). "A framework for hydrologic classification with a review of methodologies and applications in ecohydrology", Ecohydrology. 5 (4): 503-518.

Olden, J. D., Lawler, J. J., and Poff, N. L. (2008). "Machine learning methods without tears: a primer for ecologists", The Quarterly Review of Biology, 83 (2): 171-193. doi:10.1086/587826.

Palmas, G., Bachynskyi, M., Oulasvirta, A., Seidel, H. P., and Weinkauf, T. (2014). "An edge-bundling layout for interactive parallel coordinates". In IEEE Pacific Visualization Symposium (PacificVis): 57-64. Yokohama, JP.

Prior, J. (1991). "Landforms of Iowa": University of Iowa Press, Iowa City, IA, USA.

Prosser, I. Y. WJ., Rustomji, P., Hughes, A., and Moran, C. (2001). "A model of river sediment budgets as an element of river health assessment". Paper presented at the Proceedings of MODSIM 2001 - International Congress on Modelling and Simulation, Canberra, Australia.

Rosgen, D. L., and Silvey, H. L. (1996). "Applied River Morphology" (2nd edition ed.): Wildland Hydrology.

Sadeghi-Niaraki, A., Varshosaz, M., Kim, K., and Jung, J.J. (2011). "Real world representation of a road network for route planning in GIS", Expert Syst. Appl. 38 (10): 11999-12008.

Scheidegger, A. E. (1973), "Hydrogeomorphology", Journal of Hydrology. 20 (3): 193-215.

Schnier, S. T. (2016). “Data-driven analyses of watersheds as coupled human-nature systems”. Ph.D. Thesis, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL.

Schnier, S. T. and Cai, X. (2014). “Prediction of regional streamflow frequency using model tree ensembles”. Journal of Hydrology, 517(Supplement C): 298-309. doi:https://doi.org/10.1016/j.jhydrol.2014.05.029

Schumm, S.A. (1991), “To interpret the earth: ten ways to be wrong”: Cambridge, Cambridge University Press: 180.

Skupin, A. and Fabrikant, S. (2003). “Spatialization methods: A cartographic research agenda for non-geographic information visualization”, Cartography and Geographic Information Science. 30 (2): 99-119.

Tockner, K and Stanford, J. A. (2002). “Riverine flood plains: present state and future trends”. Environmental Conservation. 29 (3): 308-330.

USGS (2001). “National hydrography dataset (NHD)”. U.S. Geological Survey, Reston, VA.

USGS (2006). “National Elevation Dataset (NED)”. Retrieved from U.S. Geological Survey: http://ned.usgs.gov.

USGS and USDA (2013). “Federal standards and procedures for the national watershed boundary dataset (WBD) (Fourth Edition)”. U.S. Geological Survey Techniques and Methods 11-A3, 63, Reston, VA.

Vulevic, T. and Dragovic, N. (2017). “Multi-criteria decision analysis for sub-watersheds ranking via the PROMETHEE method”, International Research and Training Center on Erosion and Sedimentation and China Water and Power Press, International Soil and Water Conservation Research. 5(1): 50-55.

Wagener, T., Sivapalan, M., Troch, P. A., and Ross, W. A. (2007). "Catchment classification and hydrologic similarity", Geography Compass. 1 (4), 901–931.

Ward, J. V. (1989). "The four-dimensional nature of lotic ecosystems". Journal of the North American Benthological Society, 8 (1): 2-8. doi:10.2307/1467397.

Ward, J. V. (1989a). "Riverine ecosystems: the influence of man on catchment dynamics and fish ecology". Paper presented at the International Large River Symposium.

Ward, J., and Stanford, J. (1989a). "Four-dimensional nature of lotic ecosystems". Journal of the North American Benthological, 2 (8).

Ward, J., and Stanford, J. (1989b). "Riverine ecosystems: the influence of man on catchment dynamics and fish ecology". In: Proceedings of the International Large River Symposium, D.P.Dodge (Editor).Canadian Special Publication in Fisheries and Aquatic Sciences. 106: 56-64.

Wohl, E., Bledsoe, B. P, Jacobson, R. B., Poff, N. L., Rathburn, S. L., Walters, D. M., and Wilcox, A. C. (2015). "The natural sediment regime in rivers: broadening the foundation for ecosystem management". BioScience, 65 (4): 358–371.

Yadav, M., Wagener, T., and Gupta, H. (2007). "Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins", Advances in Water Resources, 30 (8), 1756-1774.

Zhang, Z., Wagener, T., Reed, P., and Bhushan, R. (2008). "Reducing uncertainty in predictions in ungauged basins by combining hydrologic indices regionalization and multiobjective optimization", Water Resour. Res. 44 (12).

CHAPTER 4 DEVELOPMENT OF CULVERT PLATFORM CYBERINFRASTRUCTURE

4.1 Overview

Modern information and communication technologies have become the fourth pillar of scientific investigations, complementing the capabilities of the traditional pillars (observation, theory, and analysis). Contemporary operational science and management of watershed resources are increasingly supported by data- and information-rich environments. These environments are poised for progression, not only because of the recent advances in science and engineering research but even more prominently by advancements in “hydroinformatics”. This emerging discipline is defined as the science of information handling for solving water-related problems (Abbott, 1991). Today’s hydroinformatics capabilities to integrate data with models and artificial intelligence enable the creation of high-fidelity, first-principle based, numerical surrogates of real systems that aid quantifiable understanding of the critical processes that characterize the water cycle in watersheds.

The backbone of hydroinformatics is the cyberinfrastructure technology. According to Wikipedia, “in scientific usage, cyberinfrastructure is a technological and sociological solution to the problem, which efficiently connects laboratories, data, computers, and people with the goal of enabling derivation of novel scientific theories and knowledge”. Contemporary cyberinfrastructure has become increasingly available, and sufficiently mature, so as to facilitate the development of digital platforms for supporting both scientific investigations and the management of various problems at the watershed scale (Demir and Beck 2009; Demir et al., 2009; Demir et al., 2015; Weber et al., 2018). These platforms have the potential to transform our capabilities of understanding how to address ecosystem changes, protect the environment, and predict and prevent natural and human disasters through knowledge-based adaptive management.

Many contents in this chapter are authorized replications of the IHRB TR-655 report (Muste and Xu, 2017), authored by my advisor Dr. Marian Muste and myself, as the funding for this Ph.D. study was provided by the Iowa Highway Research Board, and Iowa Department of Transportation (Iowa DOT), Grant TR-655.

4.2 Iowa DOT culverts platform architecture, software, and technologies

The "IowaDOT Culverts" geo-portal (Xu et al., 2018) comprises a front-end and a back-end component (the term "front-end" is referring to user's computer, while "back-end" indicates the server backing the web applications; the front-end is also known as the client-side, while the back-end is also referred as the server-side section of a web platform). As is common in most modern web-application templates, the "IowaDOT Culverts" platform adopts a three-tier architecture that includes the following components: (1) presentation, (2) logic, and (3) data. To ensure platform reliability, flexibility, extendibility, modularity, and maintainability industrial design and architecture patterns (e.g. model–view–controller - MVC) were applied in the system development. Figure 4.1 illustrates the overall architecture, along with the web, informatics, and GIS technologies that are associated with each tier.

The presentation tier is primarily rendered at the front-end in a user's web or mobile browser. It contains platform elements that a user can see and interact with. This tier provides users with Graphic User Interfaces (GUI), a map engine, and diverse visualization tools to facilitate map operations, information retrieval, workflow control, watershed planning, and communication. The presentation tier in the "IowaDOT Culverts" platform entails four components: (1) the map engine, (2) the GUI, (3) logic management, and (4) the visualization tools. The map engine is the means to visualize geo-spatial information, such as base maps, river networks, watershed boundaries, locations of the culverts, and modeling results (e.g., RUSLE). The presentation tier is developed with Leaflet JavaScript (JS) library and its extensions. The GUI provides a media for users to navigate through the platform, manage and control tools, and retrieve information. The GUI is developed using JQuery and Bootstrap JS library, which guarantees both user interactivity and compatibility for multi-screen sizes.

The logic management component contains a front-end MVC that improves fluid web page design and two-way data-binding. The main reason to have a logic management component is that our platform is designed as a Single-Page Application (SPA), which make the front-end very heavy. The front-end MVC, a JavaScript library itself, helps structure and optimize the front-end developments with practical industrial conventions, which increases the maintainability and extendibility at the front-end. Visualization tools are primarily responsible for visual communication and representations (e.g. plots, chart). They are developed with D3.js and HighChart libraries, both of which are data-driven and user-responsive. The entire presentation tier is developed using common front-end technologies (e.g. JavaScript, HTML, and CSS). To perform multiple system operations (e.g. updating data & information, displaying spatial features on a map, user log-in, saving user-defined watershed plans), the presentation tier sends Asynchronous JavaScript and XML (AJAX) requests to exchange information with the server-side applications in the form of JSON, XML, and images. A summary of the software components in the "IowaDOT Culverts" platform is provided in Table 4.1.

Unlike the presentation tier, the logic tier and data tier are deployed on the server-side (i.e., "back-end" of the platform). The logic tier is responsible for organizing the data, assembling the services based on the relationship between the user scenarios and the models, and for providing the necessary information requested by the presentation tier. The logic tier consists of three sub-components: (1) the map server and map services, (2) the application framework, and (3) the web services for real-time sensors. The map server and web services prepare and manage spatial information, as well as handle requests from the presentation tier for the map visualization. The "IowaDOT Culverts" uses GeoServer (GeoServer, 2014), an open-source map server application for hosting spatial information stored locally on the server (e.g. river, watershed boundaries). The GeoServer complies with a number of open standards, such as Web Feature Service (WFS), Web Map Service (WMS), and Web Coverage Service (WCS), which improve the interoperability of spatial data. Third-party map services from Google and ESRI are also used to increase the diversity of the base maps (e.g. satellite imagery, topo-maps, and NHD base map) within the platform. The application framework components manage the overall back-end logic (e.g. models, culvert design data, and data integration) and user-scenarios.

Many of the platform's tools and applications (e.g. the watershed search engine) are hosted in the application module. This module is responsible for managing local web services. The system design adopts a Service Oriented Architecture (SOA) to bring multiple web services together in one place. There are two types of web services in the Iowa DOT platform: (1) local web services (that are developed within the application framework on the local server), and (2) external web services (that are hosted on third-party servers). External web services in "IowaDOT Culverts" are mainly third-party data providers (e.g. USGS, IFC). The web services are important components for the presentation tier as they facilitate the communication between the presentation and the logic tier. The backbone of the application framework module is Yii (a PHP framework that also follows the MVC pattern). The data tier is located at the bottom of the architecture and consists of databases and datasets. The spatial data are stored in the PostgreSQL database powered by PostGIS extension, which offers additional support for the use and management of spatial objects.

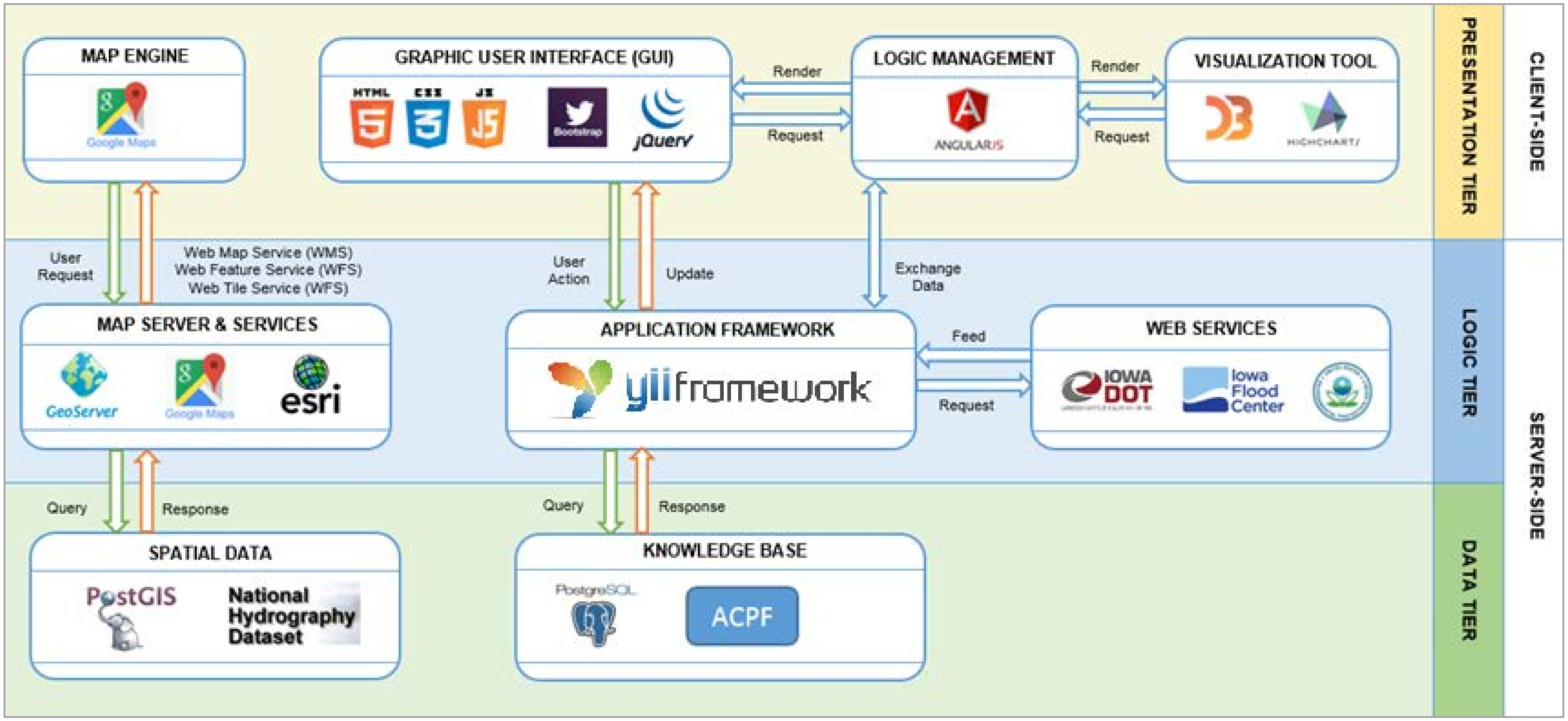


Figure 4.1 Architecture of the “IowaDOT Culverts” portal (Xu et al., 2015)

Table 4.1 Summary of software components, their role, and associated programming languages

Components	Role	Technology
Back-end		
Apache HTTP Server	Web Server	-
Apache Tomcat	Web Server	-
Geoserver	GIS Map Server	Java
PostgreSQL & PostGIS	RDBMS	SQL
PHP Yii framework	Backend MVC	PHP
Front-end		
jQuery	Cross-platform JS library	Javascript
jQuery-UI	jQuery UI design library	Javascript
Bootstrap	UI design library	Javascript
Bootstrap	Bootstrap switch	Javascript
Leaflet	Map engine	Javascript
Esri Leaflet	Leaflet plugin	Javascript
Leaflet Google	Leaflet plugin	Javascript
Leaflet Draw	Leaflet plugin	Javascript
Highcharts	Plots	Javascript
Highcharts-more	Plots-extension	Javascript
D3.js	Data visualization	Javascript
D3.js-Leaflet	Leaflet plugin	Javascript
Leaflet-OpenWeatherMap	Leaflet plugin	Javascript
Google Map API	Map engine	Javascript
typeahead.js	Text autocomplete	Javascript
Bing Map API	Map engine	Javascript
Leaflet Geolocator	Leaflet plugin	Javascript
Geolocator Control	Leaflet plugin	Javascript
Leaflet Routing Machine	Leaflet plugin	Javascript
Lagodiuk Decision Tree	Javascript machine-learning	Javascript
UI Design Double Scroll	jQuery UI design	Javascript
Three.js	3D visualization	Javascript
	Data Sources	
Iowa Flood Center	Stage sensor	PHP web-service
USGS SSURGO	Web map service	Rasters
RUSLE	Web map service	Rasters
High Resolution Land Cover	Web map service	Rasters
SIIMS	Culvert information	CSV
EPA StreamCAT	Watershed characterization	CSV/Web services

The platform uses two (front-end) mapping engines: Leaflet and Google Map API. The Leaflet is used as the primary map engine to display base maps and most of the two-dimensional features. Google Map API is used to display 45-degree bird-eye views (i.e., oblique imagery), considered a 2.5-dimensional representation for a map. The Leaflet is a popular open-source, mobile-friendly JavaScript map engine library designed for interactive web mapping. Compared to other modern map engines (e.g., OpenLayers, ArcGIS JavaScript API, Google Map), Leaflet is a light-weight and flexible map engine. Leaflet supports WFS, WMS, and WMTS produced in any geographic coordinate system, geographic features or maps in geographic projection system through the usage of customized extensions. The fundamental Leaflet provides a basic method to display a base map, overlay, and data visualizations in browsers, while its diverse extensions provide more advanced features in web mapping.

The front-end consists of a GUI and an interactive map created using HTML, CSS, and JavaScript (MDN, 2015). The GUIs are created for function implementation, workflow navigation, and platform control purposes. In other words, through the interface, the user is able to find and select specific functions or tools for certain data or analytical demands. The GUIs also create classifications for different functionalities and workflows. The platform's GUIs are rendered using HTML and CSS languages. JQuery (jQuery Foundation, 2010) and generic JavaScript language are attached to the webpage to create dynamic, interactive effects for the users. An interactive map is built to visualize geographic information and spatial data and is composed of web map engines and web map services. The platform uses Leaflet (Agafonkin, 2010), Google Map API, and various Leaflet extensions as map engines.

The back-end of the geo-portal holds the database, the GIS web server software, and the server software. The database stores local spatial data, tabular data, and model results for the platform.

Other map data are obtained with web services connected to the IDOT database and other public digital repositories. The platform uses the PostgreSQL database with PostGIS extension as the main database, which grants original Postgres with spatial capabilities. Middleware programs are used to connect the database with the front-end mapping engines and to distribute data for different applications. The platform also adopts subject-oriented data structure, which optimizes data downloading and browser caching to save system loading time. This structure is built with middleware programs. GIS web server software is used to manage and distribute spatial data.

4.3 Cyberinfrastructure for data integration

The Iowa DOT architecture uses the Digital Watershed (DW) concept as a spatial unit for assembling multi-scale, multi-domain data acquired in-situ or resulting from models produced by various agencies (Muste, 2014). To increase the data accessibility, a selection tool is therefore developed to offer multiple choices for the search (e.g., HUC or administrative keywords) and subsequently triggers visualization of watersheds associated with the selected culverts. Using the spatial boundary of the delineated drainage area, the portal dynamically generates a DW of the area entailing the watershed-related data and information stored in the system's database. Organization of the data in DW is made by indexing the watershed-related data with an identifier that relates the data with the watershed.

The backbone of the watershed characterization module is the NHDPlus dataset, which provides physical (e.g. shapefiles of rivers, catchments, and watershed boundaries) and topological (e.g. river connectivity, hierarchy of tributaries, hydrologic unit code system) descriptions of the real-world hydrologic system in a digital environment. The physical description includes features such as streams, catchments, and Hydrologic Unit Code (HUC) watersheds, which are either displayed on the map as boundaries or spatially indexed with the watershed-related data resources (e.g. raster data, point-observation, and modeling results) associated with them. Results of the previous modeling (e.g., RUSLE soil loss and associated parameters) or syntheses (e.g., StreamCat) and structural information associated with culverts are also spatially indexed. The watershed-related data and information can be easily discovered and retrieved for each indexed watershed (e.g. by structure number, or river, or road name, or HUC code). The NHDPlus covers hydrologic and hydrographic features at a national level, and it is widely used among research groups and watershed management communities in the U.S. The dataset is stored using relational formats and can be easily extended in terms of spatial coverage and new attributes. The overall structure of the data integration and their sources is provided in Figure 4.2.

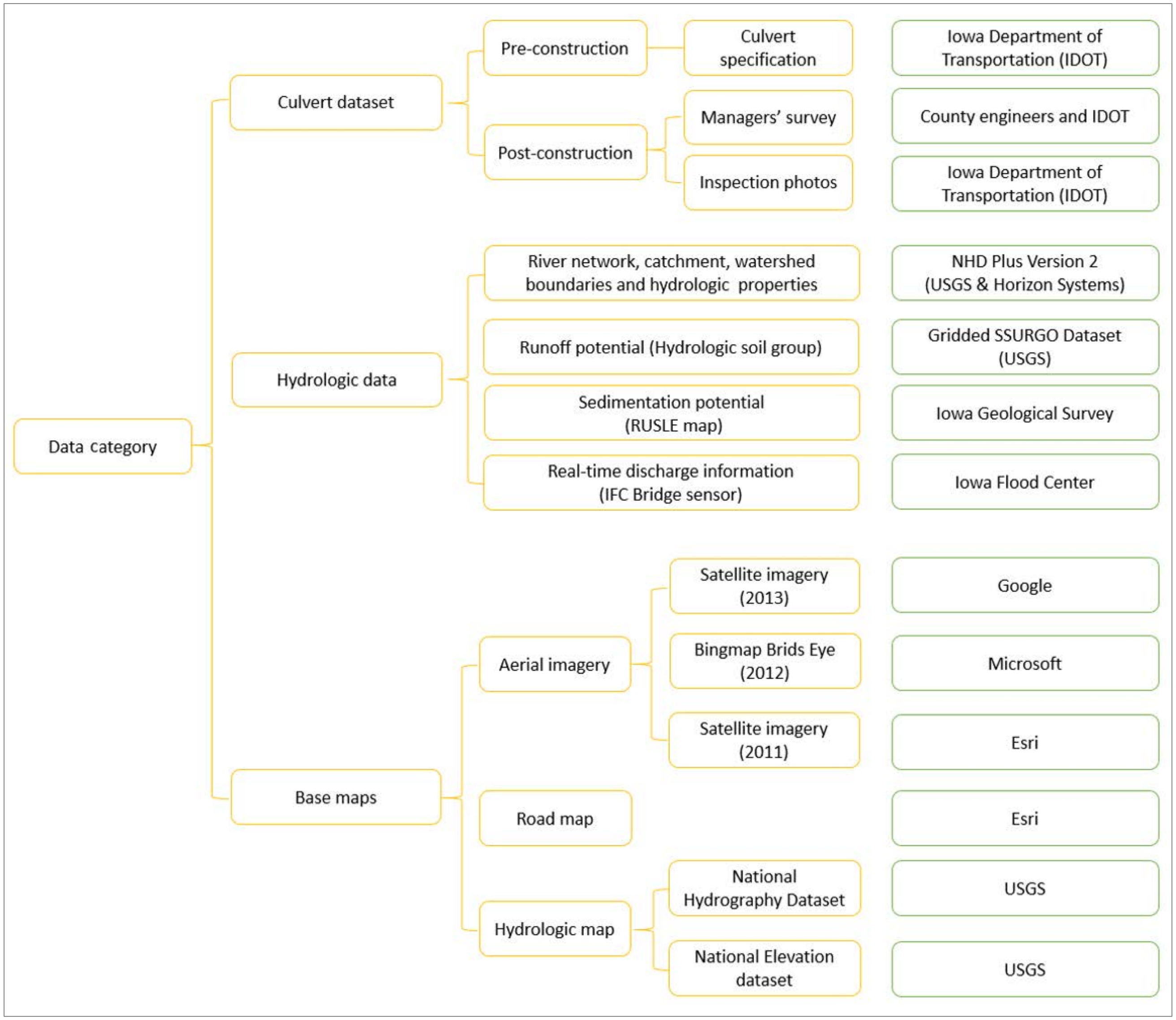


Figure 4.2 Structure of the data contained in the IOWADOT Culverts platform

The Leaflet is used for 2D map visualization while the Google Map API contains 45 degrees bird's-eye views (only available in Urban areas), which is used for presenting the surrounding environment of a Point of Interest (POI) from four different perspectives (directions). Both the mapping engines visualize the map using various types of map services from map servers, which require the use of different Leaflet Extensions, including Leaflet Projection Extension, Leaflet-Google Map API, ESRI-Leaflet API and other supporting upgrades (Agafonkin, 2010). The purpose of Leaflet Projection is to enable the Leaflet mapping engine to display web maps in miscellaneous projection systems, which are used by various web GIS servers. Leaflet-Google Map API and ESRI-Leaflet (ESRI, 2015) API are used to improve spatial data interoperability and compatibility with other GIS software and web services. Map services in the platform are materialized through Web Map Service (WMS) (OGC, 2009), Web Map Tile Service (WMTS) (OGC, 2011a), and Web Feature Services (WFS) (OGC, 2011b). The platform provides a variety of base maps including aerial imagery, street maps, topographic maps, and hydrographic maps. Aerial imageries acquired from a different time period (provided by the Iowa Geographic Map Server) to create a time series for monitoring environmental, topographical, and infrastructural changes.

Besides the base maps, the platform allows users to visualize the culvert location, river cross-sections, county and watershed boundaries, and sensor locations. All these data are converted to GeoJson files stored in the servers' directory system. The reason to use GeoJson files, and not WMS, is that vector-based geographic objects are much more interactive and modifiable than image-based WMS, allowing users to change their symbology and style (e.g., color and size) dynamically. The platform also utilizes web GIS server (map server) software to store and execute online geo-processing services. Those online geo-processing services enable the platform to calculate areas and length of dynamically created geographic components (e.g., culvert sedimentation mapping), and provide routing service (e.g., geocoding, navigation to culverts, and creating survey routes between different addresses). The OGC GeoServer and ArcGIS servers are the two GIS web server software used in the platform. For optimization of the operations on the platform, the functions provided through the ArcGIS server was replaced with custom created server-side scripts hosted on an IIHR server.

4.4 Workflows

The platform functionalities are grouped under four workflows with tools and engines serving different aspect of the management and design of the culvert. The four workflows are: General Info, Monitoring, Sedimentation Analysis, and Design Aids.

4.4.1. "General info" workflows

The information about the culverts is retrieved through two search engines. Each of the search engines delineates the boundary of the drainage area upstream of any selected culvert across Iowa, and can retrieve the associated drainage area characteristics. With this arrangement, drainage area characteristics for existing and future culvert locations can be conveniently retrieved. The dataset for existing culverts contains pre-construction and post-construction culvert data. Pre-construction data include the locations and design specifications of the 723 multi-box culverts (mainly 3 & 4-box culverts maintained by Iowa DOT). Culvert locations are provided in geographic coordinates and converted into a shape file with joint specification as attributes. Culvert specifications contain elementary culvert design data, including culvert structure numbers, owner, type information, state code, and geometry information. All of the pre-construction data are stored in a PostgreSQL database. The post-construction data contain information acquired during routine trained inspections or culvert surveys that are conducted following the occurrence of severe storm events. Maintenance information includes written comments and photographs of the culvert. This data is stored for some of the Iowa culverts in the IDOT SIIMS database. These maintenance records also contain IDOT engineer recommendations regarding the asset status. Figure 4.3 reproduces typical information stored in the IDOT SIIMS database.

The culvert database is connected to the GeoServer software using the FHWA structure number so that data requested by the platform can be transferred to the client-side as a web feature service and displayed on the map. As the caching of the entire pre-construction data at the client-side is slow, the data caching is made using a subject-oriented data structure. The location of the culverts is the only cached data during the initial loading of the platform. While users request specific functions or applications, the related culvert specifications are subsequently loaded. The culvert search workflow is controlled through friendly graphical interfaces that accommodate the needs of engineers with a minimum computer programming background. The culvert search workflow combines the culvert data repository and regional hydrologic analysis with culvert-related environmental attributes to compensate for the lack of information on sedimentation in current design specifications. The culvert monitoring workflow allows managing personnel to query comprehensive culvert information and inspect the physical status of culverts through high-quality aerial imagery taken over time. The flow diagram for culvert search and monitoring is provided in Figure 4.4. The results of the post-construction surveys conducted by the IDOT personnel are recorded under the “Managers’ Survey” in Figure 4.4.

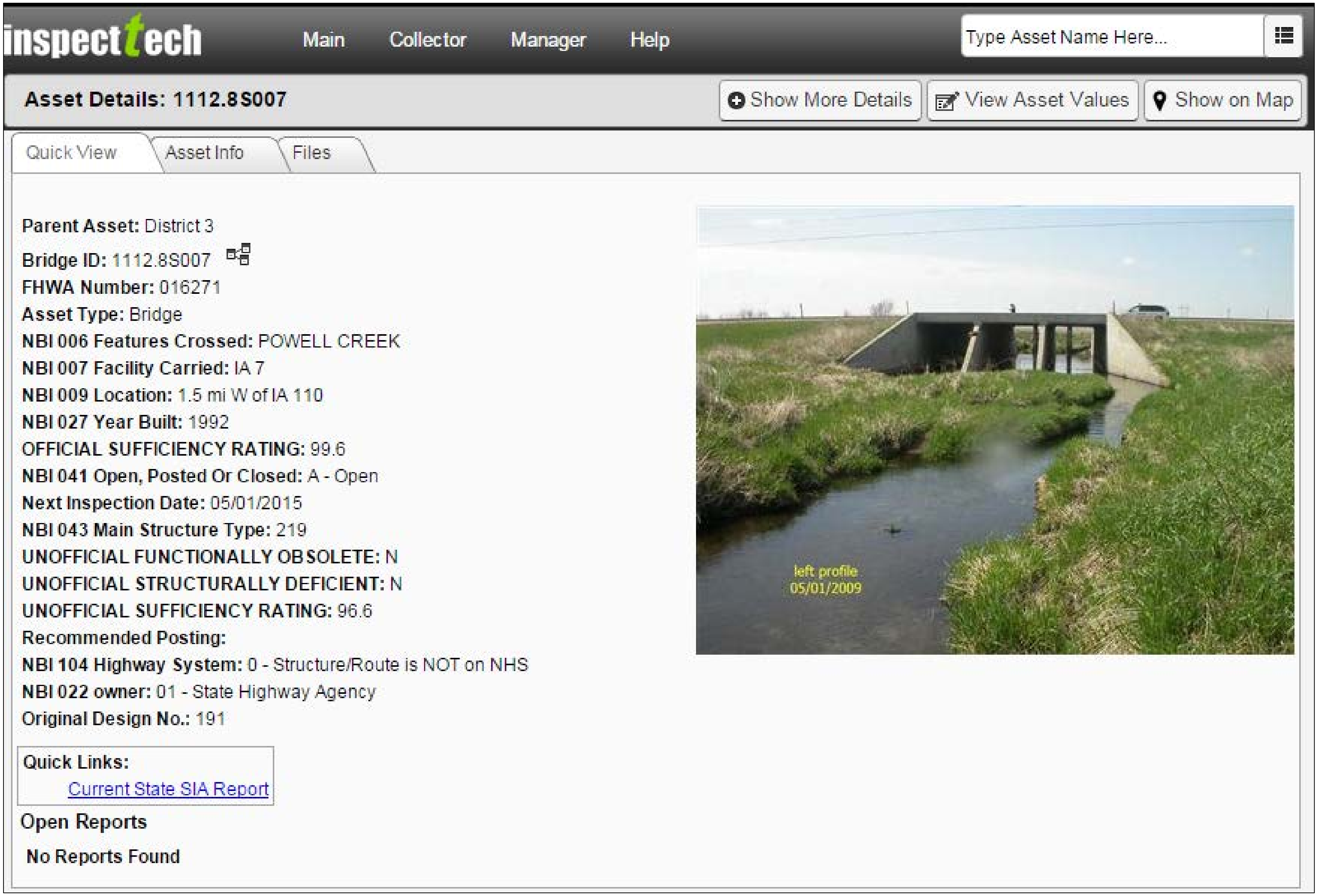


Figure 4.3 Maintenance record in the IDOT SIIMS database (SIIMS, 2015)

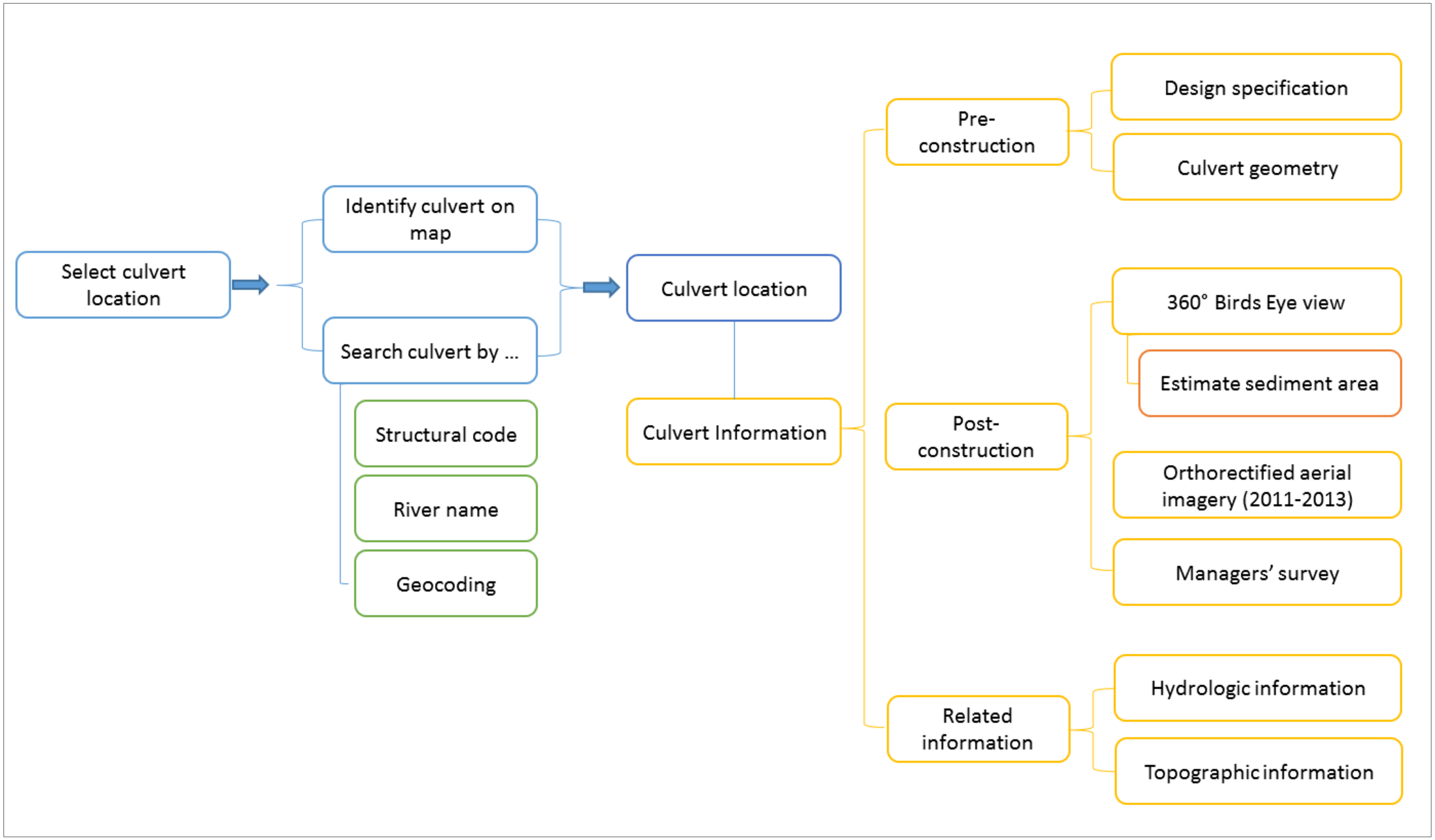


Figure 4.4 Flow diagram for the "General info" workflow

4.4.2. "Monitoring" workflows

This cluster of tools is assembled with the intent to aid culvert monitoring activities. Specifically, the workflow assists users to inspect a culvert site and observe time series and details of the surrounding terrain provided from LiDAR maps. The flow of information for this workflow is illustrated in Figure 4.5. Useful tools for navigating to sites and organizing field campaigns that cover specific purposes are also embedded in this workflow. Figure 4.6a shows itineraries created for this study to inspect the 257 culverts in the state. The "IowaDOT Culverts" platform offers a choice of background base maps, including high-resolution topographic maps derived from national elevation datasets, watershed and river information from national hydrographic datasets, and 1-meter LiDAR hill shaded maps, which provide the user with a 2.5D view of the culvert geometry and surrounding terrain. This variety of maps allows users to view various aspects of culvert sedimentation detail and its evolution. Shortly after the surveys are finalized, the monitoring results can be viewed immediately if the field data collection device (e.g., smartphone and iPad) are connected with to the platform through the internet. To view the information, the culvert search tool is used, first to access the culvert using the site metadata (structural code, geocoding and the river on which the culvert is located). After locating the culvert of interest, the system automatically displays two inspection windows for the selected site. The first window contains basic information about culvert design and geometry, as illustrated in Figure 4.6b.

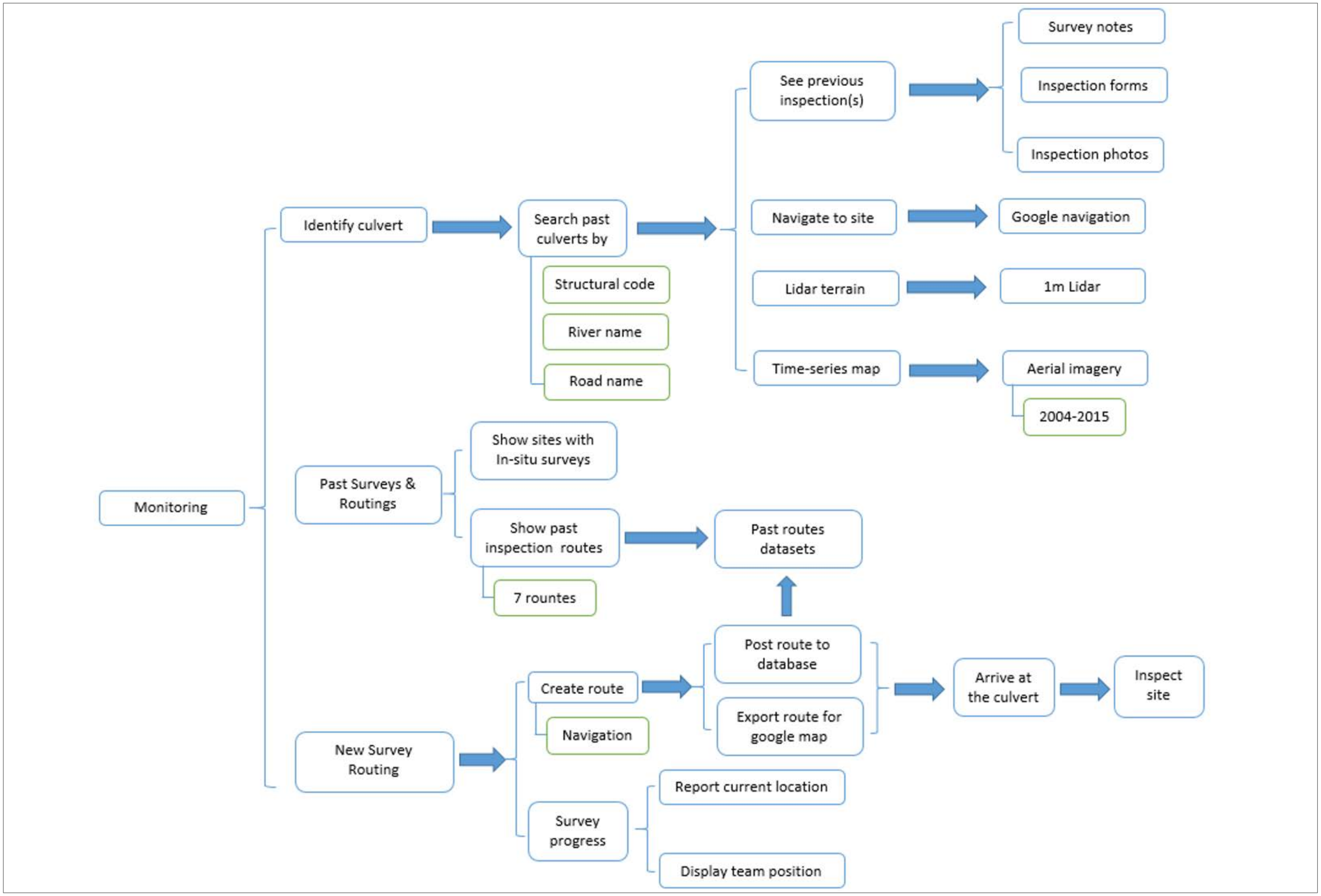


Figure 4.5 Flowchart for the "Monitoring" workflow

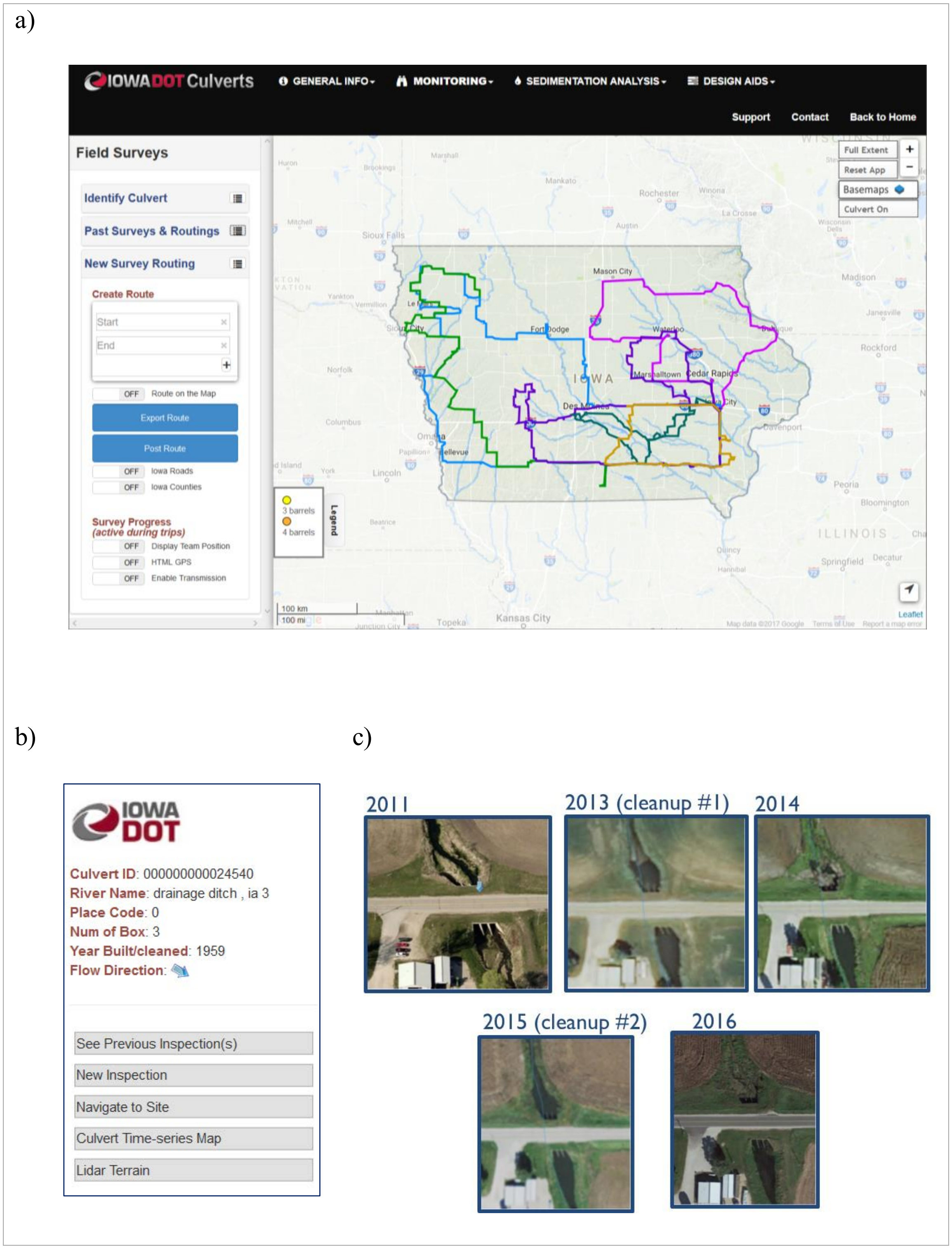


Figure 4.6 Functions associated with the "Monitoring" workflow: a) engine for aiding navigation at multiple sites in the "monitoring" workflow; and b) time series for a culvert with recurrent sedimentation

Another useful feature that takes advantage of the availability of ortho-rectified aerial imagery from 2004 to 2013 stored in the platform is the culvert sedimentation time series. For example, the time evolution of the sediment deposits can be quite reliably tracked with a customized map viewer that covers the same area in the vicinity of the culvert to make the analysis efficient. A sample time series is provided in Figure 4.6c. The availability of the aerial images (dating back to the 1940s) allows users to make abundant inferences about the sedimentation process as a whole as the retrieved images also document morphological changes in the configuration of the stream in the culvert vicinity.

4.4.3 "Sedimentation analysis" workflows

This workflow entails the central group of cyber tools in the platform that is focused solely on the quantification and evaluation of the sedimentation as an end-to-end process, i.e., the "Sedimentation Deposit Mapping" and "Multi-Criteria Decision Analysis" workflows, respectively. The first workflow contains several mapping alternatives for the estimation of the degree of sedimentation, which are subsequently used in the MCDA phase. The quantification of culvert sedimentation status is made directly on the aerial photograph contained in the display window, illustrated in Figure 4.7. Using geo-referencing techniques and geo-processing services developed specifically for this purpose, the user can delineate, using a polyline, the boundary of sediment deposit at the culvert. The map engine estimates the geo-coordinates of the polyline vertices and imports the results into a geo-processing service for area calculation. If required, the geo-processing service is capable of roughly estimating the volume of the sediment in the mapped deposit using area and elevation information available in the digital elevation model. A similar geo-processing tool is available to estimate the stream width that is subsequently used to define the critical process parameter of stream-to-culvert width ratio (see Figure 2.7). The workflow allows users to inspect the archive of sediment deposit maps, redo the mapping, or create a new sedimentation mapping for the culvert site.

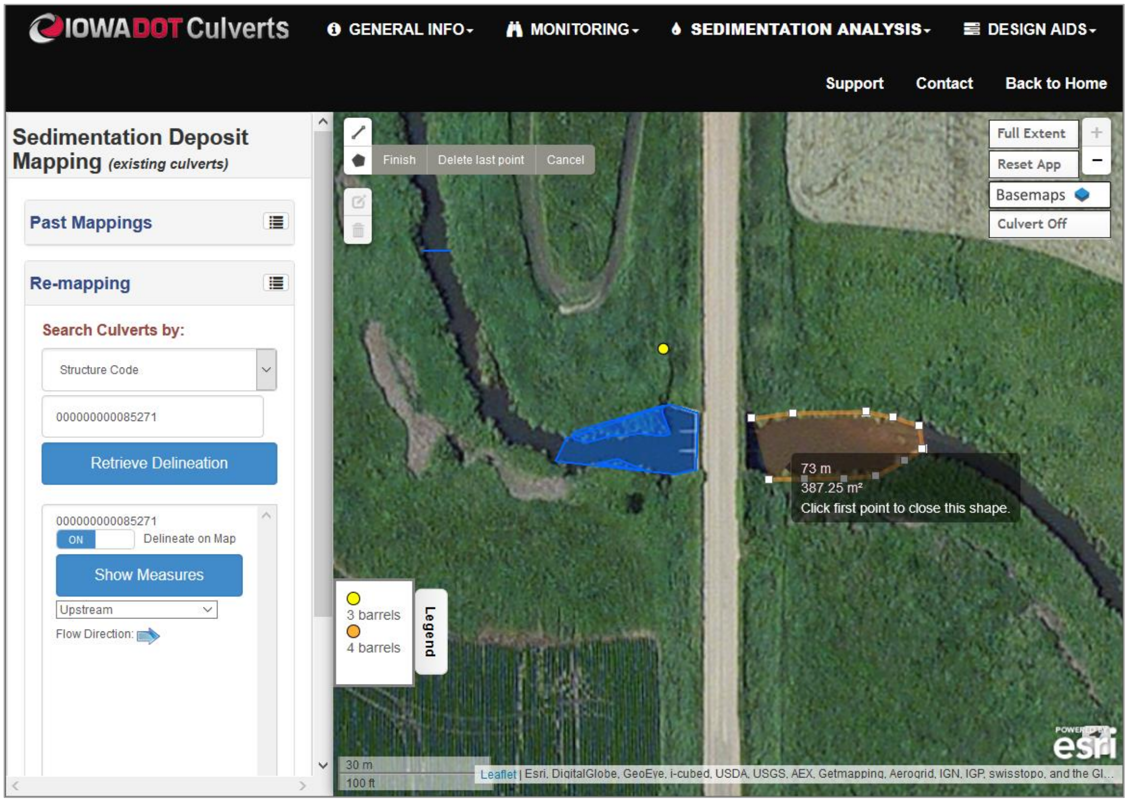


Figure 4.7 The geo-processing tools associated with the "Sediment Deposit Mapping" workflow

All numerical and textual information are retrieved from the specialized IDOT culvert database (SIIMS), while the graphical information produced during the culvert sedimentation mapping (geo-processing-based) is transmitted through Web Feature Services (WFS). As the SIIMS information is continuously updated after culvert construction, it is very useful in aiding the mapping process by providing a glimpse of the dynamics of the geomorphological processes at the culvert site. Given that SIIMS is widely used across the U.S. transportation agencies, it makes the functions developed for the "IowaDOT Culverts" platform easily extendable to other settings across the nation. The flow chart for the "Sedimentation Deposit Mapping" workflow is included in the diagram provided in Figure 4.6, with the outcomes enclosed in the "Sediment mapping" block of the figure.

The "Multi-Criteria Decision Analysis" (MCDA) workflow quantifies the relationship between the degree of sedimentation at culverts and the predictor variables integrated into the portal's database. The MCDA's main objective is to develop quantitative relationships between the degree of culvert sedimentation and the key process drivers within the drainage area of the culvert using deductive hydrological classification, powered by machine-learning and visual-analytics techniques. The quantification of the relationships in the MCDA is typically based on a tree-like hierarchy of criteria and alternatives (decision tree-based model). The overall flux of information for the "Sedimentation Analysis" workflows is illustrated in Figure 4.8. The structure of the tree for the MCDA used in our study is shown in Figure 4.9.

In addition, interactive scientific visualization, consisting of parallel coordinate plots and principle component charts, is used to improve the interpretation of the analysis and enable human judgments in the final stage of the decision support loop. It is important to distinguish between the support for MCDA methods and for the MCDA process: the first focuses on supporting the visualization and technical implementation of applying the method, and the latter provides more general suggestions and good practices for carrying out the whole process in a meaningful way (Mustajoki & Marttunen, 2017). We analyzed 23 MCDA software tools in terms of their applicability to support environmental planning processes and concluded that none of the analyzed software tools can be used without prior experience of MCDA. Our platform hides the complexity of the MCDA from the platform users by not allowing them to interfere with the analysis settings.

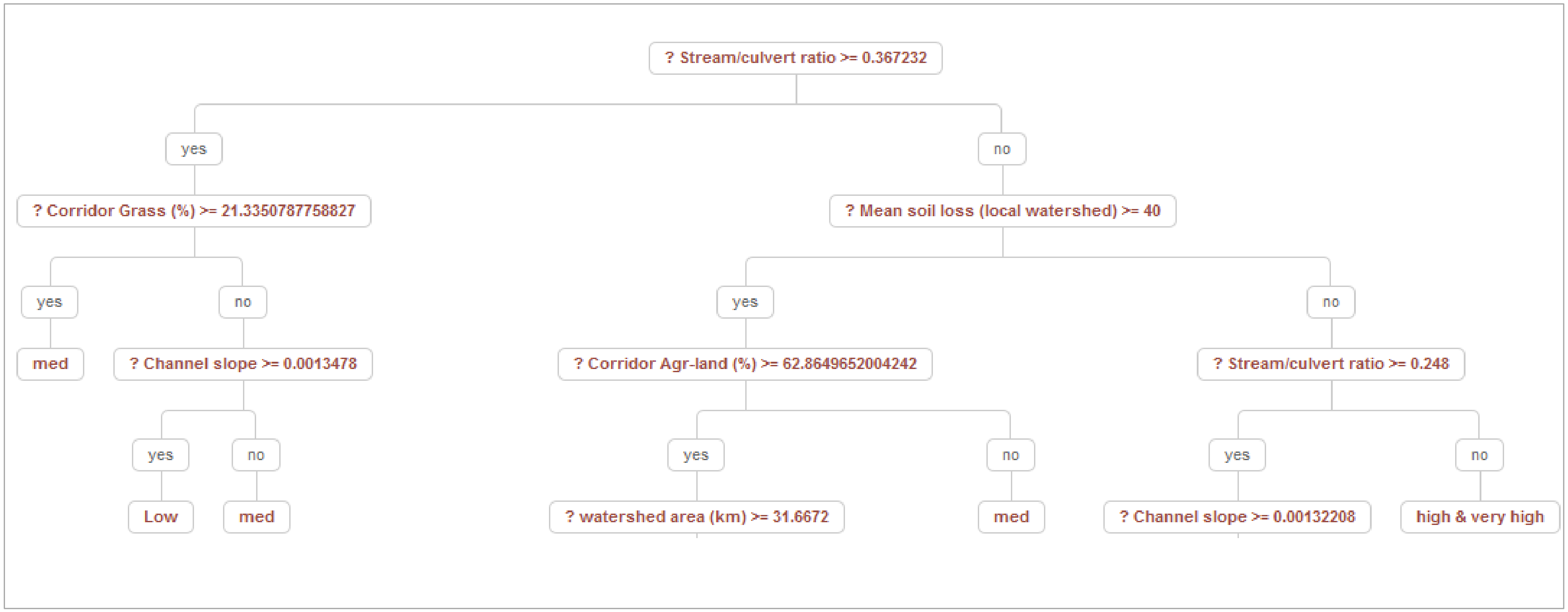


Figure 4.8 The tree-like structure used for the Multi-Criteria Decision Analysis applied to the sedimentation at culverts

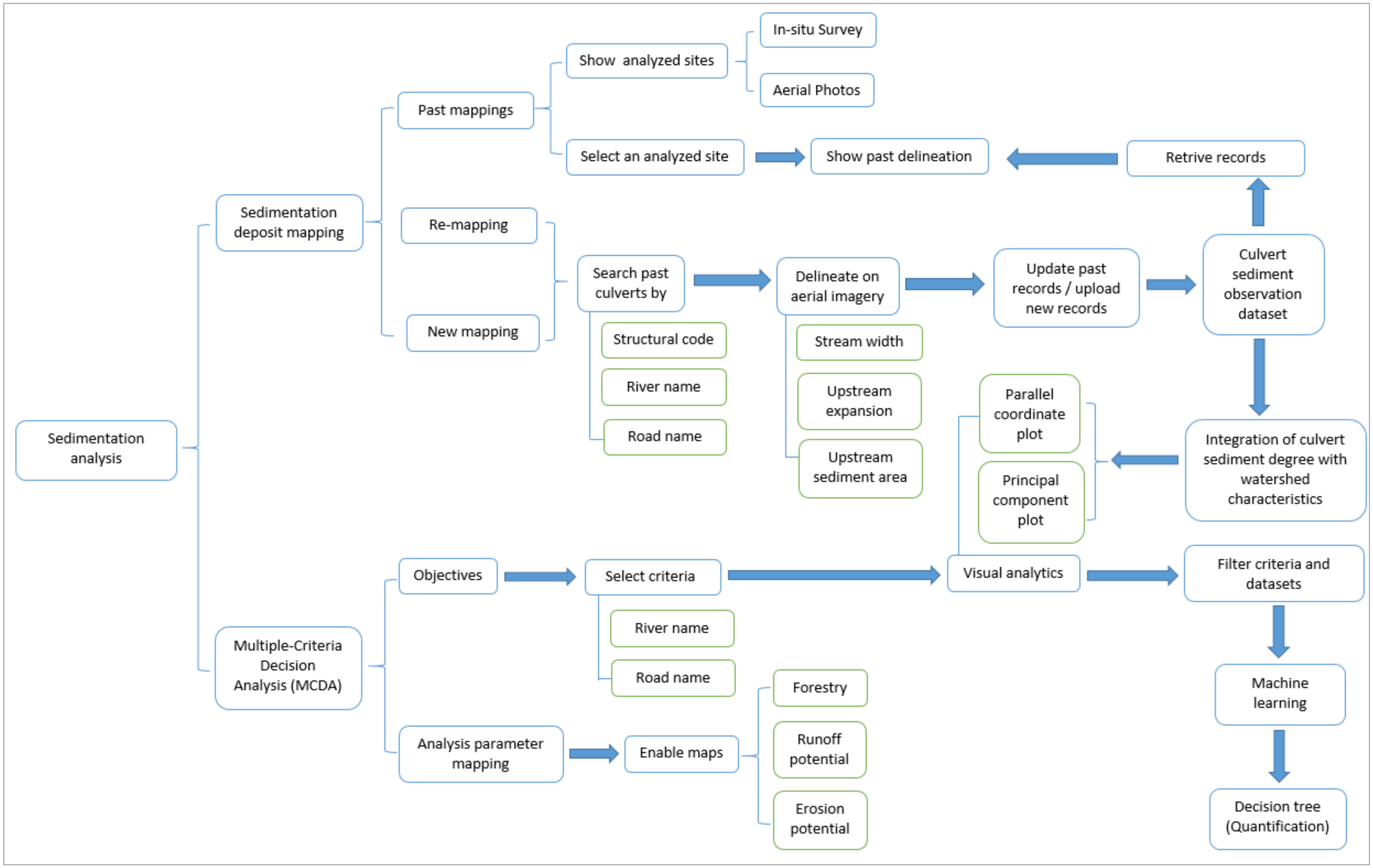


Figure 4.9 The overall flux of information for the "Sedimentation Analysis" workflows

As mentioned in Section 2.4, in this study we use a decision tree ensemble to establish reliable quantifications and cause-effect relationships between culvert sedimentation and drivers. Conceptually, this task includes watershed (drainage area) classification techniques based on supervised machine learning techniques. The techniques proposed for assessing the robustness of the sedimentation prediction is based on tree-ensembles. Tree-ensemble method uses multiple decision tree algorithms to obtain better predictive performance than could be obtained from any of the constituent trees alone. The tree-ensemble method derives an empirical relationship using a tree-like model of key drivers and their possible consequences regarding the culvert sedimentation degree.

The MCDA outcomes are presented through the visual interface of the "IowaDOT Culverts" platform, as illustrated in Figure 4.10. The visualization components presented in the figure include a Parallel Coordinate Plot (PCP), a pie chart for showing the composition of the culvert sedimentation degree, and the Leaflet web map. These components are able to display the dependencies, especially one-to-one relationships, between key process drivers and the culvert sedimentation degree for all the sites that are surveyed based-on aerial images (309 sites). The variables on the vertical axes can be selected for preferential ranges to illustrate the effect of various choices of selection. The results of these on-line analyses are dynamically calculated and updated on the interface in real time. Samples of these analyses are provided in Chapter 5.

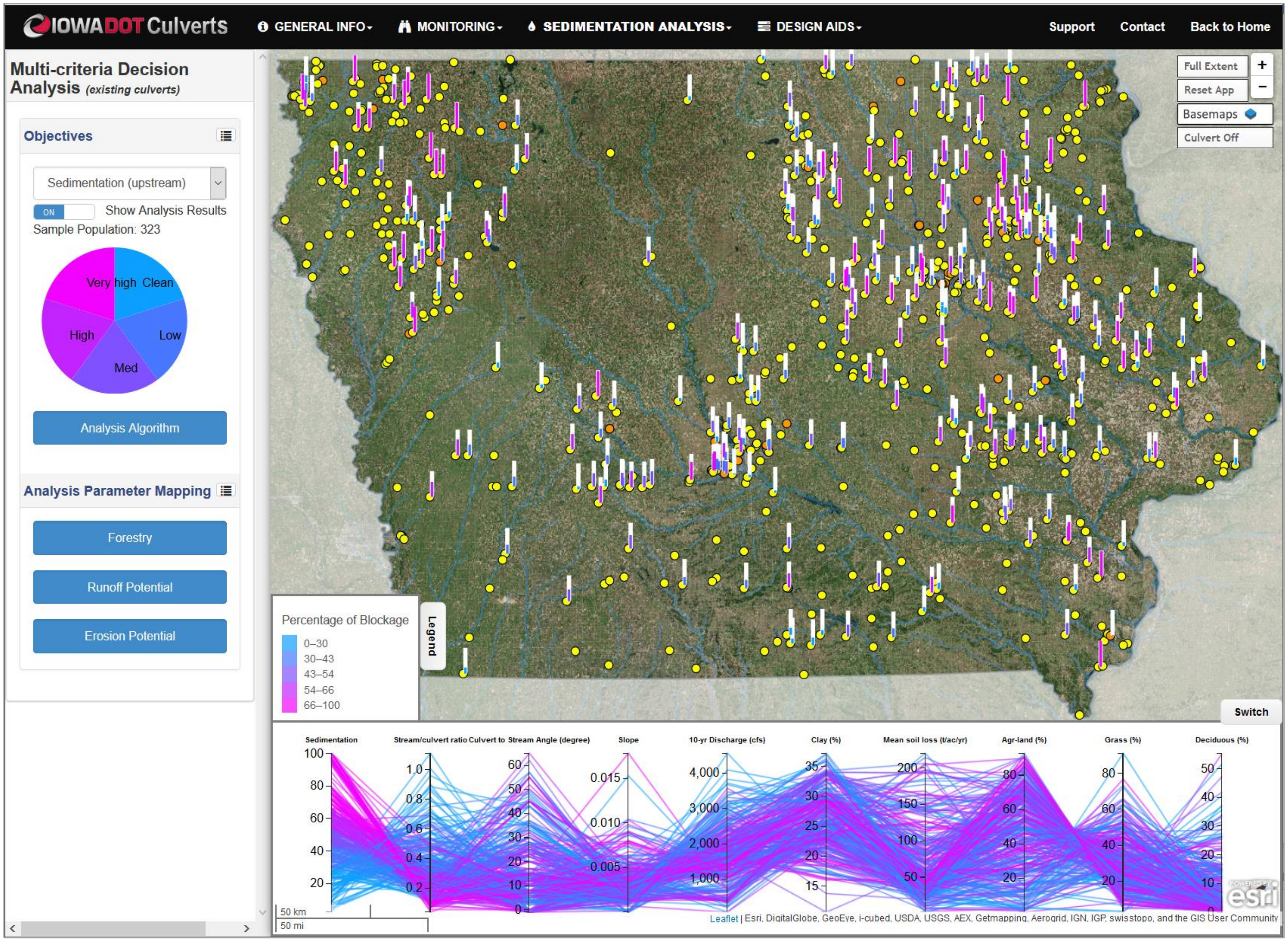


Figure 4.10 Outcomes of the MCDA as applied to all culverts (aerial imagery surveys)

4.4.4 "Decision aids" workflows

This set of workflows is practically the corollary of all the developments carried out through this study. The toolset includes useful aids for the culvert designer or operation personnel by providing the data needed for the estimation of the culvert design discharge and assessment of the degree of sedimentation at the culvert construction site. The tools can be used to estimate these variables at existing or future culvert sites. Based on the degree of sedimentation provided by the forecasting tool, the user can decide if the culvert needs to be re-evaluated in terms of hydraulic design or to be associated with protective measures to mitigate sedimentation. These measures can include practices that reduce erosion and sediment transport in the upland areas or can recommend self-cleaning solutions such as those developed in a companion study on culverts (Muste and Xu, 2017).

The first decision aid of the "IowaDOT Culverts" portal is for the estimation of design discharge that is used in the hydraulic sizing of the culvert. The discharge is calculated using the current IDOT estimation protocol whereby its value is set to the magnitude of the annual exceedance-probability discharge. There are several methods to estimate this probability. This study used the USGS Eash method based on regional regression equations applied to the culvert drainage area (Eash, 2001). The discharge estimation algorithm is the same as the one incorporated in the StreamStats software package developed by USGS to provide users with assistance when using various water-resources planning and management tools (https://water.usgs.gov/osw/streamstats).

The input information required for the design discharge calculation for a planned or existing culvert includes the watershed boundaries and hydrologic observations in the enclosed drainage area. In order to accommodate this input information, the portal uses a geo-processing service based on the watershed delineator and PostGIS functions (i.e., Geospatial extension of the PostgreSQL database) for delineating the drainage area, with consideration of the area overlay with the hydrologic regions defined by the USGS Eash method. The delineation is done using the 90 meter flow matrix-based watershed search engine developed at the Iowa Flood Center (Demir and Szczepanek, 2017). Numerical values for each drainage area are calculated using PostGIS functions called “St-Intersect” under “St-Area”. The numerical results of the area overlay with different hydrologic regions are sent to the front-end through the geo-processing service provided by a PHP script. At the front-end, these results are plugged into the respective regression equations corresponding to the hydrologic regions their divided area lies within. The outcome of the geo-processing tool is illustrated in Figure 4.11.

The forecasting of the degree of sedimentation at culverts is the premier product of this study as it embeds all the artificial intelligence tools incorporated in the “IowaDOT Culverts” platform. This workflow is labeled “Sedimentation Potential Warning” and is located under the “Design Aids” cluster of the “IowaDOT Culverts” platform. The tool provides the degree of sedimentation for existing or new culvert sites. The degree of sedimentation is defined as the ratio of the total area of the expansion upstream from the culvert divided by the area covered by sediment deposits. The forecast is currently based on the “training” of the MCDA that used all the culvert sites analyzed with aerial imagery with the “Sediment Deposit Mapping” tool. From this perspective, it is obvious that the accuracy of the forecast can be improved by adding new analyzed cases to the statistical sample used for training the forecasting. The MCDA is dynamically linked to the mapping database, therefore every addition of a culvert mapping to the database (using any of the methods described in Section 2.3.2) will increase the robustness of the forecast.

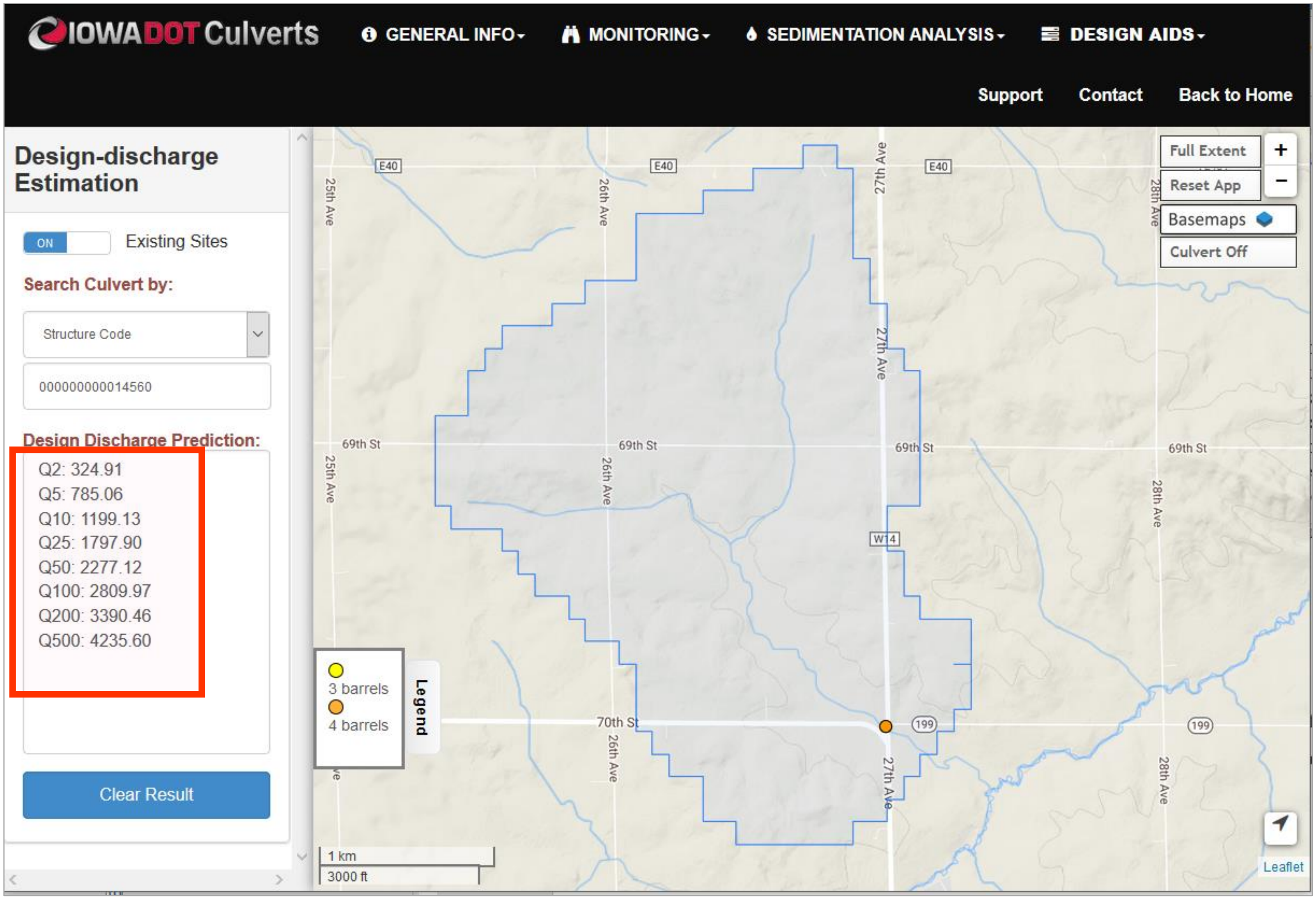


Figure 4.11 The interface for the estimation of the culvert design discharge (discharge estimates for various return periods are provided in the lower box of the left info panel)

The MCDA ingests multiple independent variables (predictors) that define the status of the dependent variable (response variable). For an existing culvert, all the independent variables are fixed as they are related to structural, watershed, and stream characteristics that do not change over short time scales (i.e., soil type, watershed physiography, vegetation cover, etc.). The role of sedimentation forecast in this situation is to validate the forecast prediction for the degree of sedimentation at a given site. For new culvert sites, however, the forecast can play a role in sizing the culvert, as the hydraulic portion of the culvert design allows some freedom in terms of selecting the culvert cross-section or changing the hydraulic controls at the culvert. The flowchart of the information within the "Sedimentation Potential Warning" is provided in Figure 4.13.

Out of all the independent variables of the MCDA, the "stream-to-culvert width" (SCW) ratio is the only variable that can vary at the design stage. Specifically, if the hydraulic design suggests a certain size cross-section, the designer has several choices to combine with the prescribed box shape and sizes that result in wider or narrower total culvert width (labeled with *w* in Figure 4.12a). The designer has also to decide if the inlet and outlet wingwalls associated with the culvert are set oblique or straight. To account for this flexibility, the forecast workflow embedded in our portal allows the user to adjust the SCW ratio when searching for the sedimentation potential. The outcomes of the MCDA analysis re different for the same culvert site commensurate with the value selected for the SCW ratio. In order to support decisions on mitigating sedimentation both in design and operations, the workflow provides the option to adjust the SCW ratio for both existing and new culvert sites (see Figures 4.12b).

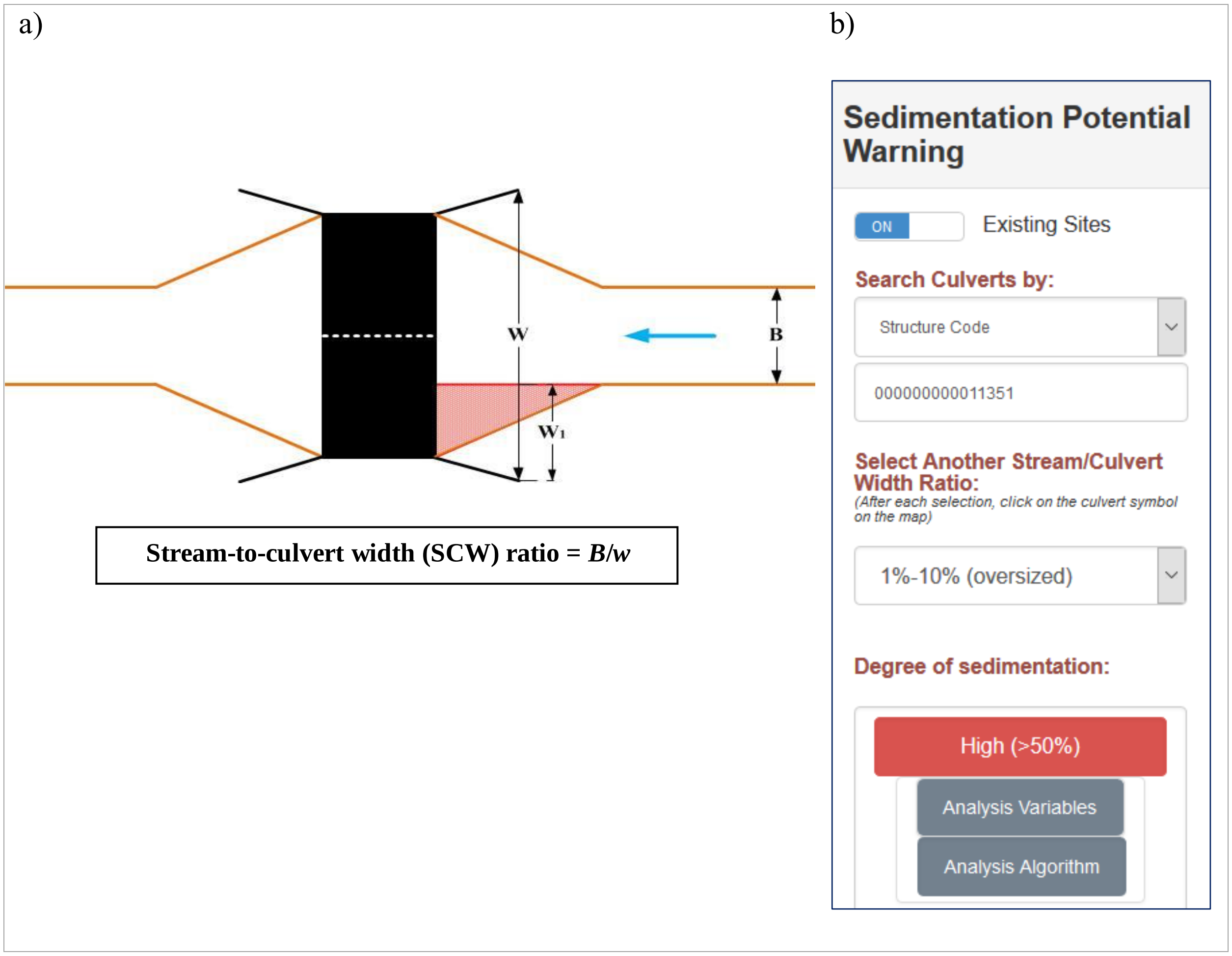


Figure 4.12 “Sedimentation Potential Warning” workflow: a) definition sketch for the stream-to-culvert (SCW) ratio; and b) forecasting interface for existing culvert sites

For new culverts, an additional step required for the MCDA input is the establishment of the stream width at the site. The workflow for the sedimentation forecast at new sites is provided in Figure 4.13. The stream width required for the forecasting input can be estimated using the workflow by using the same geo-processing tool described in the Sediment Mapping workflow. For this purpose, the length measurement tool is activated (see Figure 4.14). The stream reach approaching the culvert can be affected by morphological changes; hence, it might be difficult to detect a “representative” stream width in the upstream area.

For acquiring such a representative stream width we suggest making a visual inspection of the stream upstream and downstream from the culvert to get a good sense of the reaches where the stream appears undisturbed, i.e., lacking channel non-uniformities such as bends, naturally occurring weirs, or other obstacles in the stream. Once that inspection is finished, select a "representative" channel reach. Several cross sections are sequentially measured with the geo-processing tool on the representative stream reach. The portal geo-processing tool records each individual measurement and provides an average of the measurements that can be subsequently used to calculate the SCW ratio. Once this ratio is determined, the forecast is automatically produced in a similar manner as for an existing culvert site case. The degree of sedimentation is expressed in forecast bins corresponding to 1-10%, 10-30%, 30-50%, 50-50%, and 70-100%.

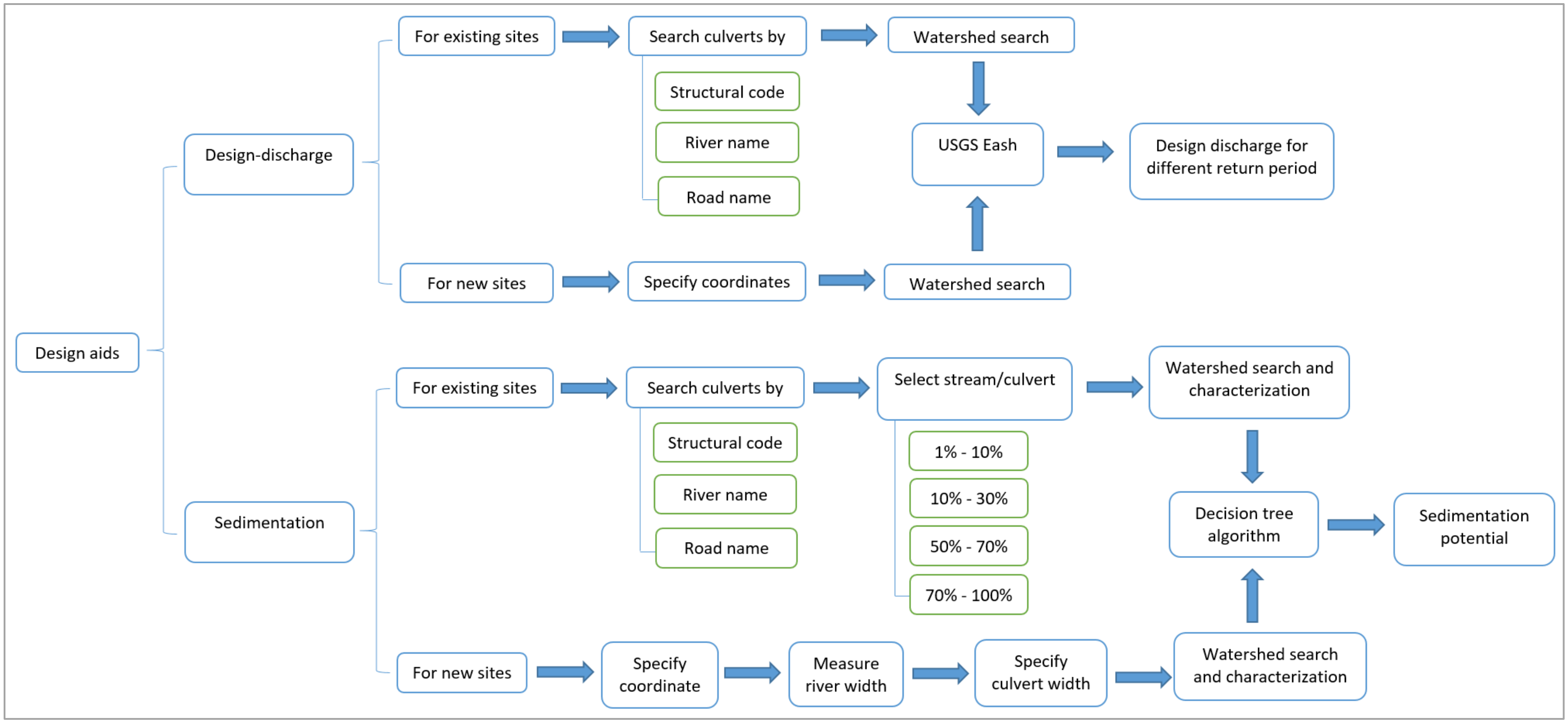


Figure 4.13 “Sedimentation Potential Warning” workflow for new culvert sites - flowchart of the workflow

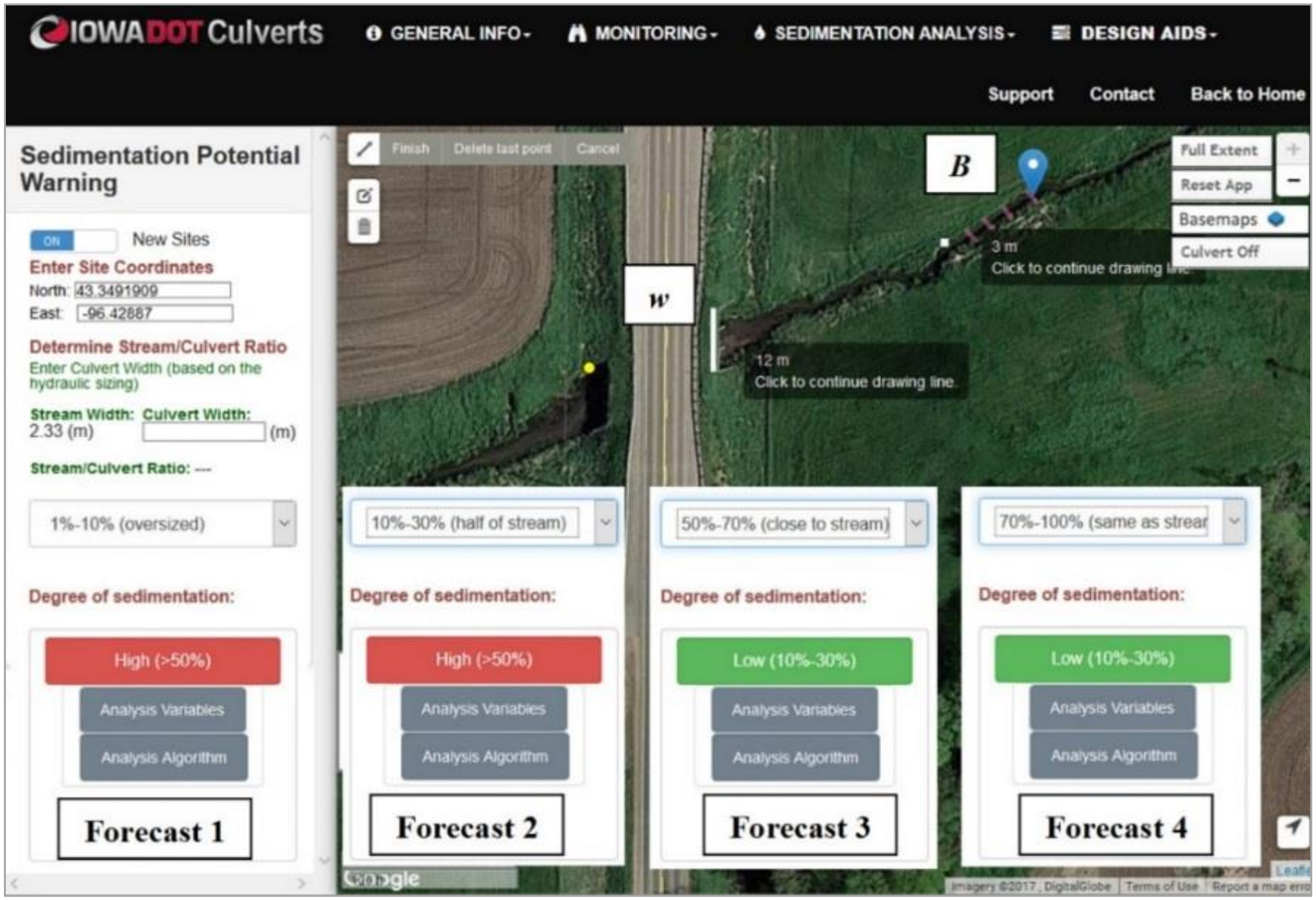


Figure 4.14 “Sedimentation Potential Warning” workflow for new culvert sites - forecasting interface for new culvert sites

4.5 References

Abbott, M. B. (1991). “Hydroinformatics: Information Technology and the Aquatic Environment”. Ashgate Publishing, LimiteAllan, J. D. (2004). Landscapes and Riverscapes: The Influence of Land Use on Stream Ecosystems. Annual Review of Ecology, Evolution, and Systematics. 35 (1): 257-284. doi:10.1146/annurev.ecolsys.35.120202.110122.

Agafonkin, V. (2010). “An open-source JavaScript library for mobile-friendly interactive maps”. Retrieved on November 7, 2017 from: leafletjs: http://leafletjs.com.

Demir, I. and Beck, M. B. (2009). “April. GWIS: A prototype information system for Georgia watersheds”. In Georgia Water Resources Conference: Regional Water Management Opportunities, UGA, Athens, GA, US.

Demir, I. and Szczepanek, R. (2017). “Optimization of river network representation data models for web‐based systems”. Earth and Space Science, 4(6): 336-347.

Demir, I., Conover, H., Krajewski, W. F., Seo, B. C., Goska, R., He, Y., McEniry, M. F., Graves, S. J. and Petersen, W. (2015). “Data-enabled field experiment planning, management, and research using cyberinfrastructure”, Journal of Hydrometeorology. 16 (3):1155-1170.

Demir, I., Jiang, F., Walker, R. V., Parker, A. K. and Beck, M. B. (2009). “October. Information systems and social legitimacy scientific visualization of water quality. In Systems, Man and Cybernetics”, 2009. SMC 2009. IEEE International Conference: 1067-1072. IEEE.

Eash, D. (2001). “Techniques for estimating flood-frequency discharges for streams in iowa”. Water-Resources Investigations Report 00-4233, U.S. Geological Survey, Reston, VA.

ESRI (2015). “ArcGIS for Server”. Retrieved on November 7, 2017. Retrieved on May 23, 2015 from: server.arcgis.com/en/server.

jQuery Foundation (2010). "jQuery: The write less, do more, JavaScript library". Retrieved on April 29, 2010 from jQuery: https://jquery.com/.

Kosicki, A. J., and Davis, S. R. (2001). "Consideration of stream morphology in culvert and bridge design", Transportation research record.1743 (8).

MDN (2015). "AJAX". Retrieved fon June 3, 2015 rom Mozilla Developer Network: https://developer.mozilla.org/en-US/docs/AJAX

Microsoft (2014). "Bing maps API". Retrieved from on November 7, 2017: https://www.microsoft.com/maps/choose-your-bing-maps-API.aspx

Mustajoki, J. and Marttunen, M. (2017). "Comparison of multi-criteria decision analytical software for supporting environmental planning processes". Environmental Modeling and Software. 93: 78–91.

Muste, M. (2014). "Information-centric systems for underpinning sustainable watershed resource management". In Ahuja S. (eds.) Comprehensive Water Quality and Purification. Elsevier: 270-298.

Muste, M., Ettema, R., Ho, H-C. and Miyawaki, S. (2009). "Development of self-cleaning box culvert design". Report for IHRB TR-545, Iowa Department of Transportation, 800 Lincoln Way, Ames, Iowa.

OGC (2009). "Web map service". Open Geospatial Consortium.

OGC (2011a). "OpenGIS Web Map Tile Service Implementation Standard". Retrieved on June 2, 2011 from: www.opengeospatial.org/standards/wmts

OGC (2011b). "Web feature service", Retrieved on November 1, 2011 from: opengeospatial.org/standards/wfs.

Weber, L. J., Muste, M., Bradley, A. A., Amado, A. A., Demir, I., Drake, C. W., Krajewski, W. F., Loeser, T. J., Politano, M. S., Shea, B. R. and Thomas, N.W. (2018). “The Iowa Watersheds Project: Iowa's prototype for engaging communities and professionals in watershed hazard mitigation”. International Journal of River Basin Management, 16 (3): 315-328.

Xu, H. (2015). “Prototyping Hydroinformatics-based systems for supporting decision-making in culvert design and monitoring”. M.S. dissertation, The University of Iowa, Iowa City, IA.

Xu, H. Shen, B. and Muste, M. (2015). “Geo-portal for sustainable culvert design and monitoring”, e-proceedings of the 36th IAHR world congress, 28 June - 3 July, 2015, Hague, The Netherlands.

CHAPTER 5 DATA-DRIVEN INFERENCES ON SEDIMENTATION PROCESSES

5.1 Global analysis

The core element of this study is the MCDA applied to culverts spread over the whole state of Iowa. The MCDA engine embedded in the "Sedimentation Analysis" workflow of the IowaDOT Culverts platform allows users to dynamically display functional relationships between pairs, selected groups, or the entire set of independent variables used for the MCDA analysis. The focal relationships for this study is the one relating the degree of sedimentation at culvert (i.e., the response variable) with all its drivers (i.e., predictors) as they enable useful insights into the processes leading to sedimentation.

To provide system-based mitigation solutions, this study uses the derived relationships from the MCDA to predict the severity of culvert sedimentation at any location and stream in Iowa. Some of the coarseness of the MCDA outcomes are due to the fact that the MCDA presented above has a global nature, i.e., attempts to give equal weight in the analysis to culverts located in areas of various erosion potential. Moreover, the inherent assumption of the global MCDA is that the sediment deposited at culverts is transported with equal chances from any location within the drainage area. Such assumptions might not be accurately replicating the actual physics involved in the detachment and transport of the particles from origin to the point of deposition. The outcome accuracy, and therefore the insights, can be further optimized if additional data-driven algorithms are used in conjunction with the MCDA. These analysis optimization steps are described in the following sections. Despite the recognition of the above-mentioned limitations, the forecasting of the sedimentation potential at culverts using the MCDA outcome as a basis has been implemented to establish a reference for the optimized analysis subsequently described.

5.1.1 MCDA outcomes

When MCDA is applied to the whole sample of culverts available in SIIMS database, useful relationships can be built atop of the set of drivers. Two such relationships are illustrated in Figure 5.1. The visualization of these relationships displays considerable scattering that obviously hampers both the analysis as well as the need inferences. The large dispersion of the data points for some of the drivers is a consequence of the combined impact of the: a) uncertainties of the MCDA method implementation, b) uncertainties in the direct observation of the independent variables, and, c) the fact that the MCDA is applied simultaneously to all the 3-box culverts in Iowa lumped in one large sample.

Using visual analysis of the trends in the scattering of the variables there are some coarse inferences that can be made as follows. By filtering the degree of sedimentation for its extreme values (heavily silted culverts) the overall dependency looks like the one plotted in 5.1a. The dependence suggests the following ranking for the process drivers' contribution (listed in their descending order): low SCW ratio, low design discharge (under 2,300 cfs), high mean soil loss, high agricultural drainage area coverage, low grass coverage, and forest coverages. The stream-to-culvert angle of incidence and the slope are not sensitive factors as they are within narrow range of variation for the State of Iowa. By filtering the degree of sedimentation for the opposite extreme (indicating clean culverts, i.e., less than 20%) the relationship appearance is changing as illustrated in Figure 5.1b. The comparison of the two plots suggests that the most obvious contributing factors to sedimentation are (in their order of importance): SCW ratio, land cover in the drainage area, i.e., agricultural usage or natural cover.

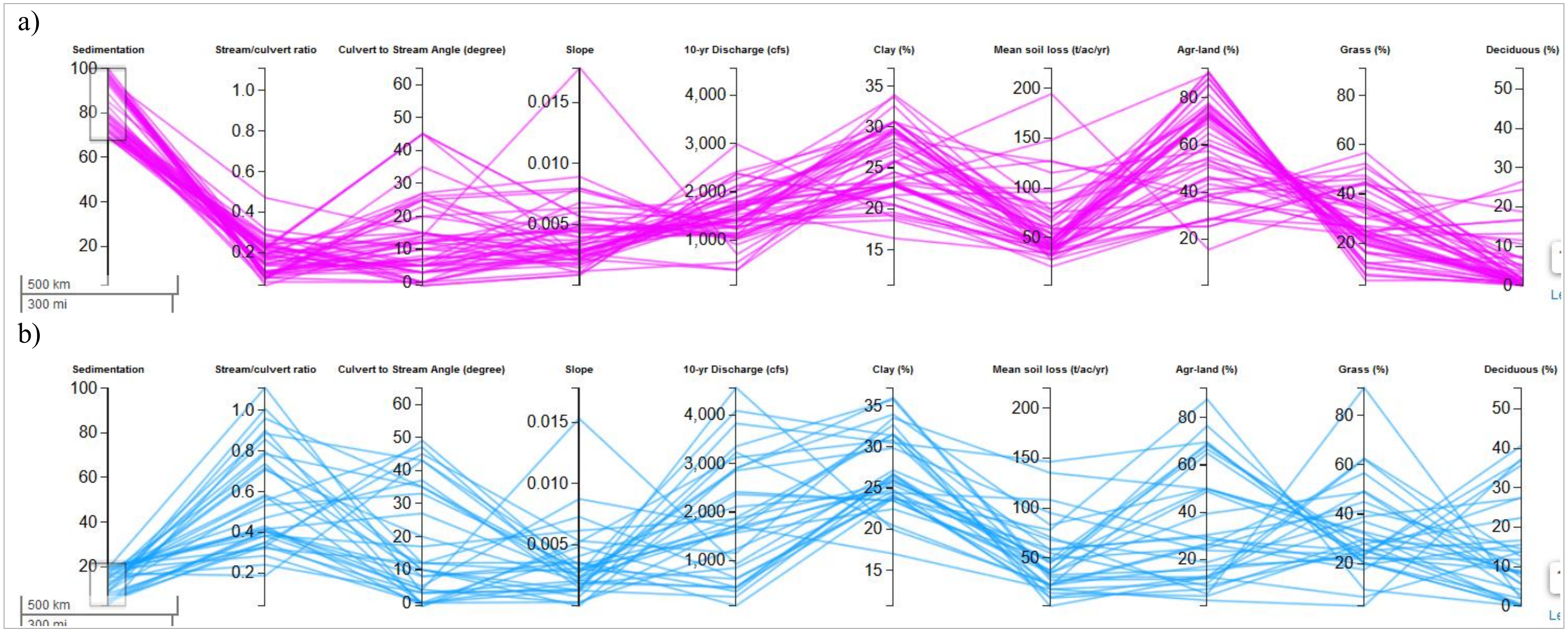


Figure 5.1 Filtering of the degree of sedimentation from the MCDA sample pool for: a) heavily silted culverts, and b) clean culverts

Further insights into the sedimentation processes can be obtained if the dependencies identified above are quantified in one-to-one relationships as illustrated in Figure 5.2. The figure displays the relationship between the degree of sedimentation and the SCW ratio, the driver found most important in the sedimentation process. Specifically, as the SCW ratio moves from small values (i.e., the stream width is considerably narrow compared to the total culvert width) to large values (the stream and culvert have similar widths), the degree of sedimentation varies from "very high" to "clean," as indicated in the figure. This relationship is expected from a physical process point of view as the smaller the SCW ratio, the larger the probability for the stream to develop low-velocity, recirculating flow in the expansion area upstream the culvert. These areas of flow non-uniformity were found critical for triggering the sedimentation process (Muste et al., 2009).

More dependencies can be obtained by "filtering" two or more of the independent variables using windows of pre-established values for the vertical axes corresponding to the analyzed variables. Dependencies of the degree of sedimentation for SCW between 0-20% (found prone to sedimentation in Figure 5.2) and different values for the culvert design discharge (a.k.a. annual exceedance-probability discharge) are illustrated in Figure 5.3. This multi-variable dependency suggests that, for Iowa conditions in general, the dependence is inconclusive as indicated in Figure 5.3 a. However, by further filtering the design discharge, the MCDA suggests that for high design discharges the probability of sedimentation is smaller than for lower discharges (See Figures 5.3 b and 5.3 c, respectively). The dependency illustrated in Figure 5.4 focuses on the estimation of the degree of sedimentation at culverts for a SCW ratio between 40 and 80% (a recommendable design value) and the type of land use in the watershed draining at the culvert. The dependency illustrated that the lower percentage agricultural use (0-25%) leads to a lower degree of sedimentation than the middle range percent value (25-60%).

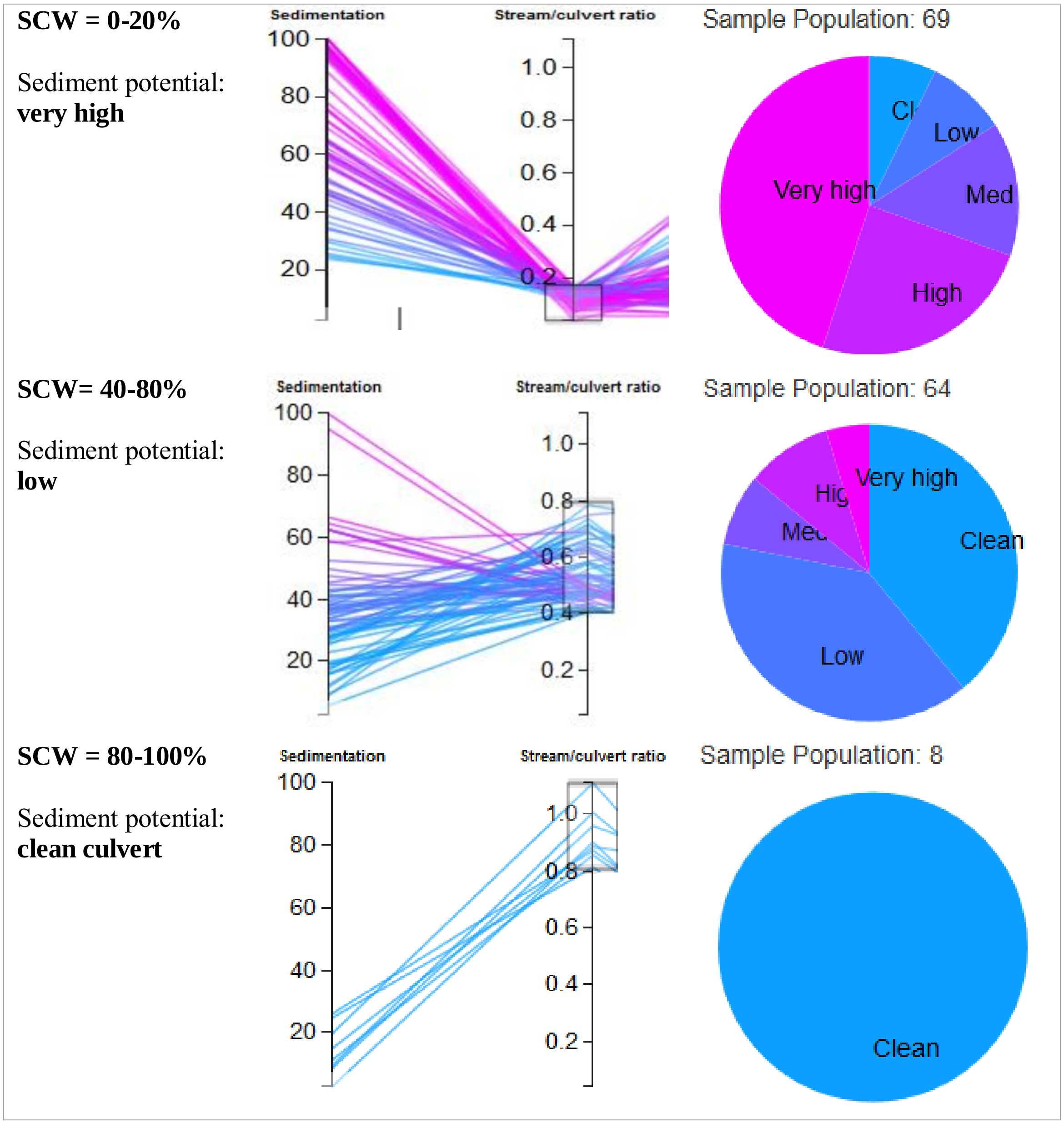


Figure 5.2 Relationship between the stream-to-culvert width (CSW) ratio and degree of sedimentation at culvert

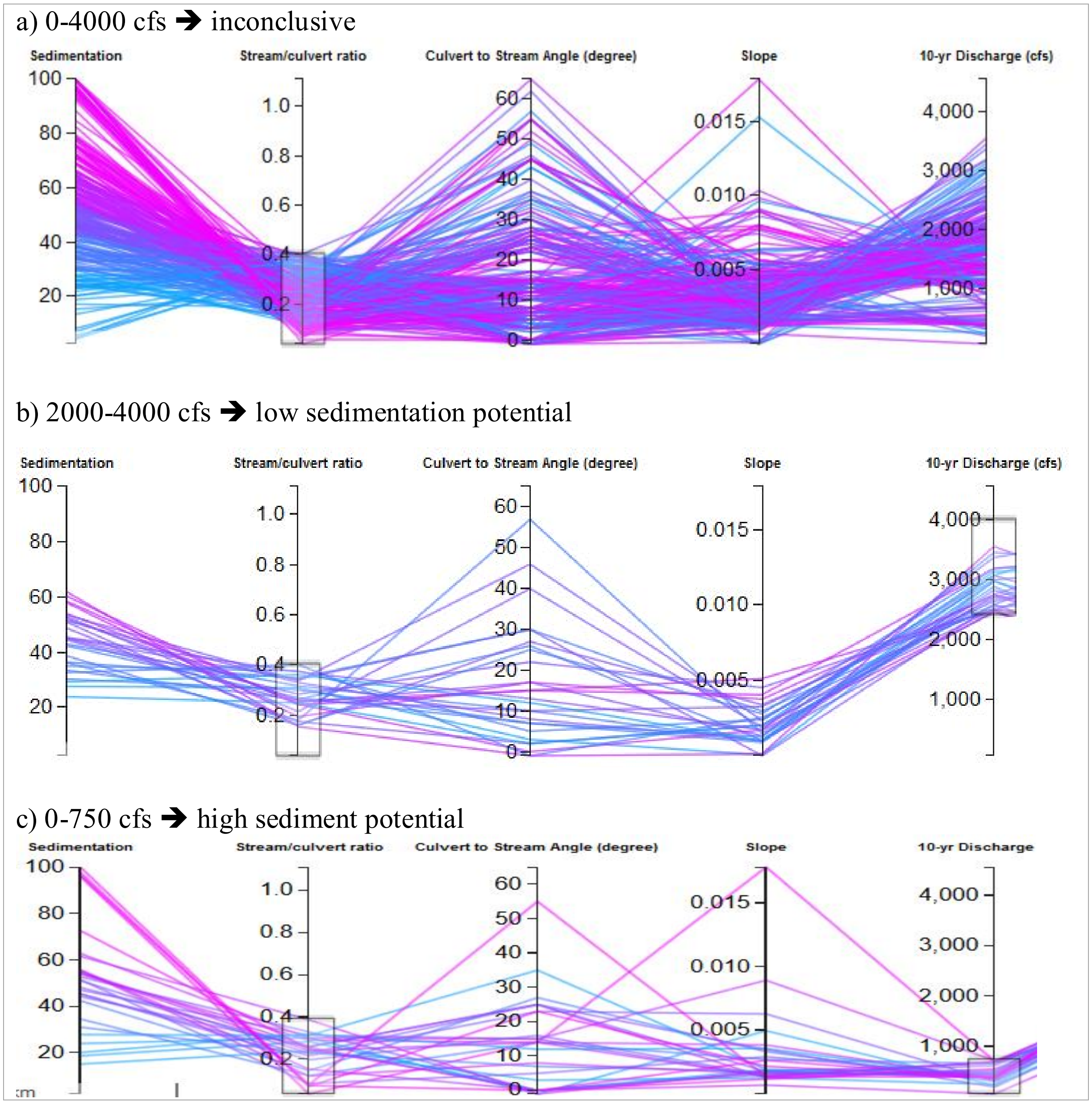


Figure 5.3 The MCDA-predicted degree of sedimentation for CSW ratio in the 0-20% range and variable design discharge

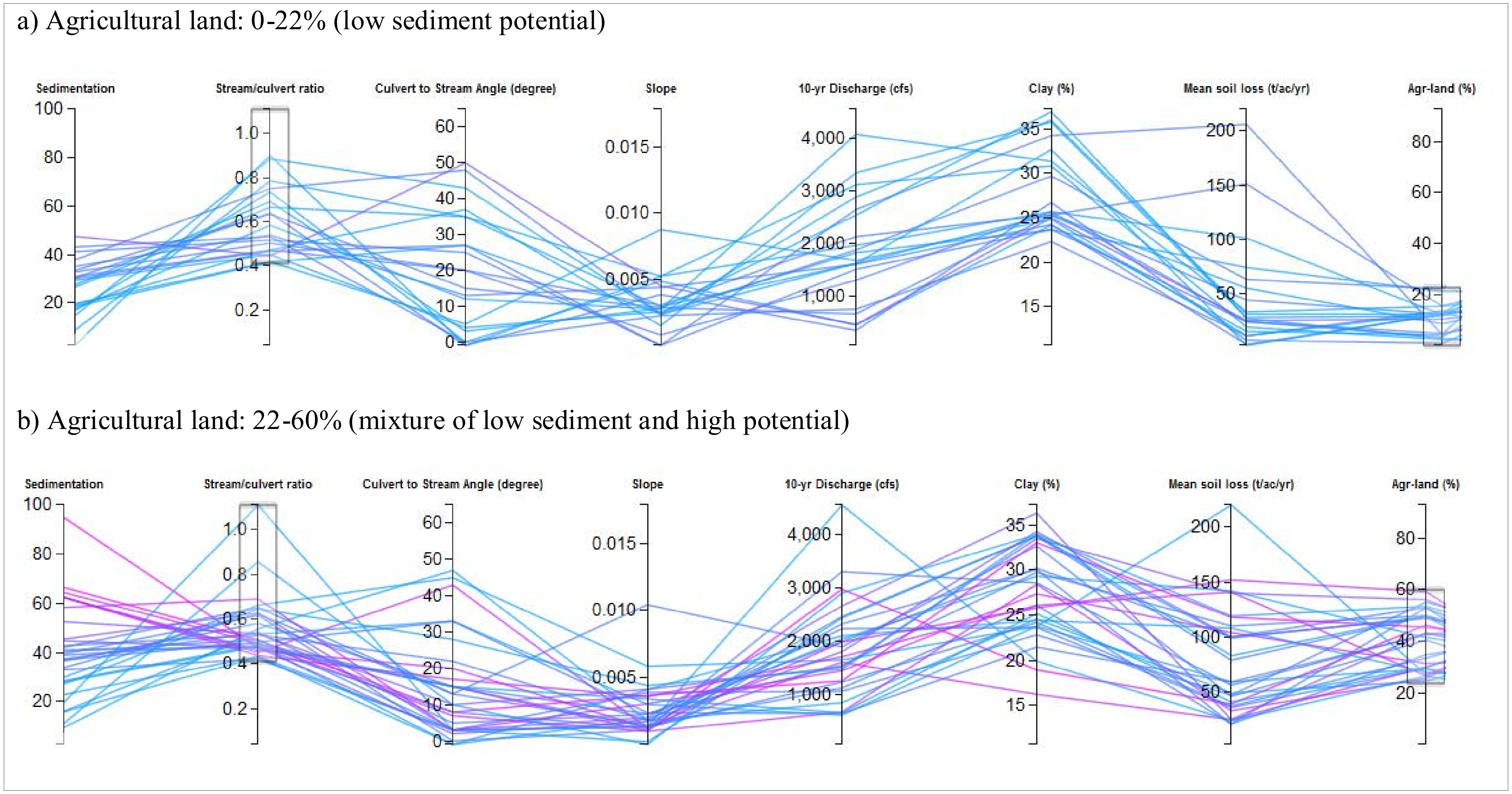


Figure 5.4 The MCDA-predicted degree of sedimentation for CSW ratio in the 40-80% range and % of agricultural use in the culvert drainage areas of: a) 0-22%, b) 22-60%

5.1.2 Performance of MCDA for sedimentation forecasting

For the purpose of estimating the performance of the MCDA using the global approach and of the culvert sedimentation potential forecast based on this approach, this study uses a *K*-Fold Cross-validation on the decision trees that are generated through the MCDA (Devijver and Kittler, 1982; Geisser, 1993). Cross-validation techniques are commonly used in supervised machine learning to evaluate the skills and performances of predictive models by partitioning the original data input into a training set to train the model, and a test set to evaluate it (McLachlan et al., 2004). When a specific value for k is chosen, say k=10, the method is then referred as a 10-fold cross-validation. In other words, k represents the fraction of data from the total sample that is used during the validation; i.e., when $k = 10$ then 1/10 (or 10%) of the data is used as testing dataset, while the rest of the data (1-1/k or 90%) is used for training the decision trees.

The tests conducted with the 10-fold (90% of data as training dataset) and 15-fold (93% of data as training dataset) cross-validation tests on the decision tree results, showing that there is no significant improvement on the model performance by increasing the number of samples used (culvert sites) for the training set with 3%. The lack of sensitivity to the amount of input data indicates that the minimum sample size required for the analysis is met. The results of the cross-validation are summarized in Table 5.1.

More reliable results and accuracy of the dependencies related to the sedimentation at culverts are expected only if more refinements are added to the MCDA analysis. Potential candidates for the MCDA optimization are the regionalization of the analysis based on the similarities of drainage area characteristics and erosion potential, and the sensitivity analysis of the spatial extents that contribute directly to the sedimentation at a specific culvert location. These refinements can be also made using data-driven algorithms as shown in Sections 5.2.2 and 5.2.3.

Table 5.1 Independent variables and spatial extents identified through the feature selection

Number of Cross-fold	Sample size for training	Accuracy
10	279 (90%)	66 - 83%
15	288 (93%)	66 - 82%

5.2 Regional analysis

5.2.1 Overview

As discussed in Chapter 3, this study performs several optimization tasks to further enhance the MCDA's ability for exploring and understanding the mechanism of culvert sedimentation. These optimization tasks aim at improving the estimations of: (1) the critical spatial extent for each key-process driver using river network feature selections, (2) the spatial variability of the culvert sedimentation potential across Iowa through regionalization based on the erosion potential, and (3) the multivariate patterns in the culvert sedimentation dataset through the clustering analysis. These optimization tasks are carried out using a more advanced visual analytics interface that is added to "Iowa DOT Culverts" management system to supplement the MCDA. The addition of the more advanced visual interface will also benefit the potential platform users. Specifically, the non-expert users (transportation practitioners and culvert designers) can gain data-driven insights into the complex mechanism of erosion and sedimentation transport while decision-makers are supported with more practical suggestions for mitigating culvert sedimentation under a variety of factor combinations. For the MCDA optimization described in this section, the number of the culvert for the sample population is increased from 309 sites to 400 sites. The sedimentation degree at additional sites is characterized using high-quality aerial imagery provided by EagleView Technologies, which are acquired from low altitude aerial photography. They complement the first set of 309 sites that were analyzed based on Google Maps wherever the available images were unobstructed by adverse conditions at the site (tall vegetation, obstructive clouds, shadows, or low image resolution). In the following sub-sections, the outcomes and the data-driven insights are produced through different optimization tasks are presented in Section 3.5.

To improve the visual representation of the dependent variable, the culvert sedimentation degree is divided into five groups (representing different sedimentation severity) using a quantile data classification method. The use of the quantile classification is justified by the following reasons: (1) the degree of culvert sedimentation is an ordinal data (sedimentation degrees from low to high), and (2) the data distribution is not extremely skewed (Figure 5.5a). Besides, this method is ideal for classifying ordinal data, as it assigns the same number of data values (culvert sedimentation degree) to each group, therefore avoids creating empty classes or classes with too few or too many values, which may cause bias in determining the regional culvert sedimentation potential across Iowa. The output of the quantile classification is presented in Figure 5.5b.

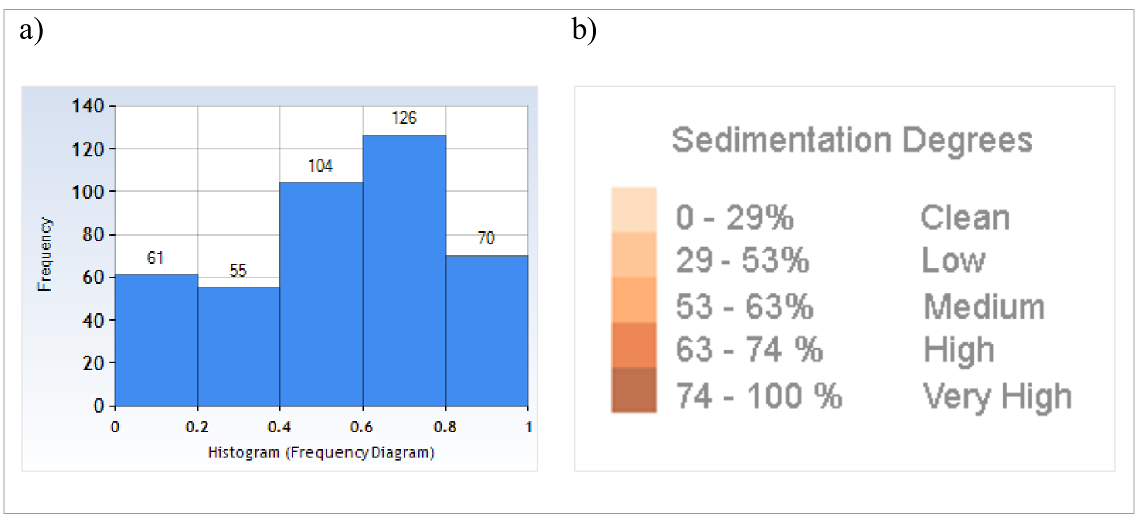


Figure 5.5 Optimization of sedimentation degree: (a) Data distribution and (b) quantile data classification of the culvert sedimentation degree

5.2.2 Sensitivity analysis of spatial extents

In the review of the theoretical considerations on the sensitivity analysis conducted in Section 3.5.2, there are four spatial extents identified as relevant for the independent variables defining erosion and sediment transport occurring within the river network outflowing at the culvert site: (a) drainage area, (b) catchment, (c) river corridor, and (d) immediate corridor upstream of the culvert. Accordingly, each independent variable in the culvert sedimentation dataset is assigned with four indicators representing the above-defined spatial extents. The sensitivity analysis tests the relevance of the finer spatial extents within culverts' drainage areas (i.e., the river corridor and the immediate corridor upstream of the culvert) and incorporates additional variables that are associated with this finer resolution description, namely average stream width and sinuosity.

Following the data preparation for each spatial extent based on their coverage within the drainage area, a tree-based feature selection algorithm (random forest) is applied to the ensemble of variables to evaluate the individual importance of each variable and their associated spatial extent. The three-based feature selection procedure is identical with the one used in Chapter 4. The evaluation is made based on the decrease of the average impurity (entropy reduction) that in turn characterizes the information gain (Buitinck et al., 2013). As the result, a ranking of the most relevant independent variables is produced and shown in Table 5.2.

Table 5.2 Ranking of the independent variables based on feature selection analysis

Variables	Variable importance*	Associated spatial extent
Stream-to-culvert width SCW ratio	0.70	N/A
Landform regions	0.27	drainage area
Design discharge	0.08	drainage area
Drainage density	0.10	drainage area
Ag land (%)	0.16	drainage area
Slope Length and Steepness Factor	0.09	Stream corridor
Forest (%)	0.01	Stream corridor
Grass (%)	0.08	Immediate upstream corridor
Stream sinuosity	0.37	Stream
Stream width	0.21	Stream
* The variable importance is an index estimated by the Python Scikit-learn algorithm based on information gain.		

Afterwards, the descriptions of individual variables involved in the data-driven analysis are characterized using the spatial extent that is indicated by the sensitivity analysis as the most sensitive to the culvert sedimentation degree. For example, the percentage of forestry is displaying the most sensitive response in the relationship for the spatial extent of stream corridor, thus in the later analysis the summary statistics of this variable is characterized at the stream corridor extent. In addition, the newly added variables that characterize channel planform (i.e., stream width and sinuosity) are deemed as important contributors to the culvert sedimentation degree, which are confirmed by the qualitative estimations provided in Figure 5.6. After the sensitivity analysis, the study updates the list of key drivers with the selected independent variables characterized with the proper spatial extents, and utilizes these drivers for the following analyses.

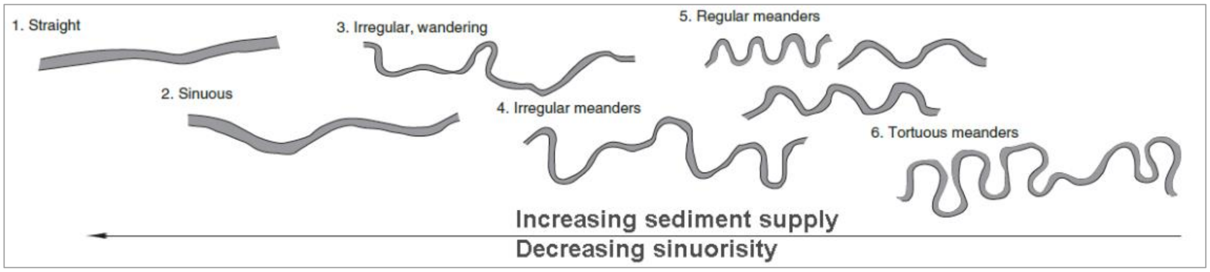


Figure 5.6 Channel patterns (reproduced from Kellerhals et al., 1976)

5.2.3 Regionalization

The rationale for regionalization stems from the idea that using all the culverts located within a large administrative-based unit (i.e., the state of Iowa) for the MCDA analysis might be not be as relevant in terms of outcomes compared to the alternative approach of identifying natural landscape units that show similarity with respect to erosion and sedimentation. In principle, through the regionalization, the study applies the MCDA analysis per such identified sub-regions. The inferences that can be made using the alternative approach are supposedly more relevant with respect to the physics of the erosion and sedimentation processes as the analysis are conducted in areas with similar vulnerabilities to the sediment production. To identify the similarities between areas with regards to the erosion potential, the study applies the Self-Organizing Map (SOM protocol described in Section 3.5.3) to all the culverts across Iowa.

The spatial patterns suggested by SOM applied to all the culverts within the Iowa state reveals a striking similarity with the areal locations of the six landform regions of Iowa (Prior, 1991). These sub-regions are: (1) Northwest Iowa Plains, (2) Des Moines Lobe, (3) Iowan Surface, (4) Paleozoic Plateau, (5) East-Central Iowa Drift Plain, and (6) Southern Iowa Drift Plain. This partition stems from the genealogy of the landscapes in specific areas in Iowa echoing their historical formation and development. Each of these regions feature quasi-similar characteristics with respect to soil type, topography, erodibility, and mobility (Scheidegger, 1973; Prior, 1991).

The map in Figure 5.7 pairs the landform regions with the SOM results with the latter reflecting the characterization of the multivariate clustering outputs. It is obvious from the map that the SOM characterization of the culvert sedimentation degree is quite different depending across different landform regions. More specifically, culverts with a high potential for sedimentation are located in the Northwest Iowa Plains and Paleozoic Plateau while the degree of sedimentation is considerably lower in the Des Moines Lobe. In between these extremes, the culverts located in the Iowan Surface and Southern Iowa Drift Plain display a mixed sedimentation potential. In the present analysis, only six of the eight Iowa sub-regions are considered as the Loess Hills and Alluvial Plains regions do not contain any culvert samples even with consideration of all samples analyzed for this study.

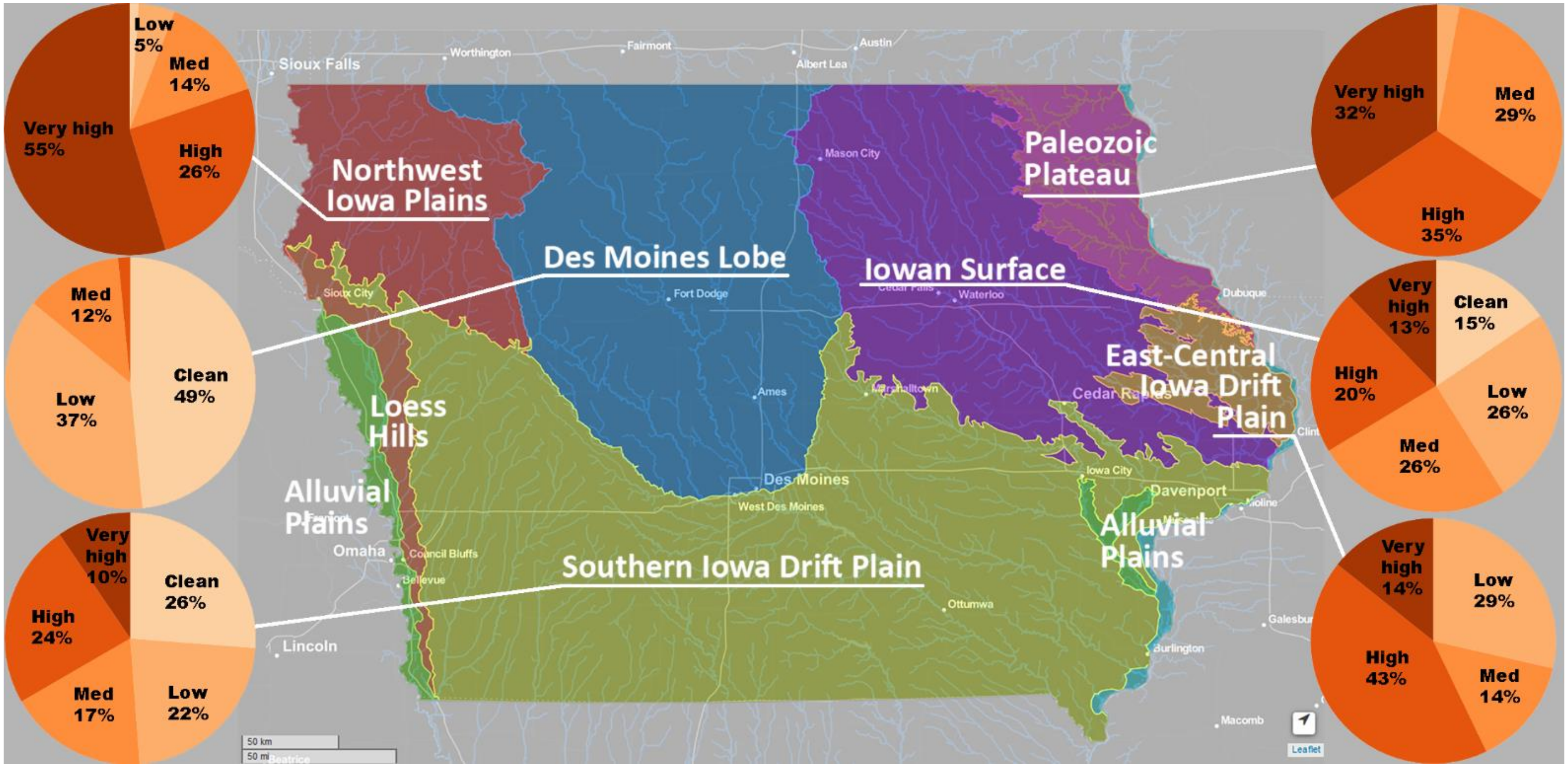


Figure 5.7 SOM estimation of the potential for culvert sedimentation across different landform regions in Iowa

It is remarkable to observe that the regionalization produced by a data-driven investigation tool matches the areal distribution of distinct regions with respect to their genealogy. This can indeed be stated as an ad-hoc validation test. Additional validations of more rigorous nature can be considered as the comparisons of SOM analysis outcomes with physics-based model surrogates. For this purpose, this study compares the culvert sedimentation potential estimated by SOM for each landform region with the erosion potential estimated by: (1) Water Erosion Prediction Project (WEPP) and (2) Revised Universal Soil Loss Equation (RUSLE). The WEPP output is acquired from the Iowa Daily Erosion Program (DEP) in the form of the average hillslope soil loss in Iowa from 2013 to 2018 (past 5 years). The RUSLE output is retrieved from the Iowa Geological Survey, being considered as a good measure of the average soil erodibility. Figure 5.8 illustrates the results obtained with WEPP and RUSLE models.

Although neither WEPP nor RUSLE directly evaluates the culvert sedimentation potential per se, their estimation of soil loss and soil erodibility for the Iowa landform regions are directly related to the magnitude of soil erosion across Iowa, thereby provide meaningful means for validating the regional culvert sedimentation potential derived through this study using the data-driven approach. The comparison of the three alternative approaches for estimating sediment-related characteristics within the Iowa landform regions is presented in Table 5.3.

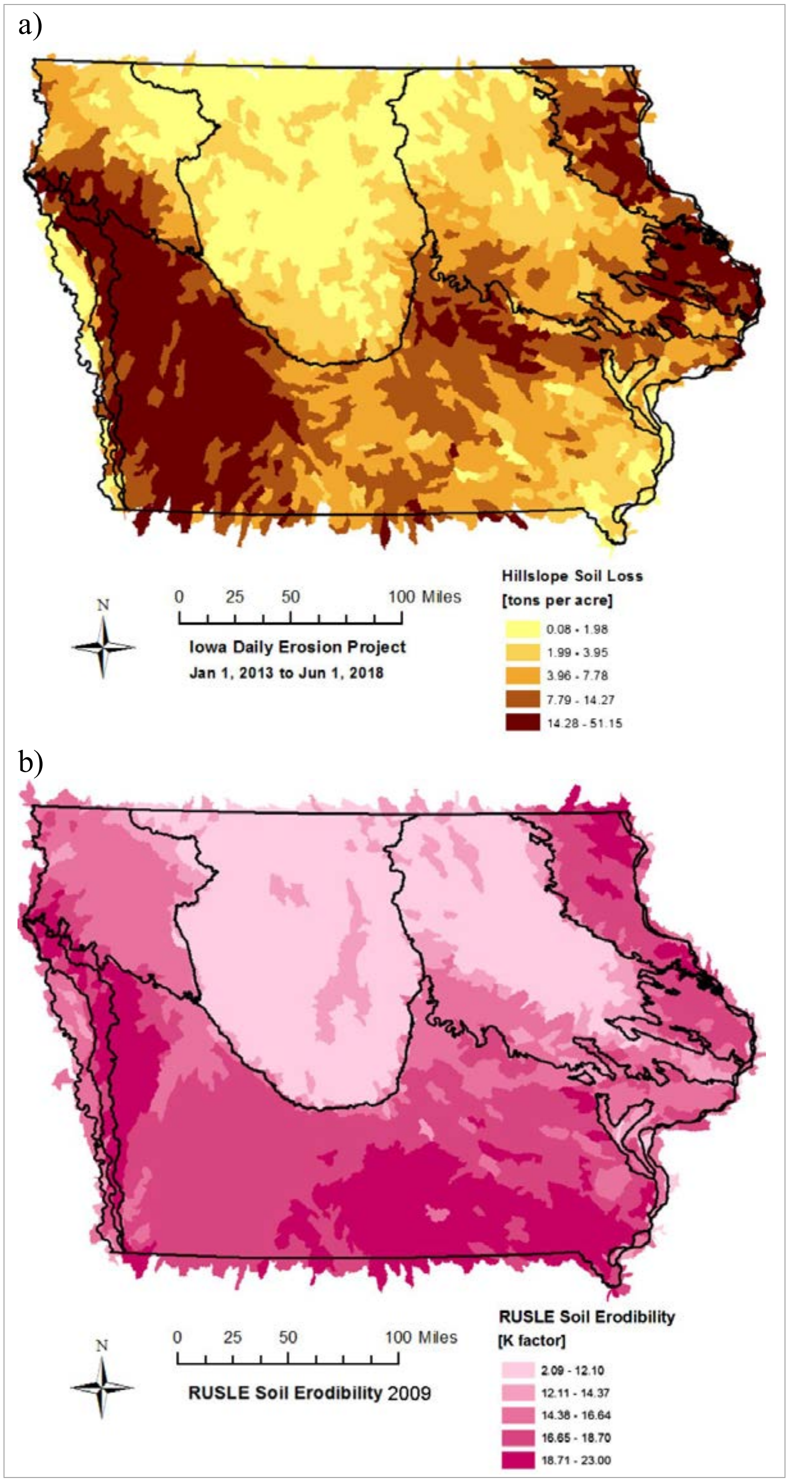


Figure 5.8 Soil erosion potential across Iowa landform regions that are estimated through: (a) average hillslope soil loss (from 2013-2018) estimated with WEPP, and (b) soil erodibility (K factor) estimated with RUSLE

Table 5.3 Comparison between the culvert sedimentation potential and soil erosion potential estimated by RUSLE and WEPP

Landform regions	**Culvert sedimentation potential**	**Average soil loss from WEPP**	**Soil erodibility from RUSLE**
Des Moines Lobe	low	low	low
Northwest Iowa Plains	Very high	mixed	high
Paleozoic Plateau	high	high	high
Iowan Surface	mixed	mixed	mixed
Southern Iowa Drift Plain	mixed	mixed	high
East-Central Iowa Plain	high	high	high

It is rational to assume that the potential for culvert sedimentation is positively correlated with the erosion potentials estimated from the two models, as a higher soil erosion potential usually implies that culverts within a region receive more sediment supply. By comparing the result of this study with the outputs from RUSLE and WEPP, it can be confirmed that, for most landform regions, the culvert sedimentation potential estimated by SOM is consistent with the erosion potential estimated by both models. The only two regions where the SOM results are inconsistent with the modeling results are the Northwest Iowa Plains and Southern Iowa Drift Plain. The Northwest Iowa Plains displays a "very high" culvert sedimentation potential predicted by the data-driven approach, but is quantified as "mixed" in terms of erosion potential by WEPP. Conversely, the Southern Iowa Drift Plain exhibits a "mixed" culvert sedimentation potential predicted by the data-driven approach, while the RUSLE soil erodibility indicates that the region suffers from a high soil erosion.

The differences in the results between SOM and models in the Northwest Iowa Plains and Southern Iowa Drift Plain regions are not solely attributable to lower SOM performance. The first source for the difference is the fact that most of soil erosion models (including WEPP and RUSLE) do not capture the processes related to sediment deposition near culverts, which is affected by many local in-stream conditions. Another source for those differences can be attributed to certain limitations of the two soil erosion models. This latter source of differences in the SOM-model comparison is discussed below. Figure 5.9 compares the average trend of culvert sedimentation drivers in different landform regions using parallel coordinates plots (PCP). The color-coded lines in the plot represent averages of the key drivers within the six Iowa landform regions, which are spatially portioned through SOM.

From this comparison, it can be observed that the Northwest Iowa Plains (represented by the red line in Figure 5.9) has some extreme values for some of the independent variables (i.e., the highest percentage in Ag land, and the lowest percentage in forestry and grass coverage). These extreme landscape characteristics are considered as major contributors (stressor) to severe soil erosions (Blanco-Canqui and Lal, 2010; Montgomery, 2007). Since watersheds in the Northwest Iowa Plains has the highest average agricultural land coverage, it is likely that the erosion potential in the landform region is underestimated by WEPP, as the model does not account for gully erosion (Merritt et al., 2013). This type of erosion is known to be a significant cause of soil loss from agricultural fields (Jahantigh and Pessarakli, 2011; Zaimes and Schultz, 2012). Similarly, RUSLE trends to overestimate the soil erosion in the Southern Iowa Drift Plain region (represented by the green line). This is because this region has a large portion of wooded terrain in its southern areas (a.k.a. Central Irregular Plains). One of RUSLE's major limitations is that it cannot estimate soil erosions in undisturbed forest lands (NSERA, 1995; Renard et al., 1997).

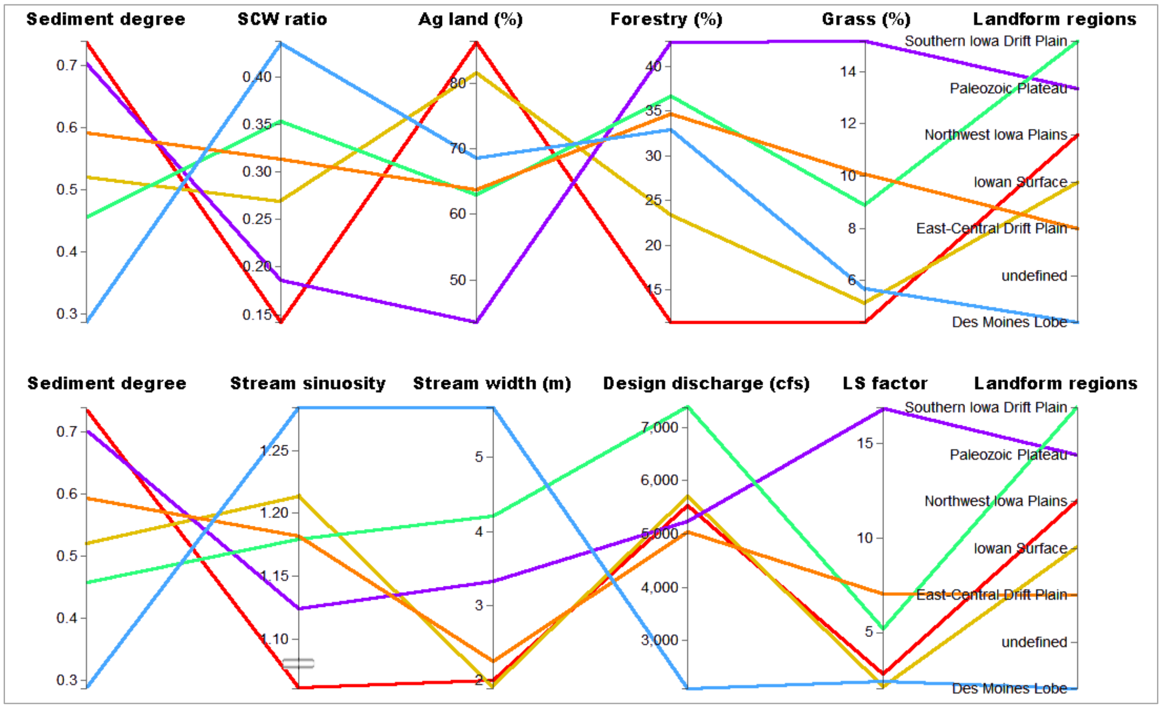


Figure 5.9 Relationships between the degree of sedimentation and the average values of the of key drivers for the Iowa landform regions

In summary, the MCDA optimization using the regional analysis based on the geological landform regions seems to be a reliable building block for developing system-based mitigation strategies for sedimentation at culverts by connecting the spatial variability of sedimentation processes with the landscape classifications from where the sediments are originated. The use of regionalization maps such as the one shown in Figure 5.7 can guide the culvert design and management accounting for sedimentation at culvert by offering data-driven insights on: (1) culvert sedimentation-prone regions in Iowa, and (2) the overall relationships between the average trend of drivers and the culvert sedimentation degree within each landform region (as shown in Figure 5.9). These data-driven inferences can be encoded into culvert design suggestions to help transportation agencies make rapid assessments on sedimentation risk at new culverts based on landscape features in the drainage area and surroundings, as well as to prioritize sedimentation mitigation at existing culverts across Iowa. These inferences demonstrate the capabilities of the data-driven approach to aid the efforts on mitigating sedimentation, a subject where there is no alternative method to inform the design and management of these structures.

5.2.4 Multivariate analysis within individual regions

In addition to the spatial sensitivity and regional analysis discussed above, further optimization of the analysis and forecasting of the culvert sedimentation potential can be expected by investigating what drivers, or combination of them, are responsible for the mix of culvert sedimentation within individual regions. This additional optimization is especially useful in areas such as the Iowan Surface region and Southern Iowa Drift Plain regions where the inter-variable relationships are not so well defined.

The above optimization goal can be attained by employing multivariate clustering analyses. These analyses drill down from the regional to the local level (culverts within each landform region) to: (1) interpret how different landscape features affect culvert sedimentation-related processes, (2) identify the one-to-one relationship (dependency) between the degrees of culvert sedimentation and the dominant drivers, and, (3) identify the multivariate patterns (relationships) between the culvert sedimentation and the drivers.

5.2.4.1 Northwest Iowa Plains

Located in the northwest corner of Iowa, the Northwest Iowa Plains exhibits a diversity of terrain features and geologic materials that are individually dominant in other landform regions (as shown in Figure 5.10a). The region has open rolling hills formed through the intense erosion that accompanied the period of maximum glacial cold (Prior, 1991). This erosional history has left the region with an integrated drainage network and an abundant loess mantle that is nearly continuous across the region (IGS, 2017; Prior, 1991). Some of the landscape features in this region are considered as main contributors to its "high" culvert sedimentation potential (see Figure 5.10b).

Loessial-dominated regions are known for: (1) high susceptibility to heavy erosion (Kung and Chiang, 1977; Piest and Ziernnicki, 1979) and (2) high soil fertility (Catt, 2001), thereby fostering heavy agricultural activities that are generating additional erosion. The information provided by the multivariate analysis in Figure 5.11 enforces the inferences on the sediment potential in this regions by visualizing on the geographic and SOM maps, the distributions and relationships among drivers and the sedimentation degree for the culverts in this region.

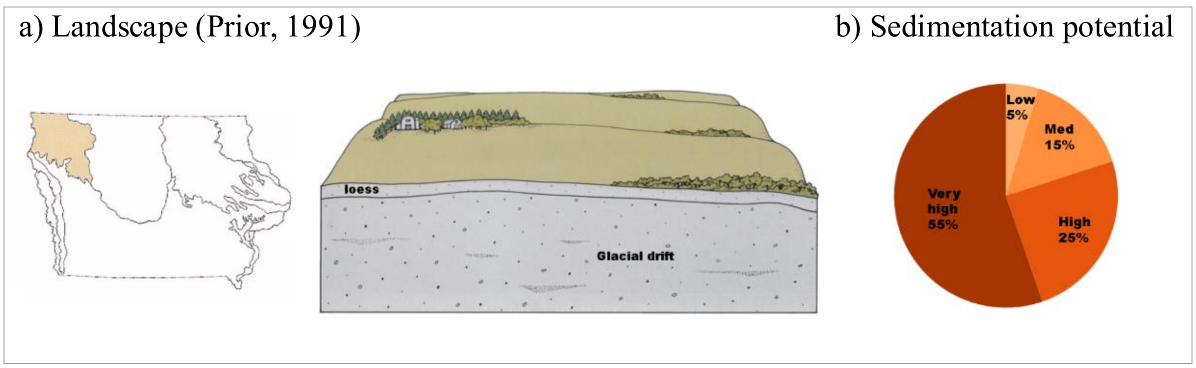
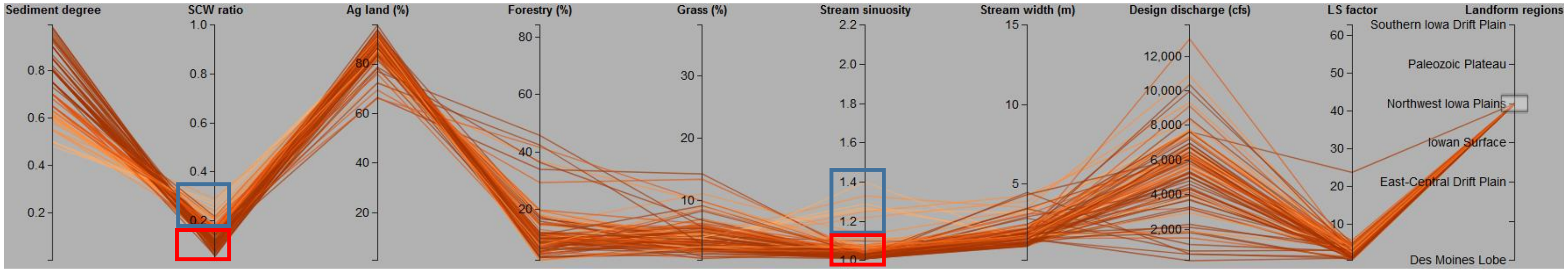
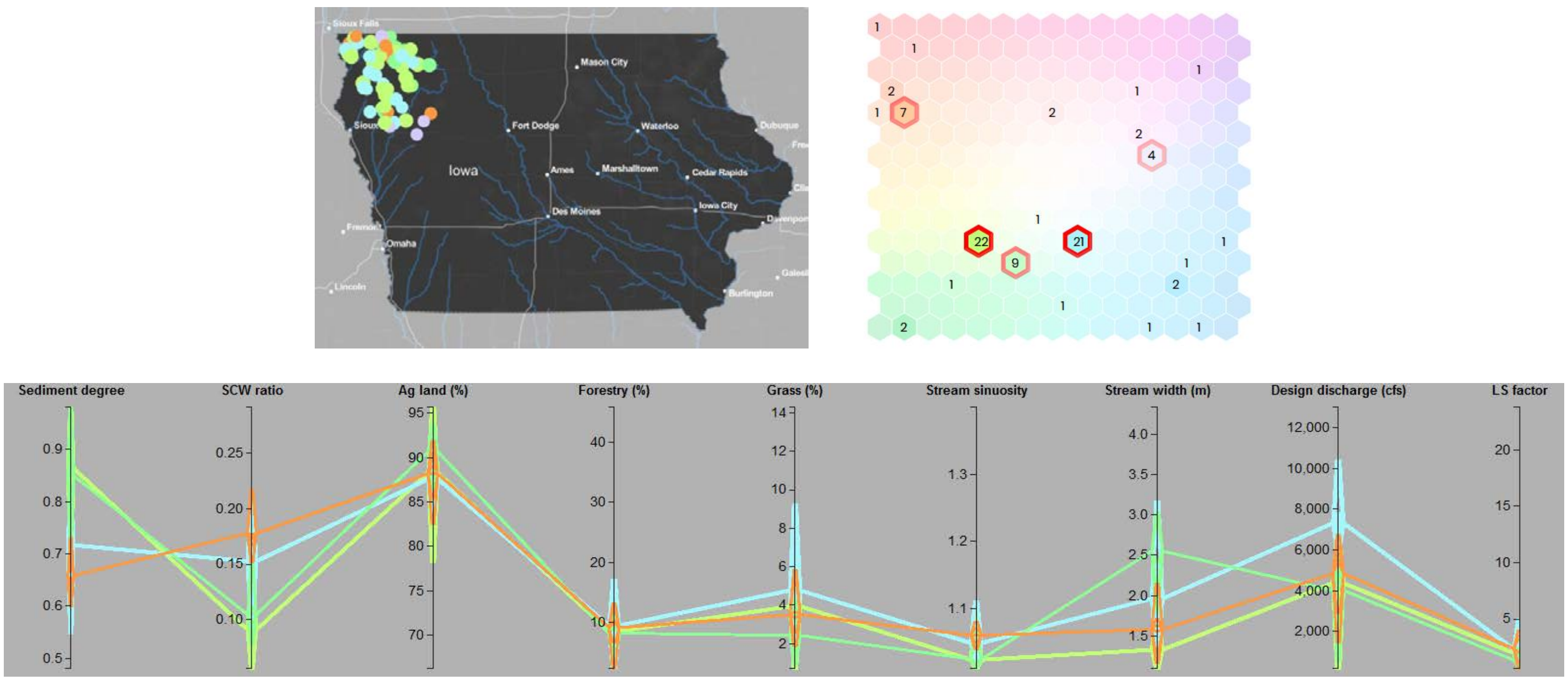


Figure 5.10 Northwest Iowa Plains Region

a) Essential one-to-one relationships (highlighted by the blue and red boxes)



b) Multivariate patterns visualized on the geographic and SOM maps



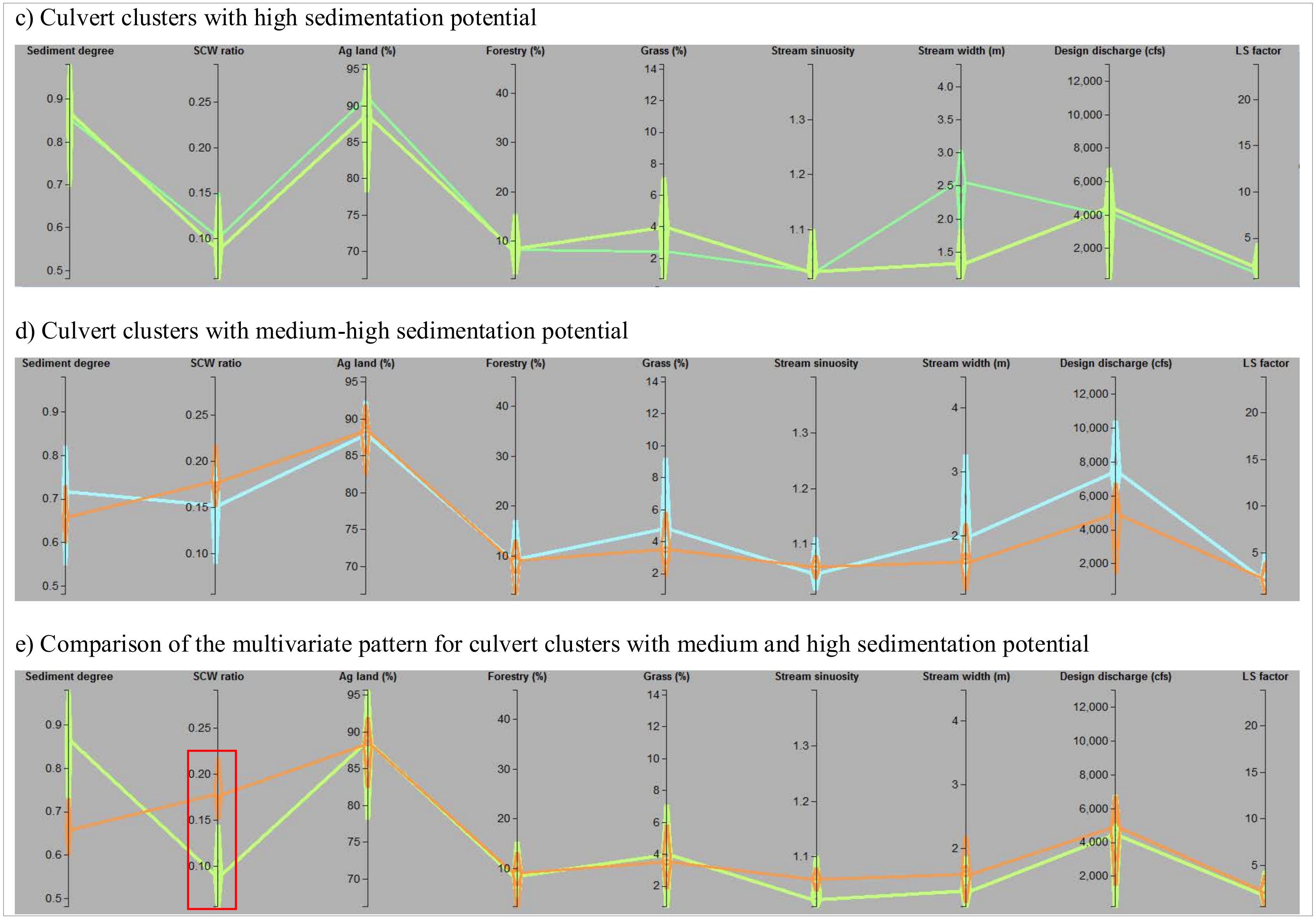


Figure 5.11 Relationship between the culvert sedimentation degree and drivers within the Northwest Iowa Plains region

Figure 5.11a indicates that the SCW ratio has the strongest one-to-one dependency with the culvert sedimentation potential in this region. Outcomes of the multivariate clustering analysis are visualized in Figure 5.11b in the form of the visual interface presented in Section 3.5. The visual interface consists of a SOM map, a geographic map, and an edge-bundling PCP. Each hexagon in the SOM map represents a culvert cluster. Each cluster has assigned a unique color through a unique 2D color scheme shown in Figure 5.11b. The distance between two hexagons (clusters) in the SOM map represents the multivariate dissimilarity between the two culvert clusters. The sample size of each cluster is shown on top of the corresponding hexagon. On the geographic map, each culvert location is displayed using the color-coded convention associated with the corresponding culvert clusters. The edge-bundling PCP presents the multivariate pattern of significant clusters (data item size greater than 5) in the multivariate attribute space, indicating the combined effects of multiple process drivers on the culvert sedimentation degree.

These significant (major) clusters reflect the dominant culvert sedimentation behavior within the landform region, thereby provide useful information for mitigating sedimentation depositions in that region. As an example shown in Figure 5.11c, the dominant two green clusters representing culverts with high sedimentation potential (70-100%) are quite uniform in many attributes (e.g., relatively low stream-to-culvert ratio, high agriculture land, and low forestry) that are also considered as the major contributors to sediment deposition in this region. By contrary, the less significant clusters (i.e., cluster containing only 1 or 2 data items) represent culverts with untypical sedimentation behavior, which are further discussed in Section 5.2.5. Consequently, these culverts are associated with deviations from the general nature of the watershed (e.g., urban culverts with no upstream agriculture land) or the presence of special structural modification at the culvert (e.g., curtain wall, downstream weir, and debris control structures).

Additionally, the visual interface allows users to acquire data-driven insights by comparing multivariate patterns of different culvert clusters, as for examples, the patterns shown in Figures 5.11c and 5.11d. Moreover, visualization such as the edge-bundling PCP in Figure 5.11e allows the platform users to compare the multivariate patterns within the two culvert clusters (color-coded in orange and green) of different sedimentation potential (i.e., very high vs medium potential). This comparison reveals that when some of the drivers (i.e., Ag land, forestry, grass, and sinuosity) are overlapping significantly as shown in Figure 5.11e, the green culvert cluster characterized by a lower stream-culvert width ratio are prone to high sediment deposition. Based on this comparison, the culvert designer is notified by the platform that a stream-culvert width ratio that is larger than 0.15 should be used to reduce the potential for sedimentation.

Sample site inspections are provided below to validate and substantiate the inferences garnered from the multivariate clustering analysis. For this purpose, Table 5.4 summarizes inspection photos collected at several culverts in the Northwest Iowa Plains region. The photos taken in the field confirm some of the landscape features of this landform region, such as high agricultural land coverage in the culverts surroundings and the relatively straight channel (low sinuosity) approaching the culvert. These features are also reflected in the PCP and SOM analysis. Most of the sediment deposits in this area consist of fine-grained particles (detailed by site: 273041) and are covered with thick vegetation, implying the fertility of the deposited soil. The presentation of the visual interfaces in this section illustrates how the platform users can directly interact with the SOM outputs and how they can further explore more data-driven insights and driver-specific solutions through the human-computer interaction. The number of possible stand-alone or combination of questions that can be addressed is numerous, but, as the previously detailed queries demonstrate, they are easy to pose. The platform return to the queries is immediate, selfexplanatory, and insightful. Given that the interactive exploration is so abundantly and explicitly visualized by the multivariate clustering interface, in the following sections, the number of queries will be limited. A more in-depth comparison of the representative clusters for each region is presented in Appendix A.

Table 5.4 Field photos of sediment deposit at culverts in the Northwest Iowa Plains region

Structure code	**Culvert inlet**	**Culvert outlet**	**Upstream channel**
26160 (high sediment)			
	Comment: located in agricultural watersheds, low sinuosity		
48511 (high sediment)			
	Comment: located in agricultural watersheds, low sinuosity		
273041 (high sediment)			
187150 (medium - high sediment)	Culvert inlet		Sediment material
			

5.2.4.2 Des Moines Lobe

The Des Moines Lobe covers most landscapes in northern and central Iowa, and was carved out by the Wisconsinan Glacier nearly 12,000-14,000 years ago (Prior, 1991). Most of the deposits underlying this region are made of glaciers materials, known as drifts and tills (IGS, 2017). As the glacier retreated, a poorly drained landscape underlain by pebbly clay as well as sand and gravel from swift meltwater streams was created (IGS, 2017; Prior, 1991). Other remarkable features of this region include large areas of broad and flat land (covering hundreds of square miles in southern Kossuth and Hancock counties) and wetlands (shown in Figure 5.12a). The region was subject to the extensive application of the tile drainage system (Crumpton et al., 2012). Due to these landscape features, the Des Moines exhibits a relatively low soil erodibility across the entire region, a feature also validated by the WEPP and RUSLE modeling outcomes.

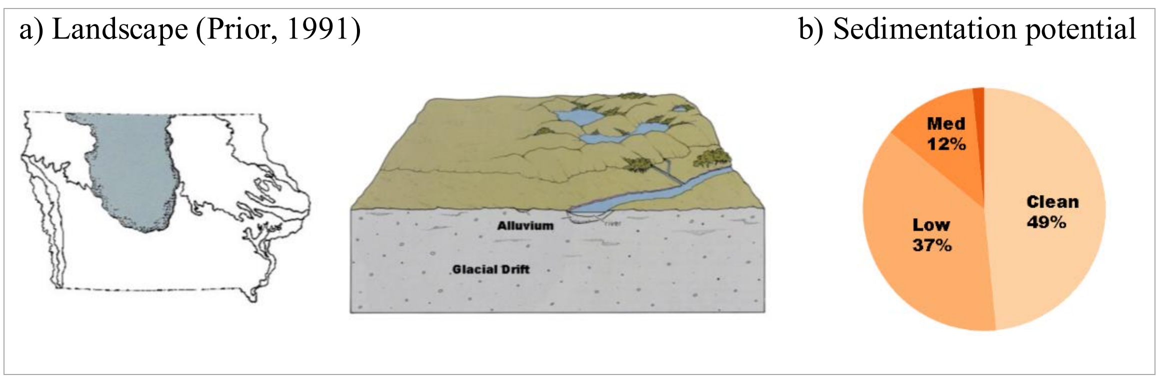
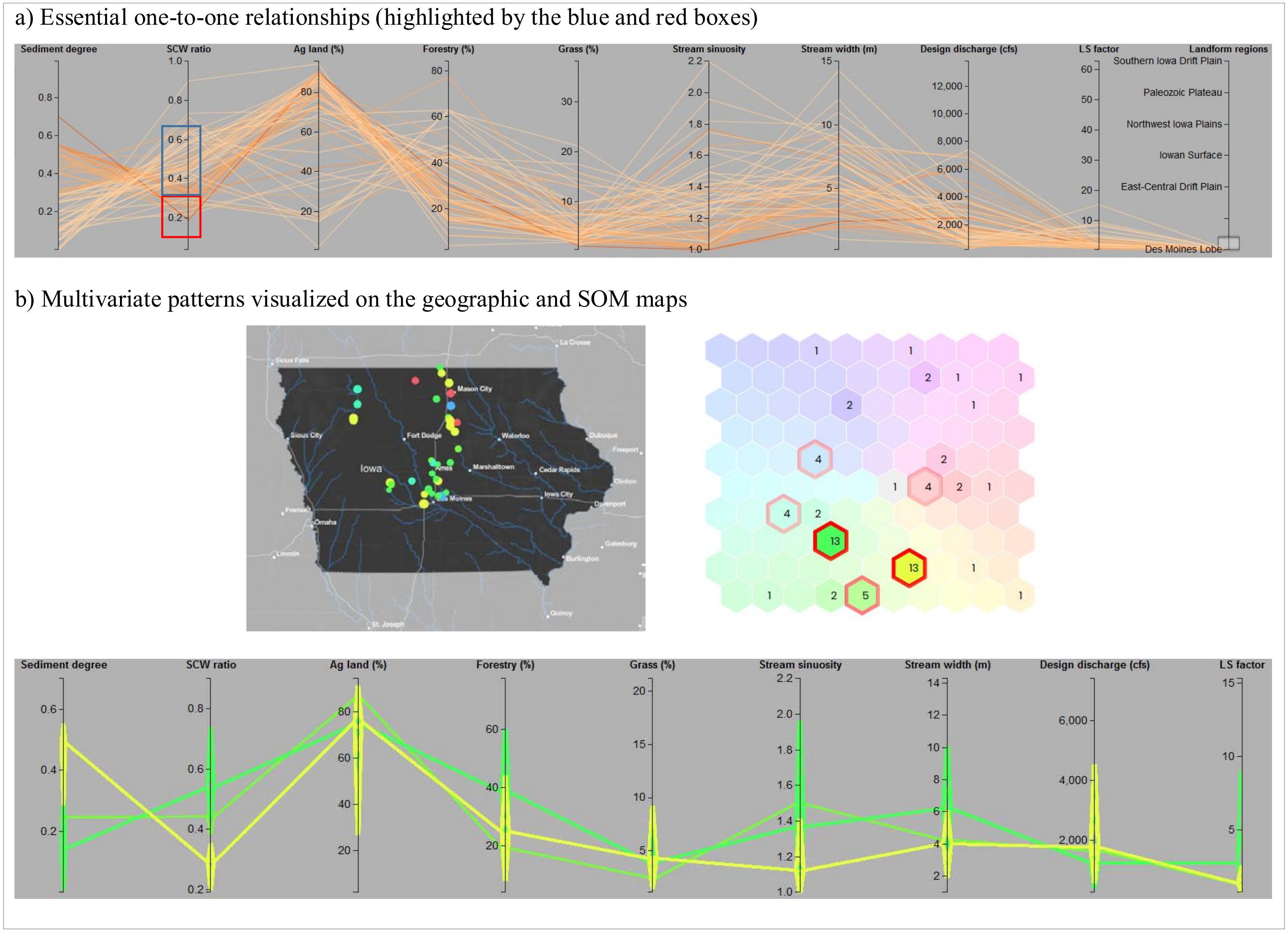


Figure 5.12 Des Moines Lobe Region

To further explore how the dominant drivers and their combinations affect the culvert sedimentation degree in the Des Moines Lobe region, the multivariate analysis used for the Northwest Iowa Plains region is applied to the culvert sample available in this region. The outputs of the analysis are presented collectively in Figure 5.13. Figure 5.13a reiterates that the SCW ratio is the most sensitive independent variable to the culvert sedimentation degree. The comparison between culvert clusters with low and medium-low sedimentation potential (green and yellow culverts in Figure 5.13c) reveals that the SCW ratio is the most critical factor in determining the culvert sedimentation degree when the other drivers display large overlapping in the edge-bundling PCP.

Several useful field observations acquired in the Des Moines Lobe are provided in Table 5.5 for confirming and substantiating the outcomes of the multivariate cluster analysis. It can be observed that, on average, the streams in this region are wider compared with streams in the Northwest Iowa Plains region (shown in Table 5.4) resulting in higher average SCW ratio for this region (as culvert width are closer to the stream width). This feature is also captured by the edge-bundling PCP provided in Figure 5.10a. The sediment deposits displayed in Table 5.5 (see for example site: 041101) appear to consist of coarser particles (sand and clay), which is different from the fine-material observed in the Northwest Iowa Plains region (see for example site 187150 in Table 5.4).


a) Essential one-to-one relationships (highlighted by the blue and red boxes)
Sediment degree
SCW ratio
Ag land (%)
Forestry (%)
Grass (%)
Stream sinuosity
Stream width (m)
Design discharge (cfs)
LS factor
Landform regions
Southern Iowa Drift Plain
Paleozoic Plateau
Northwest Iowa Plains
Iowan Surface
East-Central Drift Plain
Des Moines Lobe
b) Multivariate patterns visualized on the geographic and SOM maps
Iowa
Mason City
Fort Dodge
Waterloo
Sioux City
Ames
Marshalltown
Cedar Rapids
Iowa City
Dubuque
Davenport
Omaha
Burlington
Sediment degree
SCW ratio
Ag land (%)
Forestry (%)
Grass (%)
Stream sinuosity
Stream width (m)
Design discharge (cfs)
LS factor

c) Culvert clusters with low and medium-low sedimentation potential

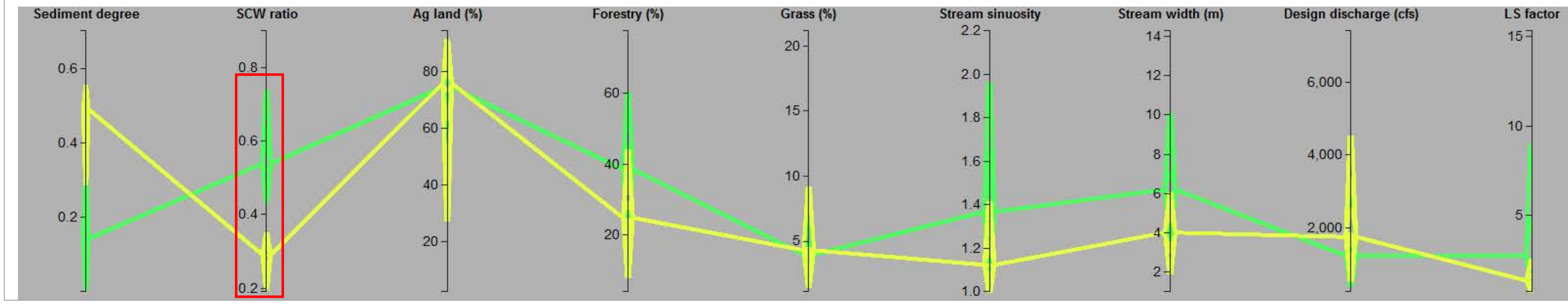


Figure 5.13 Relationships between the culvert sedimentation degree and drivers within the Des Moines Lobe

Table 5.5 Field photos of sediment deposit at culverts in the Des Moines Lobe

Structure code	Culvert inlet	Culvert outlet	Upstream channel
161941 (low sediment)			
157831 (low sediment)			
316301 (low sediment)			
041101 (medium sediment)	Culvert inlet		Sediment material
			
	Comment: coarse-grained sediment (the mixture of sand and clay)		

5.2.4.3 Iowan Surface

The Iowan Surface covers a large portion of areas in northeast Iowa and was formed during Wisconsinan glaciation events 16,500-21,000 years ago (IGS, 2017; Prior, 1991). The region is characterized by gently rolling terrain (illustrated in Figure 5.14a), which reflects various episodes of vigorous weathering and soil development, erosion, and loess deposition during the last period of intense glacial cold (IGS, 2017). In terms of landscape features, the Iowan Surface region has a thin discontinuous loess or loam covering the glacial drift and well-established drainage network across the region.

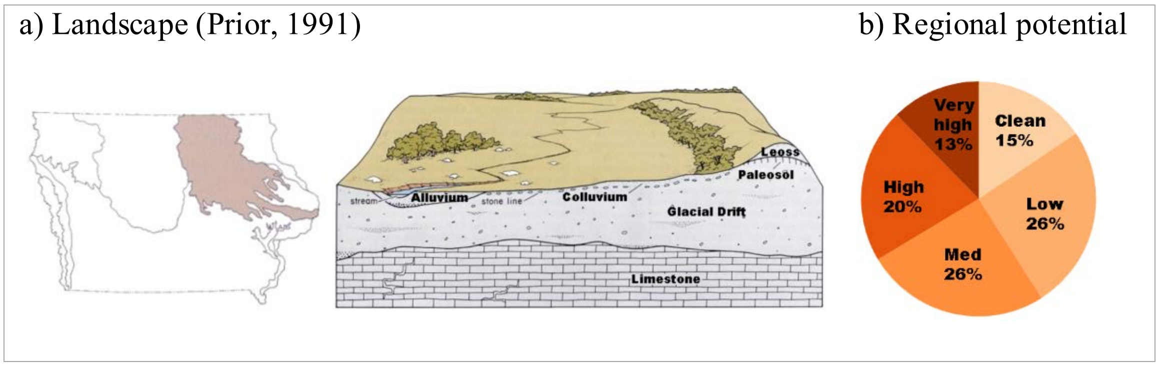
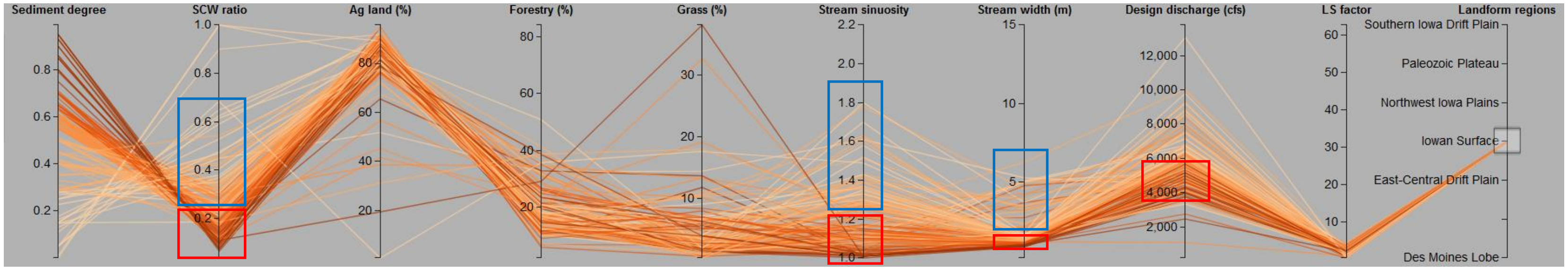


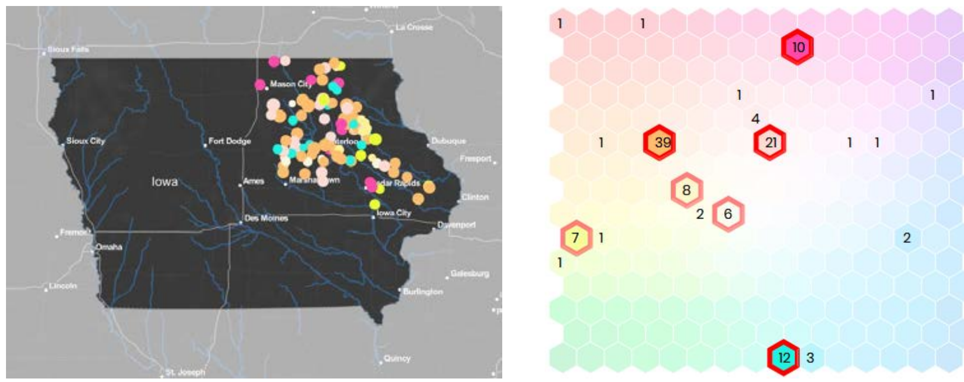
Figure 5.14 Iowan Surface Region

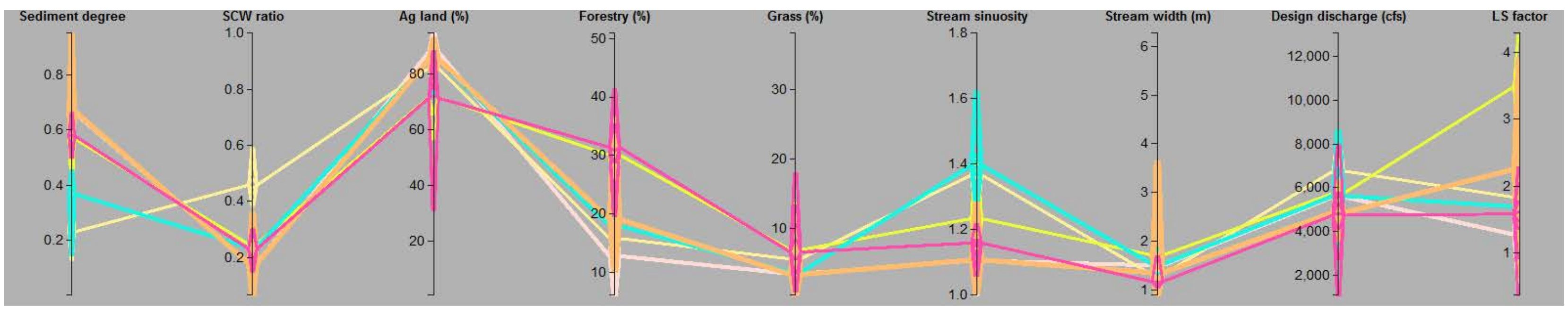
According to IGS (2017), the Iowan Surface region and Northwest Iowa Plains region have a similar erosional history, therefore, it is no wonder that, during the early regionalization stage, many culverts in the Iowan Surface region and the Northwest Iowa Plains region are grouped into the same cluster by the multivariate clustering analysis (as presented in Section 3.5.3). The culvert sedimentation potential is mixed in this region (shown as Figure 5.14b), therefore additional analysis is required to identify the dominant drivers and their combinations that are responsible for the mix in this region. The outputs of the additional analysis are presented collectively in Figure 5.15.

a) Critical one-to-one relationships (highlighted by the blue and red boxes)



b) Overview of the multivariate patterns through the geographic map and SOM map





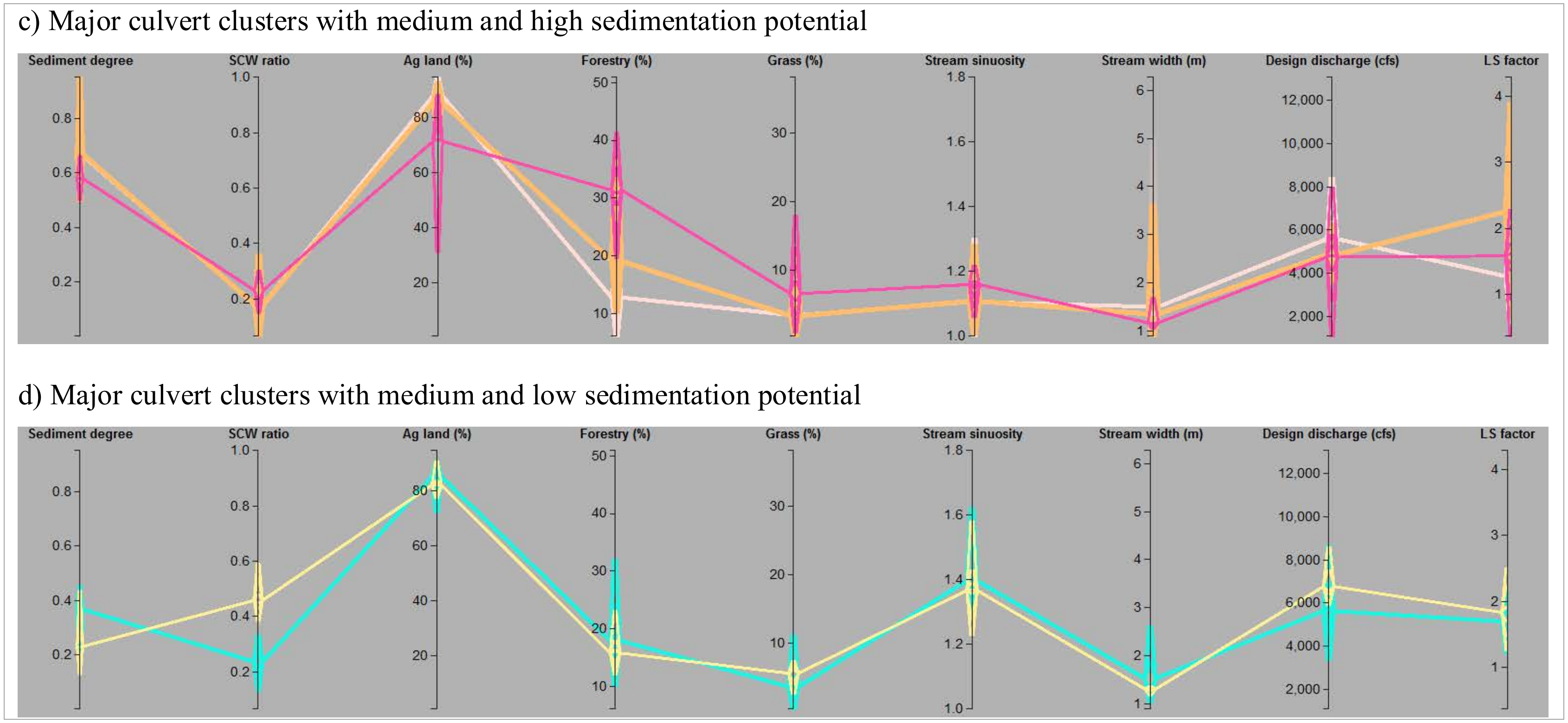
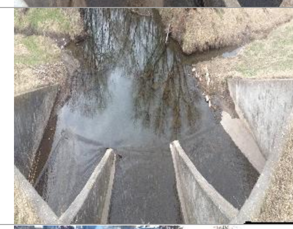


Figure 5.15 Relationships between the culvert sedimentation degree and drivers within the Iowan Surface Region

The field observations also confirm the results of the analysis, showing that the Iowan Surface region has a significant number of culverts with either low or high sediment deposition. Samples of these observations are shown in Table 5.6. In terms of the sediment material, site inspection photos show that sediment depositions at many culverts in this region are composed of fine-grained materials such as silt (site 220371 and site 75671 in Table 5.6). It is worth mentioning that, through the field inspection, the present study detected several culverts in the Iowan Surface region have local structure modifications that help reduce sediment deposition near the culvert structure, which are further discussed in Section 5.2.5.

Table 5.6 Field observations of sediment depositions at culverts in the Iowan Surface

Structure code	**Culvert inlet**	**Culvert outlet**	**Upstream channel**
700290 (low sediment)			
23241 (low sediment)			
165881 (high sediment)			
Structure code	Culvert inlet	Culvert outlet	Sediment material
220371 (high sediment)			
	Comment: fine-grained sediment		
75671 (low sediment)			

5.2.4.4 Southern Iowa Drift Plain

Located in the southern part of Iowa, the Southern Iowa Drift Plain is certainly the largest of Iowa's landform regions (IGS, 2017; Prior, 1991). The region is composed of weathered glacial drift (gravel, sand, or clay that is picked up and deposited by the glacier as it moves) covered with a windblown mantle of loess, as illustrated in Figure 5.16a (Prior, 1991). In terms of landscape features, the region is characterized by steeply rolling hills and valleys, and has integrated drainage networks that consist numerous rills, creeks, and rivers branch out across the region (IGS, 2017). The southern part of the region (central irregular plains) is covered with wooded terrain.

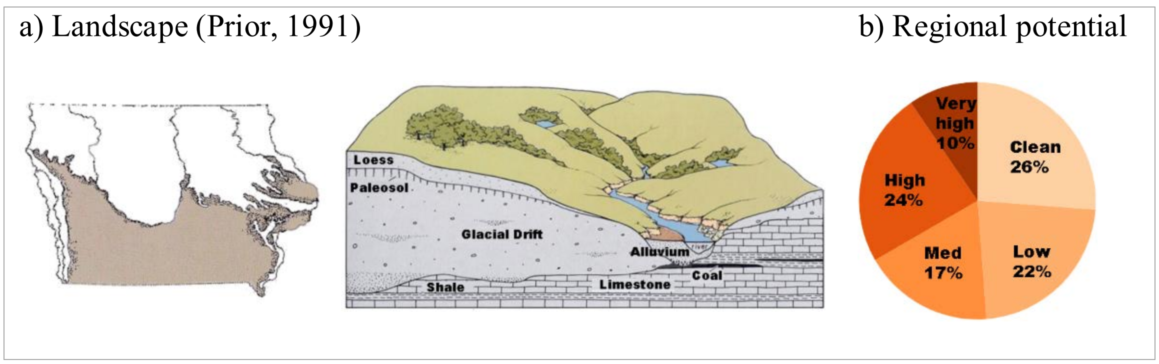
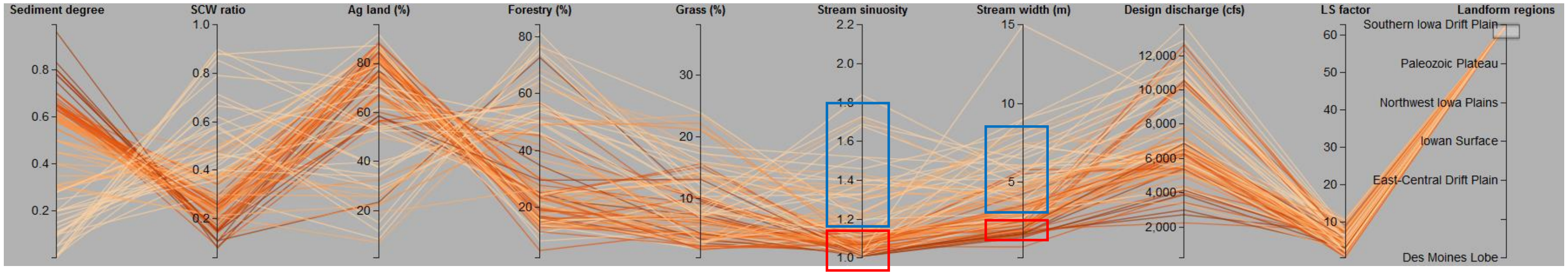


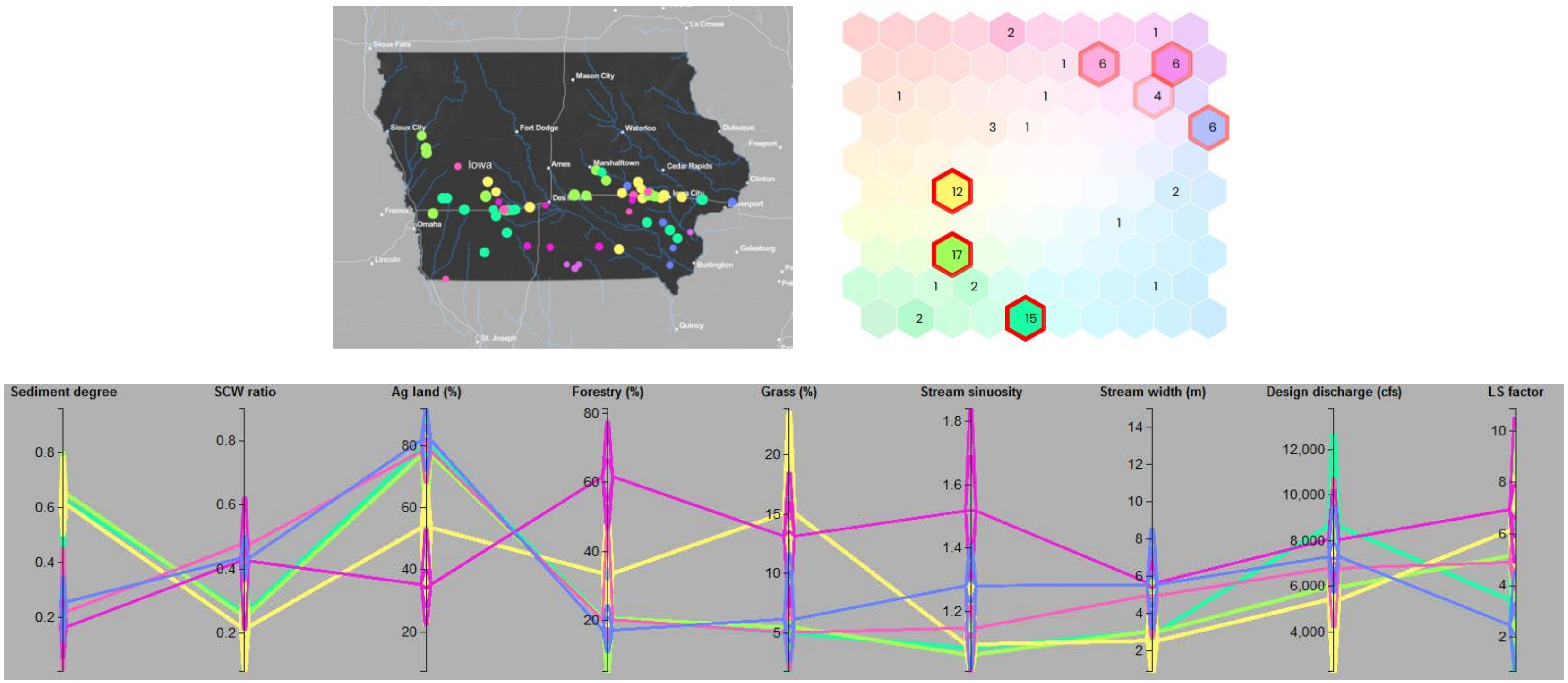
Figure 5.16 Southern Iowa Drift Plain

Similar to the Iowan Surface region, the Southern Iowa Drift Plain has a mixed culvert sedimentation potential (shown as Figure 5.16b). The outputs of the additional analysis are presented collectively in Figure 5.17.

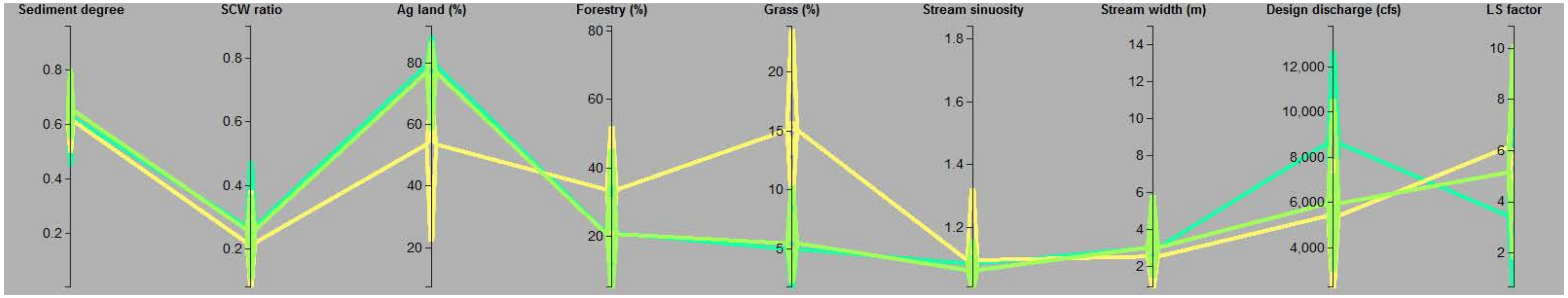
a) One-to-one relationships (highlighted by the blue and red boxes)



b) Multivariate patterns (overall trends)



c) Multivariate patterns of clusters with medium-high culvert sedimentation potential



d) Multivariate patterns of clusters with medium-low culvert sedimentation potential

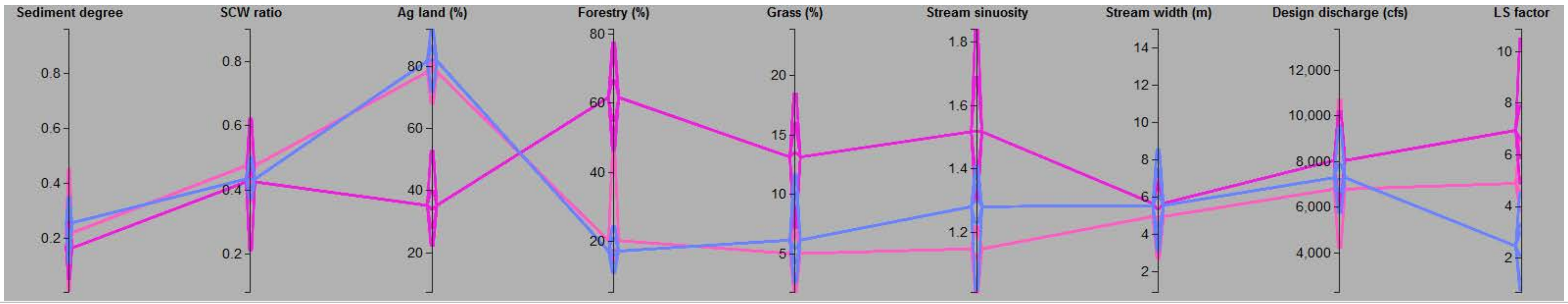


Figure 5.17 Relationships between the culvert sedimentation degree and drivers within the Southern Iowa Drift Plain

Field surveys also prove that the degrees of culvert sedimentation are mixed in the Southern Iowa Drift Plain region. Sample inspection photos taken at different culvert locations in this region are provided in Table 5.7. Through these inspections, it is discovered that many watersheds in the Southern Iowa Drift Plain region are extensively covered with grass and forestry, and culverts within these watersheds trend to have relatively low sedimentation degrees (site 64311 and 19992). This observation validates the multivariate relationship presented in Figure 5.17d that indicates that the forestry in culvert drainage basin can reduce the degree of culvert sedimentation in the Southern Iowa Drift Plain region. Another important finding from the field inspection is that culverts located in the western areas in the region are more likely to be blocked by sediments. This situation can be explained as the western parts of the region act as the recipient of the windblown loess soil that is transported from the neighboring Northwest Iowa Plains and loess hill.

Table 5.7 Field observations of sediment depositions at culverts in the Southern Iowa Drift Plain region

Structure code	Culvert inlet	Culvert outlet	Upstream channel
64311 (low sediment)			
	Comment: located in a forested watershed		
19992 (low sediment)			
	Comment: located in a forested watershed		
503870 (low sediment)			
	Comment: located in a urban watershed		
271840 (high sediment)			
	Comment: located in the western parts of the region		
47960 (high sediment)			
	Comment: located in the western parts of the region		

5.2.4.5 Paleozoic Plateau

The Paleozoic Plateau is recognized as the most distinct of the landform regions in Iowa, as its landscapes are characterized with many deep, narrow valleys containing cool, fast-flowing streams, and abundant woodlands (as are illustrated in Figure 5.18a) (IGS, 2017; Prior, 1991). This spectacular high-relief landscape is located in the northeast corner of Iowa, and is formed from erosions through rock strata of Paleozoic age (Prior, 1991). In this region, stream erosion and hillslope development are magnified and complicated by the bedrock-controlled relief and steep slopes (IGS, 2017). Trimble (2013) studied the sediment budget across the Paleozoic Plateau and observed that many areas (e.g., upper main valley and lower main valley) in this region are prone to severe aggradation and channel and bank erosion. This observation further confirms the high culvert sedimentation potential discovered in this study.

a) Landscape (Prior, 1991) b) Regional potential

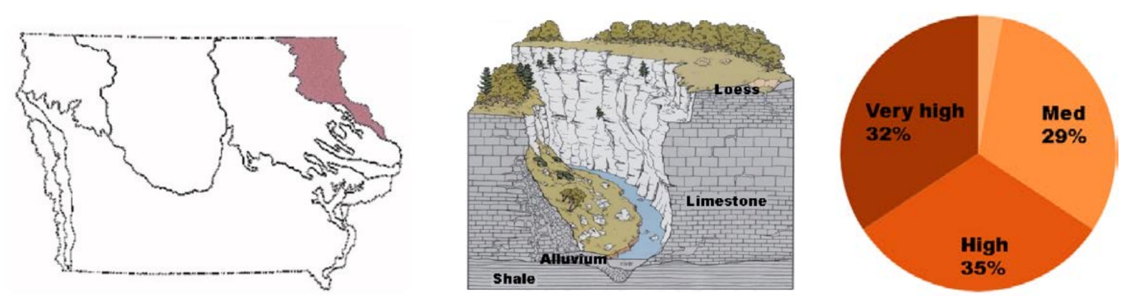
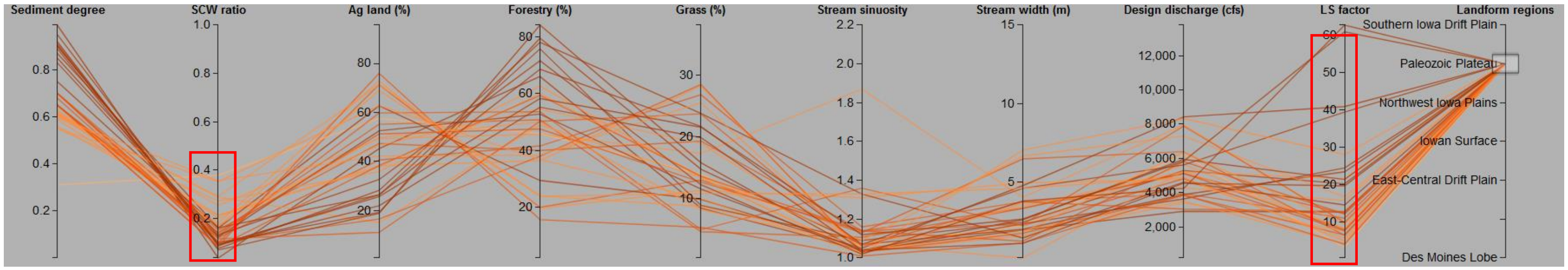


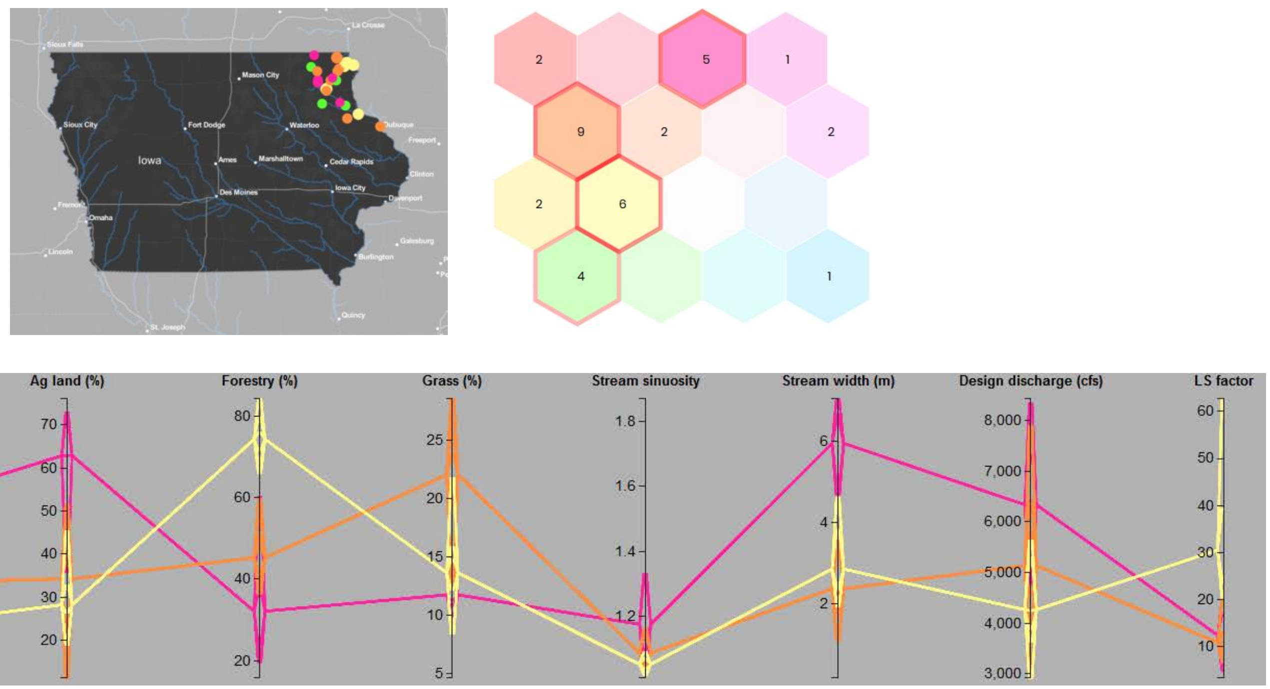
Figure 5.18 Paleozoic Plateau

Culverts in the Paleozoic Plateau disclose a uniformly medium to high sedimentation potential (shown in Figure 5.18b). Figure 5.19 demonstrates the outcomes of the multivariate analysis.

a) One-to-one relationships (highlighted by the blue and red boxes)



b) Multivariate patterns (overall trends)



c) Multivariate pattern comparisons between clusters with medium and high sedimentation potentials)

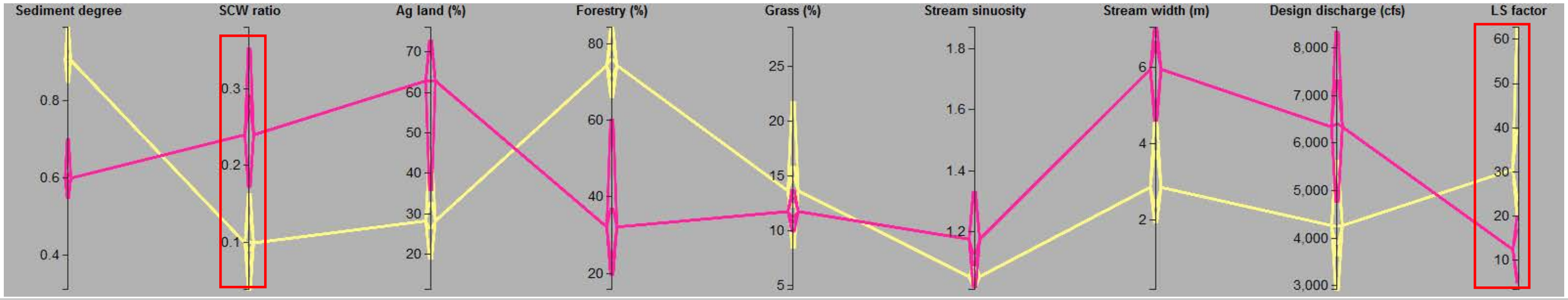
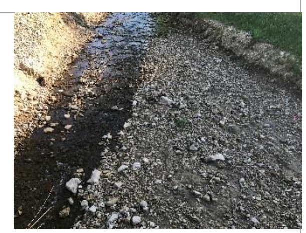


Figure 5.19 Relationships between the culvert sedimentation degree and drivers within the Paleozoic Plateau

Through a series of culvert inspections, the study discovers that culvert sediment deposits in this landform region are dominated by both rocks (gravels and cobbles) and fine-grained sediments, as illustrated in Table 5.8. Due to the bedrock-dominated terrain in this region (Prior, 1991), many stream channels in this region are semi-controlled (Schumm, 1985), displaying a very different sediment response compared with alluvial streams in other landform regions.

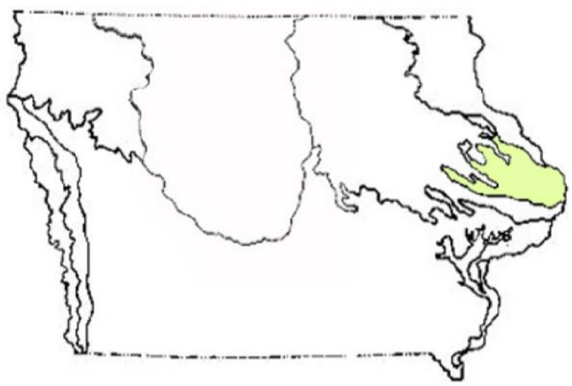
Table 5.8 Field observations of sediment depositions at culverts in the Paleozoic Plateau

Structure code	Culvert inlet	Culvert outlet	Upstream channel
13511 (high sediment)			
	Comment: gravels and cobbles, steep slope		
61921 (medium-high sediment)			
	Comment: gravels and cobbles, steep slope		
62115 (medium-high sediment)			
	Comment: gravels and cobbles, steep slope		
121150 (high sediment)			
	Comment: fine-grained sediment, mild slope		

5.2.4.6 East-Central Iowa Drift Plain

The East-Central Iowa Drift plain was a sub-region renamed from the Southern Iowa Drift plain, as its geographic location is apart from the rest of the drift plain (Prior and Kohrt, 2006). Influenced by the nearby Iowan Surface region to the south-west and the Paleozoic Plateau to the north, the terrain of the East-Central Drift Plain is partitioned by deeply-cut streams (IGS, 2017).

a) Landscape (Prior, 1991)



b) Regional potential

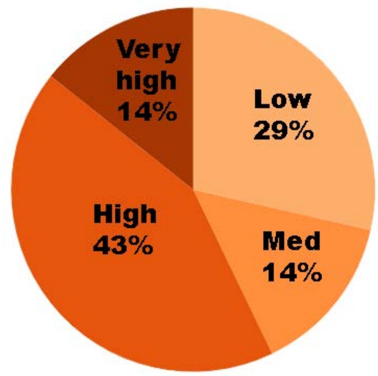


Figure 5.20 The East-Central Iowa Drift plain

The East-Central Drift Plain is a relatively small region (shown in Figure 5.20a) with only seven culverts having adequate information for the data-driven analysis. It appears that this region has the majority of its culvert with the “high” and “very high” sedimentation potential (shown in Figure 5.20b). Due to its small size and limited culvert sample, the derived relationships that are shown Figure 5.21 for the region are considered not significant and merged with the Southern Iowa Drift plain for further analysis.

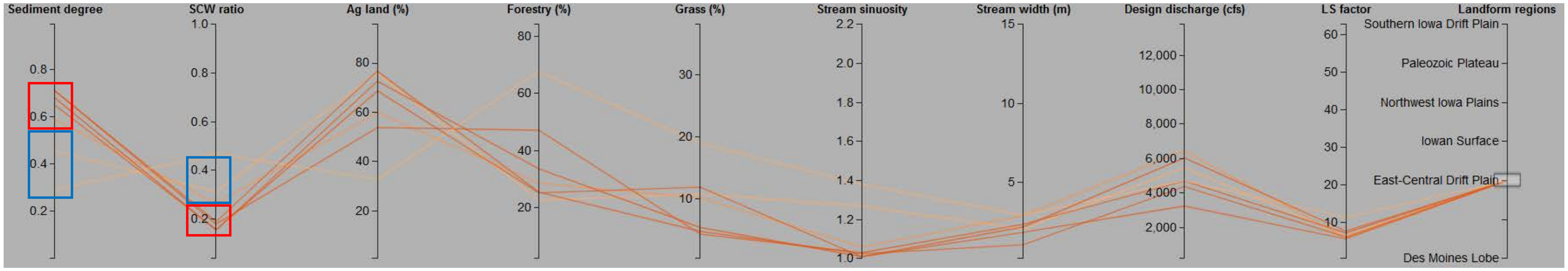


Figure 5.21 Relationships between the culvert sedimentation degree and drivers within the East-Central Iowa Drift Plain

5.2.5 Data-driven inferences

Using the outcomes from the spatial extent, regional, and multivariate clustering analyses, this section summarizes the data-driven insights derived in different landform regions. There are two types of useful insights on the type of clusters ensuing from the SOM maps, namely: (1) major clusters (i.e., highly relevant clusters that assemble sample population greater than 5 culverts), and (2) minor clusters (i.e., exceptions of clusters with only 1 or 2 data items). Samples of clusters of the two types are visualized in Figures 5.22a and 5.22b, respectively. The pertinent inferences for practical purposes will be discussed around these two categories of clusters.

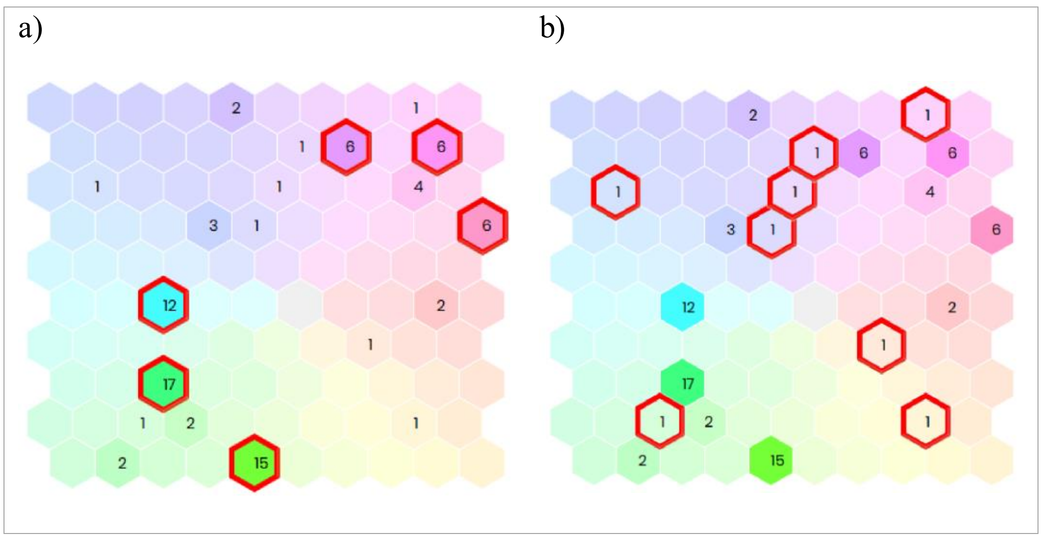


Figure 5.22 Sample SOM maps: (a) major clusters, (b) minor clusters (exceptions)

5.2.5.1 Data-driven Insights from Major Clusters

The protocols to determine the major clusters and their specific multivariate patterns for individual landform regions are discussed in the previous subsections. Similar to the decision tree associated with the MCDA, the outcomes from the regionalization and SOM can be readily converted into quantitative criteria to provide specific suggestions for culvert sedimentation mitigation in different landform regions. Appendix A summarizes these trends for each region and provides design suggestions stemming from the derived data-driven insights to assist in the management of existing culverts and the design of new culverts.

5.2.5.2 Data-driven Insights from Minor Clusters

During this study, the multivariate clustering analysis identifies a number of culverts that have distinct patterns compared with the rest of the culverts in the same landform region. These culverts are grouped either by themselves or with a few other sites into minor clusters than do not from samples larger than 4 items in the sample. Although discrete clusters are normally considered as outliners in many data-driven studies, the present study finds that many of the culverts in these minor clusters are actually sites with customized structural features (modification on structures or channels) or unique environmental attributes (special stream locations and conditions). Figure 5.23 illustrates two such examples. This study also discovers that, in the minor clusters, culverts without the presence of additional modification usually have extreme values in one or several of its drivers. Figure 5.24 provides few examples of such culvert sites. Although culverts in these minor clusters do not reflect the major trends of sediment response in a landform region, their unique features do provide valuable insights that help infrastructure managers develop effective strategies for either mitigating sedimentation at existing culverts or inhibiting sediment accumulation at new culverts. Through the minor clusters produced from the SOM, the present study identifies several unique features at culvert locations that can help reduce sedimentation risks at culverts. These features are summarized in Appendix B.

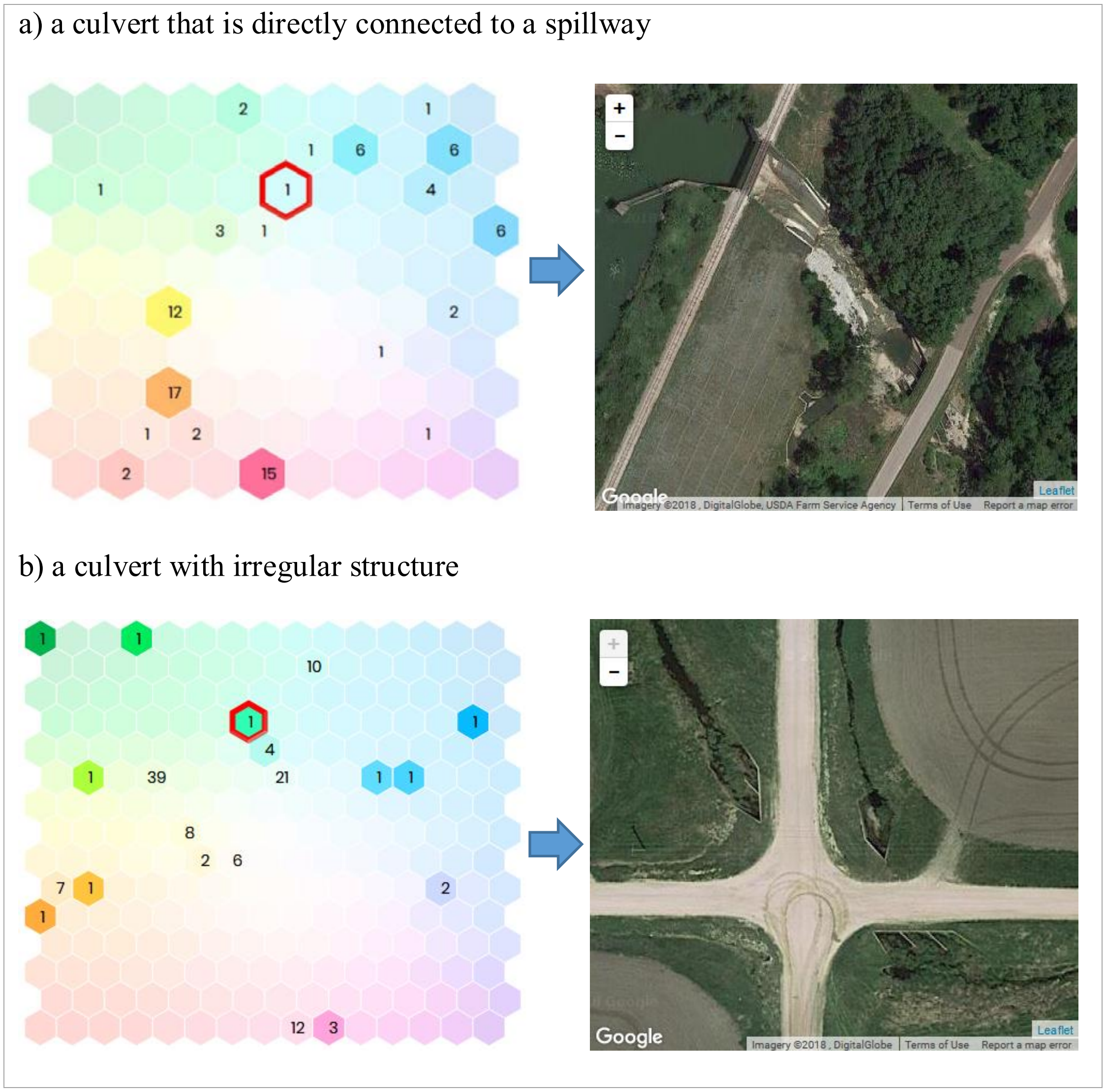
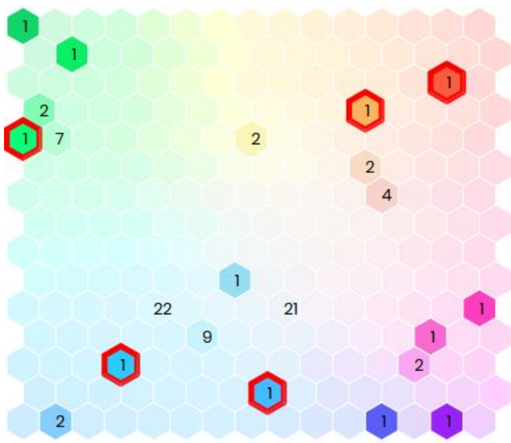
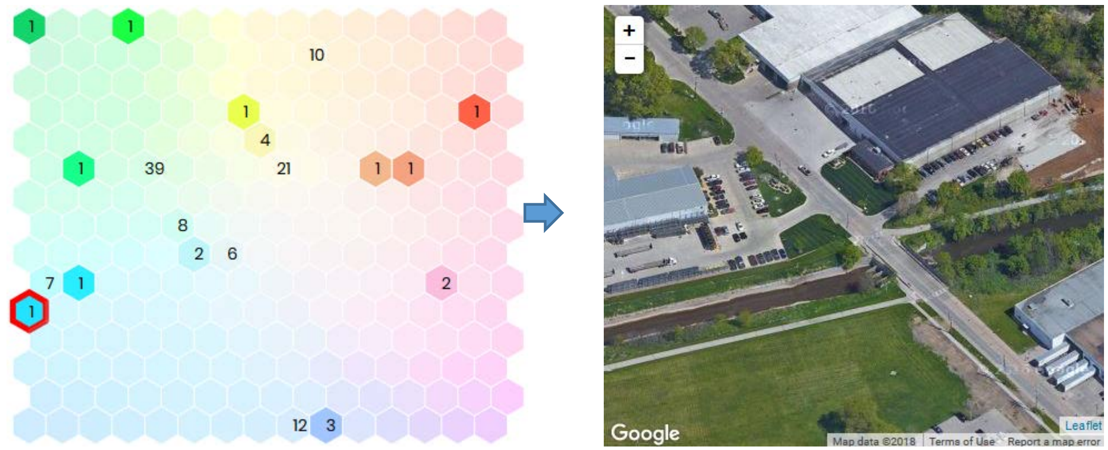


Figure 5.23 Demonstration of special culvert cases

a) minor clusters in the SOM map



c) a clean culvert with an extreme value in its drivers (no agricultural land use in its drainage basin)



a) extreme values of process-drivers contained in minor clusters

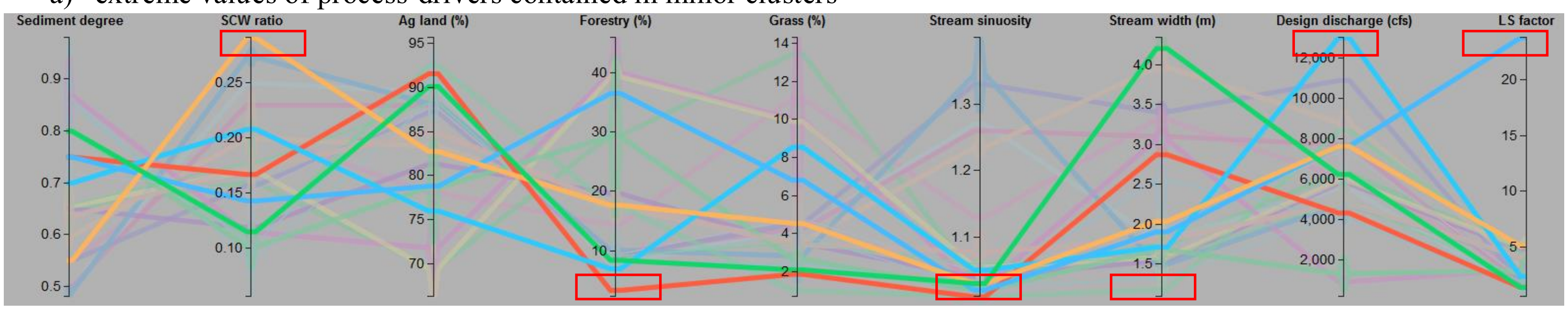


Figure 5.24 Demonstration of culverts with extreme values in the environmental drivers

5.3 References

Abbott, M. B. (1991). "Hydroinformatics: information technology and the aquatic environment". Ashgate Publishing, LimiteAllan, J. D. (2004). Landscapes and Riverscapes: The Influence of Land Use on Stream Ecosystems. Annual Review of Ecology, Evolution, and Systematics. *35* (1): 257-284.

Agafonkin, V. (2010). "An open-source JavaScript library for mobile-friendly interactive maps". Retrieved on November 7, 2017 from: leafletjs: http://leafletjs.com.

Blanco-Canqui H. and Lal, R. (2010). "Soil erosion under forests". In: Principles of Soil Conservation and Management. Springer, Dordrecht.

Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., and Grisel, O., et al. (2013). "API design for machine learning software: experiences from the scikit-learn project". arXiv Preprint arXiv:1309.0238.

Crumpton, W., Helmers, Matthew, J., Stenback, G., Lemke, D., and Richmond, S. (2012). "Integrated drainage wetland systems for reducing nitrate loads from Des Moines lobe watersheds", Agricultural and Biosystems Engineering Technical Reports and White Papers. 13.

Devijver, P. A. and Kittler, J (1982). "Pattern recognition: a statistical approach". London, GB: Prentice-Hall.

Eash, D. (2001). "Techniques for estimating flood-frequency discharges for streams in Iowa", Water-Resources Investigations Report 00-4233, U.S. Geological Survey, Reston, VA.

ESRI (2015). "ArcGIS for server". Retrieved on November 7, 2017. Retrieved on May 23, 2015 from: server.arcgis.com/en/server

Geisser, S. (1993). "Predictive inference". New York, NY: Chapman and Hall. ISBN 0-412-03471-9.

Iowa Geological Survey (2017). “Landform regions of Iowa”, Retrieved on Nov 27, 2018 from https://www.iihr.uiowa.edu/igs/publications/uploads/2017-04-27_15-04-11_em44.pdf.

Jahantigh, M. and Pessarakli, M. (2011) “Causes and effects of gully erosion on agricultural lands and the environment”, Communications in Soil Science and Plant Analysis. 42 (18): 2250-2255, DOI: 10.1080/00103624.2011.602456

Kiang, M. Y., Hu, M. Y., Fisher, D. M., and Chi, R. T. (2005). “The effect of sample size on the extended self-organizing map network for market segmentation”, Proceedings of the 38th Annual Hawaii International Conference on System Sciences. 09, January 03 - 06, 2005.

Kellerhals, R., Church, M., and Bray, D.I. (1976). “Classification and analysis of river processes”, Journal of the Hydraulics Division, American Society of Civil Engineers 102: 813–829.

McLachlan, G. J., Do, K. A.,and Ambroise, C. (2004). “Analyzing microarray gene expression data”. Wiley.

Montgomery, D. (2007). “Soil erosion and agricultural sustainability”, Proceedings of the National Academy of Sciences of the United States of America (PNAS). 104 (33):13268-13272.

Muste, M., Ettema, R., Ho, H-C. and Miyawaki, S. (2009). “Development of self-cleaning box culvert design”. Report for IHRB TR-545, Iowa Department of Transportation, 800 Lincoln Way, Ames, Iowa.

Mustajoki, J. and Marttunen, M. (2017). “Comparison of multi-criteria decision analytical software for supporting environmental planning processes”. Environmental Modeling and Software. 93:78–91.

MDN (2015). “AJAX”. Retrieved fon June 3, 2018 rom Mozilla Developer Network: https://developer.mozilla.org/en-US/docs/AJAX.

Microsoft (2014). “Bing maps API”. Retrieved from on November 7, 2018: https://www.microsoft.com/maps/choose-your-bing-maps-API.aspx.

National Soil Erosion Research Laboratory (NSERA), USDA. (1995). “USDA-Water Erosion Prediction Project (WEPP) user summary”. NSERL Report N0.11: 1-126.

OGC (2009). “Web map service”. Open Geospatial Consortium.

OGC (2011a). “OpenGIS Web Map Tile Service Implementation Standard”. Retrieved on June 2, 2011 from: www.opengeospatial.org/standards/wmts.

OGC (2011b) “Web feature service”, Retrieved on November 1, 2018 from: opengeospatial.org/standards/wfs.

Piest, R. F. and Ziernnicki, S. (1979). “Comparative erosion rates of loess soils”, the American Society of Agricultural Engineers, St. Joseph, Michigan.

Project, T. J. (2010). “jQuery: The write less, do more, JavaScript library”. Retrieved on April 29, 2018 from jQuery: https://jquery.com/

Prior, J. (1991). “Landforms of Iowa”: University of Iowa Press, Iowa City, IA, USA.

Prior, J. C., and Korht, C.J. (2006), “The landform regions of Iowa, Iowa Geological Survey”, digital map, available on IDNR GIS Library: ftp://ftp.igsb.uiowa.edu/gis_library/ia_state/geologic/landform/landform_regions.zip; http://www.igsb.uiowa.edu/nrgislibx/

Renard, K. G., Foster, G. R., Weesies, G. A., McCool, D. K., Yoder, D. C. (1997). “Predicting soil erosion by water: a guide to conservation planning with the Revised Universal Soil Loss Equation (RUSLE)”. USDA Agr. Handb.703: 1-225.

Schumm, S. A. (1985). “Patterns of alluvial rivers”, Department of Earth Resources, Colorado State University, Fort Collins, Colorado 80523.

Trimble, S. W. (2013). “The distributed sediment budget model and watershed management in the Paleozoic plateau of the upper Midwestern United States”, Physical Geography. 14 (3).

Xu, H. (2015). “Prototyping Hydroinformatics-based systems for supporting decision-making in culvert design and monitoring”, M.S. dissertation, The University of Iowa, Iowa City, IA.

Xu, H. Shen, B. and Muste, M. (2015). “Geo-portal for sustainable culvert design and monitoring”, e-proceedings of the 36th IAHR world congress, 28 June - 3 July, 2015, Hague, The Netherlands.

Zaimes, G. N. and Schultz, R. C. (2012). “Assessing riparian conservation land management practice impacts on gully erosion in Iowa”, Environmental Management 49: 1009. https://doi.org/10.1007/s00267-012-9830-9

CHAPTER 6 CONCLUSIONS AND FUTURE WORK

6.1 Summary

Sedimentation at multi-barrel culverts has become a widespread problem in the US Midwest as a direct consequence of the changes in landscape produced by human intervention as well as climate change (arguably also driven by human activity in rural and urban settings). The current design for culverts is driven by analyses of the hydrologic and hydraulic conditions at the construction site. Less attention is typically given to the assessment of the potential for sedimentation adjacent to and within the culvert, a process that is triggered by a myriad of complex and interlinked geo-morphological processes. This is unfortunate as sediment cleaning at culvert is expensive, labor intensive, and sometimes need to be repeated over time.

Given that it is difficult to investigate and solve the sedimentation problem with conventional approaches (i.e., laboratory-based experimental methods, and physics-based erosion and sediment transport models), currently, there is no sufficient knowledge to develop systematic strategies or formulate suggestions for preventing and mitigating the detrimental effects of culvert sedimentation. This research aims to fill this problem-solving gap by using a data-driven approach for identifying key-process drivers and link them in quantitative relationships that eventually can aid the conceptualization of the culvert sedimentation process. The targeted relationships for this study are between the degree of sedimentation at the culvert site and the independent variables characterizing their drainage area along with culvert structural parameters related to erosion and sediment transport.

The study is carried out in a progressive manner by gradually increasing the complexity of the data-driven methods to, eventually, enable the extraction of relevant information that can improve the capabilities of the culvert designers and operators to cope with the problem of sedimentation at culverts. In the early stage of the study, the exploration of the multivariable relationships is made through the use of Parallel Coordinate Plots (PCP) and a coarse Multi-Criteria Decision Analysis (MCDA) applied to a dataset describing the sedimentation status at 309 culverts across Iowa. PCP is a contemporary multivariate data visualization technique offering opportunities to improve one's qualitative comprehension of the system processes encoded within highly dimensional datasets. The MCDA utilizes a decision tree-based technique, a popular method in data-mining classification and prediction applications, to quantify the empirical relationships between erosion and sediment transport drivers and the culvert sedimentation degree.

In a subsequent stage, the MCDA is applied to a sample population of 400 culverts with the addition of aerial images of culverts located in areas not well covered by the original dataset. Using the such-obtained more uniformly distributed Iowa culvert dataset, a series of optimization tasks is undertaken by regionalizing the multivariate analysis. The optimization phase includes data exploration techniques aimed at: (a) establishing the effective spatial extent of each key-process driver, (b) testing the spatial variability of the culvert sedimentation potential across different landform regions in Iowa, and (c) identifying the patterns (relationships) between the culvert sedimentation and its drivers within each landform region. The optimization applied to the extended set of data has visibly improved the derivation of the relationships between drivers and the culvert sedimentation degree when the selected drivers were assessed at the stream and stream corridors scales (i.e., at the level of the stream floodplain). The SOM implementation clearly substantiates the existence of distinct erosion and sedimentation potential Iowa regions. Finally, the SOM applied on top of the regionalization sheds light on the importance of the local factors in ranking the sedimentation process drivers by identifying major and minor clusters of culvert sedimentation response within each region, hence offering valuable aids for developing mitigation strategies specific to each culvert location.

Given that the present study has an immediate practical implementation aspect besides the scientific exploration of sedimentation at culvert processes, the developed data-driven tools and functions are embedded in a customized web-based Problem Solving Environments (PSE) labeled the "IowaDOT Culverts" platform. This platform assembles in one place pre- and post-construction data and information, irrespective of their provenance. The portal enables four workflows: 1) storage and query of culvert specifications and ancillary information; 2) monitoring of sedimentation at culverts using in-situ or remote sensing technologies; 3) analysis of the sedimentation at culverts; and (4) support of culvert sedimentation-sensitive design that reduces future culvert operation hazards through intelligent-aided analysis. User-friendly portal interfaces allow users to prepare a systematic plan for culvert monitoring and offer means for quantitative assessment of the potential for sediment deposit formation. The workflows can be applied to existing or potential culvert sites, therefore assisting both operations and design purposes. The platform can be accessed anytime from anywhere via the internet, therefore, it can be updated as soon as data and information are produced. The availability of a continuous updated central repository can timely and readily inform on the state of the culverts in Iowa.

6.2 Conclusions

The research presented in this study illustrates through a practical hydro-sociological example how data-driven modeling improves our current understanding of the complex processes involved in culvert sedimentation as well as supports the optimization of the management and design of multi-box culverts with consideration of sediment accumulation in their vicinity. The outcomes of this research illustrate the role of data-driven modeling to complement, rather than compete with, the more robust physically-based modeling. The study emphasizes the place of this newer investigative approach within the continuum of modeling approaches revealing in the same time the extent to which hypotheses can be formulated to better inform the modeling development (Mount et al. 2016). The improved understanding stems from the capabilities of data-driven modeling to use an inductive approach in the formulation of research hypotheses. This aspect is critically important for an area where the robust and comprehensive conceptualization of the processes involved is difficult using alternative means (Mekonnen et al, 2015).

The application of machine learning and computationally intelligent algorithms to characterize and quantify the myriad of implicit structures and relationships embedded within complex, multivariate datasets is not however without challenges. The most prominent challenge is to secure the data-rich environment regarding the investigated processes. The abundance of the data is not sufficient to offer a pathway for formulating new understandings. It is essential that the data should be well-tailored to the pursued investigation and to be acquired with adequate spatial and temporal resolution. Using increased data-driven model complexity is not a substitute for poor, limited range, and unknown uncertainty of the datasets used for the analysis. Finally, it should be mentioned that elucidating relationships between natural and man-made hydrologic and hydraulic variables through data-driven knowledge discovery is not the same as validating them, as the outcomes of a data-driven investigation are just as good as the data inputted in the analysis.

Ideally, the MCDA utilized in this study identifies one-to-one relationships between individual variables (e.g., stream-to-culvert width ratio and discharge). However, there are instances when the method is prone to overfitting, i.e., situations when the same MCDA outcome is produced by one-to-one or combined effects of the drivers, limiting the validity of the method altogether. Lack of sufficient data and overfitting the training dataset are however downsides that can be avoided by using further improvements. This is done in the present work by enhancing the relevance and the amount of the input datasets and implementing optimization of the data-driven analysis consisting of regionalization, clustering, and sensitivity analysis of the independent variables prior to implementation of the MCDA. The synthesis of research findings following the optimization of the MCDA is provided in Tables 6.1 and 6.2. More detailed suggestions (as quantitative criteria) involved in the optimizations are provided in Appendix A.

Table 6.1 Synthesis of the MCDA and forecasting outcomes using regionalization

Regions	Sedimentation potential	Key drivers
Northwestern Iowa Plains	Very high	• stream-culvert width ratio • stream width • culvert design discharge
Des Moines Lobe	Low	• stream-culvert width ratio • stream sinuosity • stream width • LS factor
Iowan Surface	Mixed	• stream-culvert width ratio • stream sinuosity • forestry • design discharge • LS factor
Southern Iowa Drift Plain	Mixed	• stream-culvert width ratio • forestry • sinuosity • design discharge • stream width
Paleozoic Plateau	High	• stream-culvert width ratio • stream width • LS factor

Table 6.2 Data-driven practical insights for mitigating culvert sedimentation using regionalization

Regions	Design recommendations	Recommended practice
Northwestern Iowa Plains	Structural engineers should *avoid placing an over-sized channel expansion area* near culvert inlet by attempting to attain a stream-culvert width ratio above 0.15.	Use pipe culverts
Des Moines Lobe	• Structural engineers should *avoid placing an over-sized channel expansion area* near culvert inlet by attempting to attain a stream-culvert width ratio above 0.40. • Land managers should avoid channelization in the culvert upstream by attempting to attain a sinuosity larger than 1.4. *Re-meandering channelized and straightened streams* could be another suggestion for preventing culvert sedimentation.	• Retrofit bathymetry in the culvert area
Iowan Surface	• Land managers should avoid channelization in the culvert upstream by attempting to attain a sinuosity larger than 1.3. *Re-meandering channelized and straightened streams* could be another suggestion for preventing culvert sedimentation. • Culvert designers should, through land management practices, *increase the percentage of forestry* in culverts' stream corridor (larger than 20%) to reduce the risk of culvert sedimentation. • Structural engineers should *avoid placing an over-sized channel expansion area* near culvert inlet by attempting to attain a stream-culvert width ratio above 0.40.	• Retrofit bathymetry in the culvert area • Install upstream curtain wall • Install downstream weir
Southern Iowa Drift Plain	• Structural engineers should *avoid placing an over-sized channel expansion area* near culvert inlet by attempting to attain a stream-culvert width ratio above 0.40. • Land managers should avoid channelization in the culvert upstream by attempting to attain a sinuosity larger than 1.3. *Re-meandering* channelized and *straightened streams could be another suggestion for preventing culvert sedimentation.* • Culvert designers should, through land management practices, increase the percentage of forestry in culverts' stream corridor (represented by the blue cluster) to reduce the risk of culvert sedimentation.	• Retrofit bathymetry in the culvert area • Install upstream curtain wall • Install downstream weir
Paleozoic Plateau	Structural engineers should *avoid placing an over-sized channel expansion area* near culvert inlet by attempting to attain a stream-culvert width ratio above 0.40.	Use pipe-arched culvert design

The present research has brought several original research contributions that can be organized around three aspects: (1) formulation of a data-driven framework for investigating complex transport problems in the natural environment, (2) elicitation of data-driven insights for understanding the culvert sedimentation process, and (3) development of the "IowaDOT Culverts" web-based geospatial platform for aiding culvert design and management. While the first two contributions are of a more exploratory nature and susceptible to more refinements, the most notable practical contribution of this study is the development of the web-based portal as an innovative decision-support tool that helps the authorized personnel gather useful information and data-driven insights related to the causality of sedimentation as well as to develop effective strategies for mitigating sediment depositions at culverts. The web-accessible portal is a continuously updated resource about the culvert life-cycle that plays the following roles: a) informs and facilitates the design of culverts by accounting not only for the hydraulic aspects of the sizing but also for the sedimentation that is not currently considered in the design stage of these structures, b) guides the culvert maintenance program, and c) serves as a repository that can be mined to infer aspects of current designs as applied for large geographic areas. The IowaDOT Culverts cyber system is developed using a flexible and extendable structure that can be slightly modified to address similar transportation assets management concerns, such as scouring and wood debris accumulations at culverts and bridges. The platform is built on an extensive third-party dataset with open-source technologies that makes the system lightweight, low-cost, and easily expanded at the national level. The prototyped "IowaDOT Culverts" web-portal is currently transferred onto the Iowa DOT's servers for supporting routine culvert management and sedimentation mitigations.

The more generic outcome of this research consists in setting the foundation for a generalized data-driven framework that combines data-mining with machine learning and expert knowledge to investigate highly complex, large-scale transport problems in the natural environment. The implementation case used here for illustrating the usefulness of the generalized data-driven-framework is culvert sedimentation, an operational nuisance for transportation agencies in many parts of the U.S. The implementation case maintains the holistic, systems approach of the general investigation but it does so with more cost-efficient and effective means compared with the conventional investigative alternatives. The method can best be applied in data-rich watersheds or in areas where surrogates for that data are available. Fortunately, these data are increasingly available through the expansion of remote sensing technologies that survey watershed properties over large scales at a fraction of the cost compared with conventional observational means. Provided with adequate data, this case study and the framework are scalable to the Contiguous US, and can be readily adapted to address sedimentation or other river management issues, such as habitat deterioration and water pollution.

6.3 Recommendations for future studies

There are many accounts proving that the data-driven approach plays an important role in enabling heuristic exploration of the complex, large-scale environmental systems where there is limited a priori knowledge to support conceptual or physically-based modeling (Mount et al., 2016; Roderick, 2015). A necessary assumption for the feasibility of the data-driven modeling is that the data is sufficiently detailed and of good quality to enable hypothetical insights and conceptualizations. If this assumption is fulfilled, the analysis can not only shed lights on how different elemental processes in the overall system function, but it can also generate new hypotheses to allow the investigation of individual elements through conceptual or physically-based modeling. In recent years, the co-production of data-driven modeling, and conceptual or physically-based modeling is gaining popularity among hydrological and sediment research. This symbiotic synergy between the two modeling approaches is often referred as Hybrid modeling (e.g. Mekonnen et al. 2015).

From these perspectives, the data-driven framework presented in this research can be extended in a number of ways:

- Expansion of the conducted analysis to culvert sites located in various regions of the continental U.S.
- Using the data-driven insights provided through this study (selection of the relevant variables and their parameterization) to further guide sedimentation at culvert investigations through physically-based analysis. In this scenario, the data-driven framework developed through this study aids scientists and engineers in hypothesis-generation efforts on a much larger scale.
- The forecasting of culvert sedimentation potential can be further enhanced by integrating regression models with regionalized SOM analysis. Use of these models will enable more precise predictions of the sedimentation degree at specific culverts (e.g., 67% rather than an interval like 63-70%).
- Development of a modular design for the web-based platform to facilitate the extent of its current coverage to the Contiguous US scale.

6.4 References

Mekonnen, B.A., et al., 2015. Hybrid modeling approach to prairie hydrology: fusing data-driven and process-based hydrological models. Hydrological Sciences Journal, 60, 1473–1489. (in press). doi:10.1080/02626667.2014.935778.

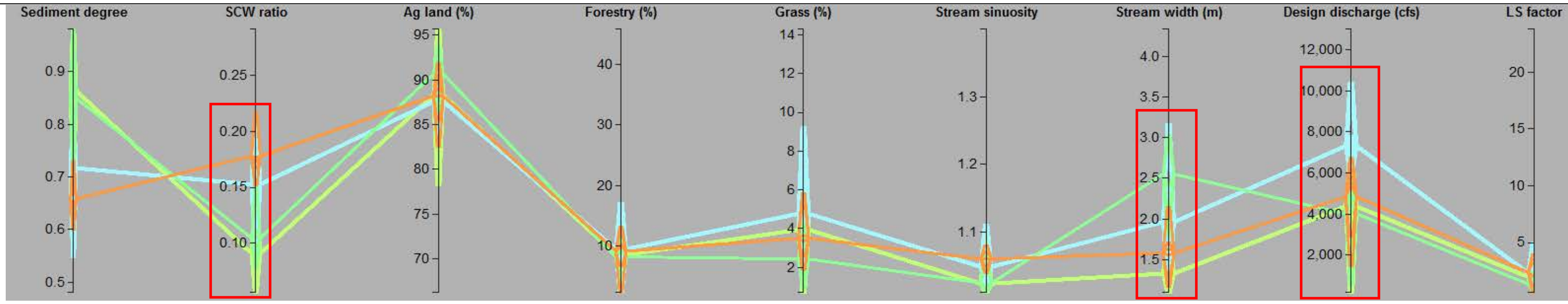
Mount, N. J., Maier, H.R., Toth, E., Elshorbagy, A., Solomatine, D., Chang, F.-J., and Abrahart, R. J. 2016. Data-driven modeling approaches for socio-hydrology: opportunities and challenges within the Panta Rhei Science Plan, Hydrological Sciences Journal: 61 (7). Pp 1192-1208.

Roderick, M. (2015). Mathematical and Computational Modeling: With Applications in Natural and Social Sciences, Engineering, and the Arts. Wiley. ISBN 978-1-118-85398-6.

APPENDIX A

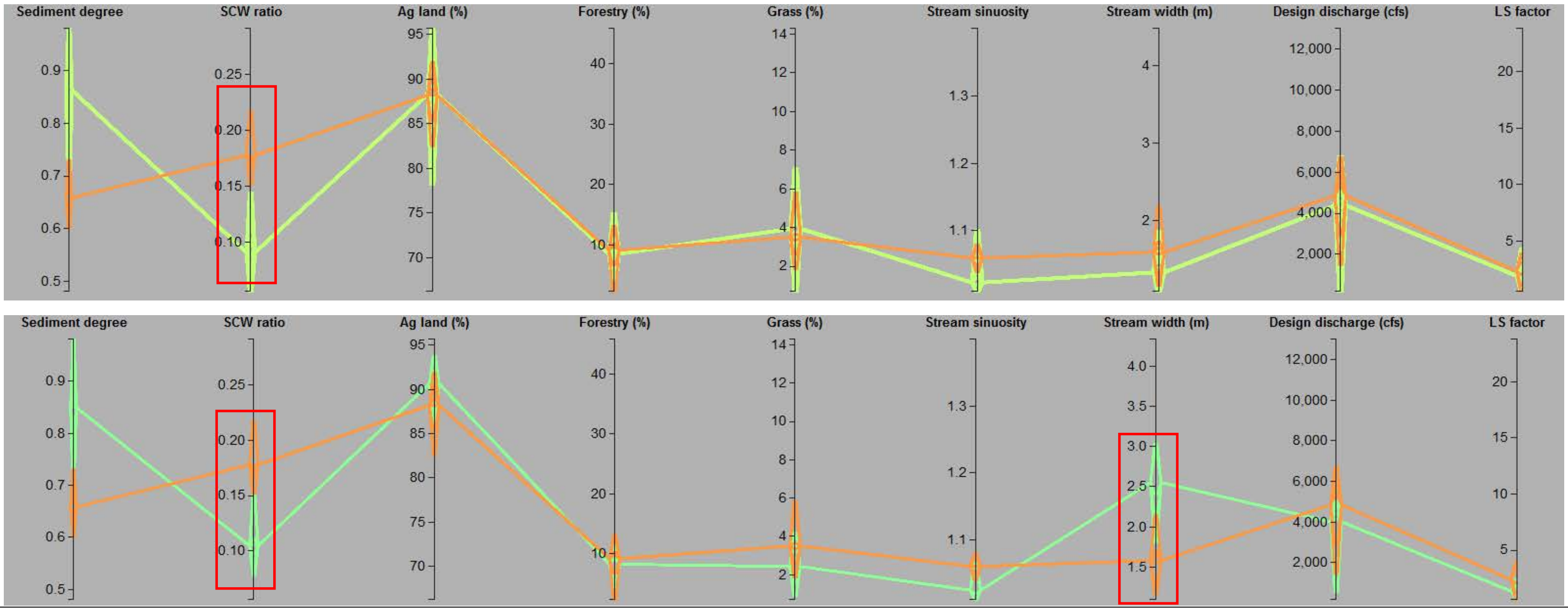
Table A.1 Data-driven insights for designing culverts in the Northwestern Iowa Plains region

Northwestern Iowa Plains



Key drivers: (1) stream-culvert width ratio (SCW), (2) stream width, and (3) design discharge. Excepting for watershed interventions for reducing the erosion and sediment transport in the culvert drainage area, the only factor that can be controlled for mitigation of sedimentation at culverts is the selection of an appropriate stream-culvert width ratio during culvert design.

(Table A.1 Continued)



Design suggestions:

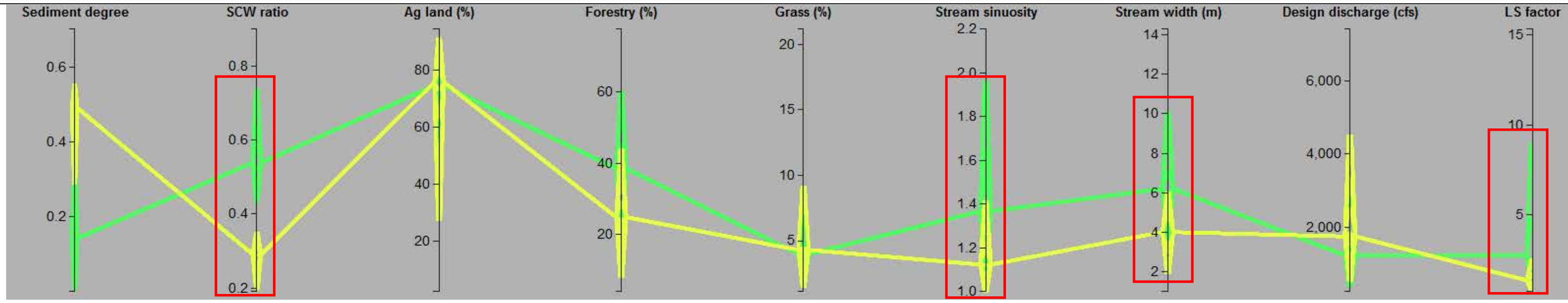
- Structural engineers should *avoid placing an over-sized channel expansion area* near culvert inlet by attempting to attain a stream-culvert width ratio above 0.15.

Table A.2 Multivariate dependency between the culvert sedimentation degree and drivers in the Northwestern Iowa Plains region

Sediment degree	SCW ratio	Ag land (%)	Forestry (%)	Grass (%)	Stream sinuosity	Stream width (m)	Design discharge (cfs)	LS factor
0.70-0.98	0.06-0.14	78-95	3-15	0-7	1.01-1.10	1.1-1.8	184-6758	0.63-4.30
0.70-0.98	0.08-0.15	86-93	3-11	1-4	1.01-1.07	1.9-3.1	579-6206	0.61-2.39
0.60-0.73	0.15-0.22	82-91	2-13	2-6	1.04-1.08	1.2-2.2	1506-6722	0.59-3.98
0.60-0.70	0.16-0.19	84-85	15-20	2-6	1.06-1.11	1.7-2	1870-6437	3.84-4.48
0.55-0.82	0.09-0.21	83-92	2-17	2-9	1.02-1.11	1.2-3.3	5971-10363	0.82-4.99

Table A.3 Data-driven insights for designing culverts in the Des Moines Lobe region

Des Moines Lobe



Key drivers: (1) stream-culvert width ratio, (2) stream sinuosity, (3) stream width, and (4) LS factor are important factors that are sensitive to the culvert sedimentation degree in the Des Moines Lobe. In this region, the only human-controllable factors are (a) SWC ratio and (b) stream sinuosity.

Design suggestions:

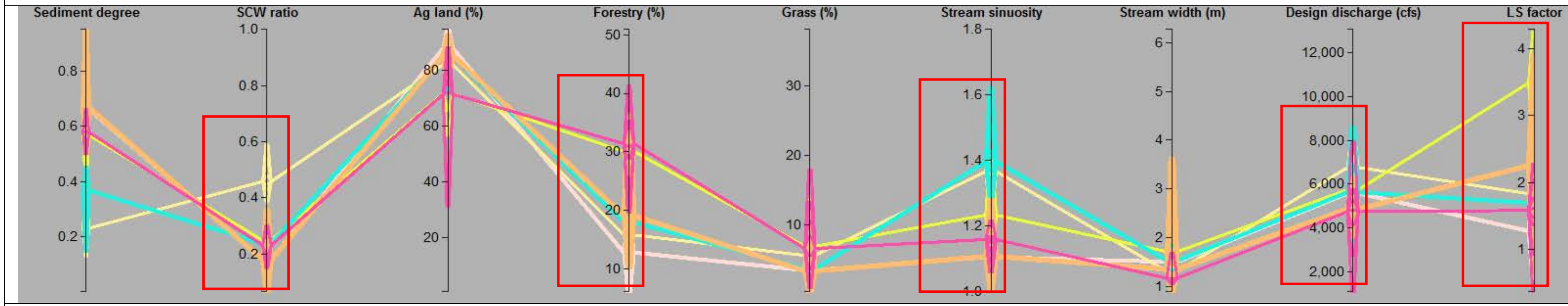
- Structural engineers should *avoid placing an over-sized channel expansion area* near culvert inlet by attempting to attain a stream-culvert width ratio above 0.40.
- Land managers should *avoid channelization* in the culvert upstream by attempting to attain a sinuosity larger than 1.4. *Re-meandering* channelized and *straightened streams could be another suggestion for preventing culvert sedimentation.*
- Subsurface drainage tiles are extensively used in the Des Moines Lobe (Crumpton et al. 2012), which in turn reduce surface runoff and erosion, therefore offering a margin of freedom in the design.

Table A.4 Multivariate dependency between the culvert sedimentation degree and drivers in the Des Moines Lobe region

Sediment degree	**SCW ratio**	**Ag land (%)**	**Forestry (%)**	**Grass (%)**	**Stream sinuosity**	**Stream width (m)**	**Design discharge (cfs)**	**LS factor**
0.29-0.55	0.20-0.35	27-91	8-44	1-9	1.00-1.41	2-6	573-4491	0.76-2.48
0.28-0.55	0.32-0.50	74-78	27-29	2-5	1.07-1.77	3.4-8	3770-7412	0.77-4.40
0.26-0.50	0.46-0.56	85-89	14-26	2-4	1.03-1.66	7-9.5	1085-2984	0.91-1.15
0.15-0.42	0.45-0.65	65-86	16-23	2-7	1.01-1.38	2.3-8	2453-5084	0.80-1.30
0.01-0.32	0.44-0.73	50-90	19-60	2-7	1.05-1.96	3.5-10	429-3006	0.70-8.88
0.15-0.29	0.39-0.49	78-88	8-35	1-3	1.23-1.75	2.79-7.5	690-3771	0.86-1.65

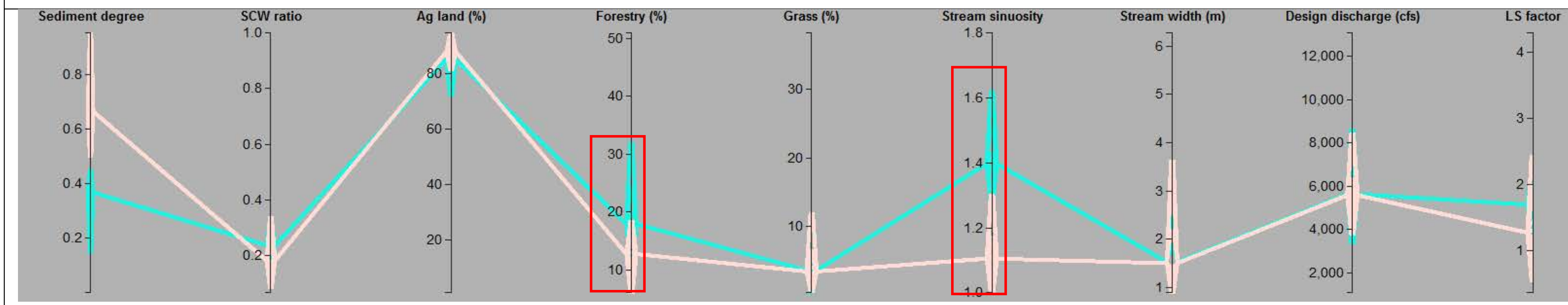
Table A.5 Data-driven insights for designing culverts in the Iowan Surface region

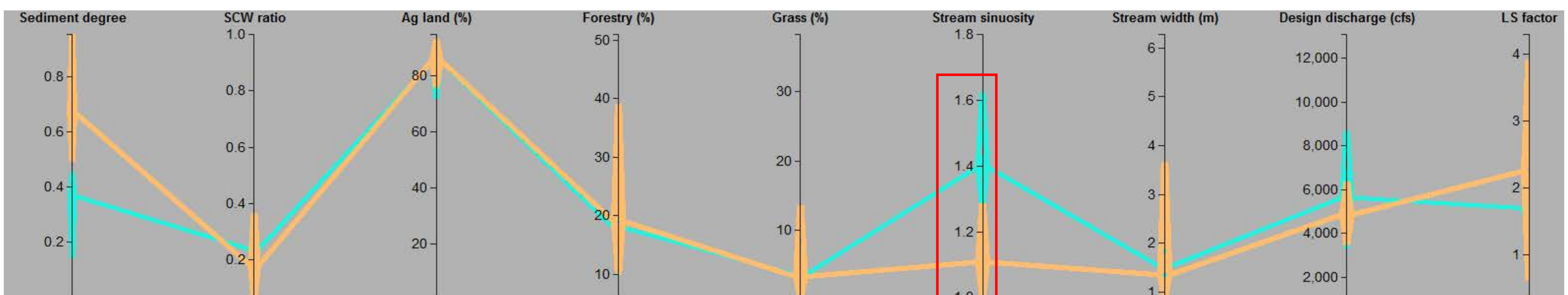
Iowan Surface



Key drivers:

The overview of major clusters indicates that (1) stream-culvert width ratio, (2) stream sinuosity, (4) forestry, (5) design discharge, and (6) LS factor, are the most sensitive factors to the culvert sedimentation degree in the Iowan Surface Region. Among these factors, *stream-culvert width ratio*, *stream sinuosity*, and *forestry* are human-controllable factors.

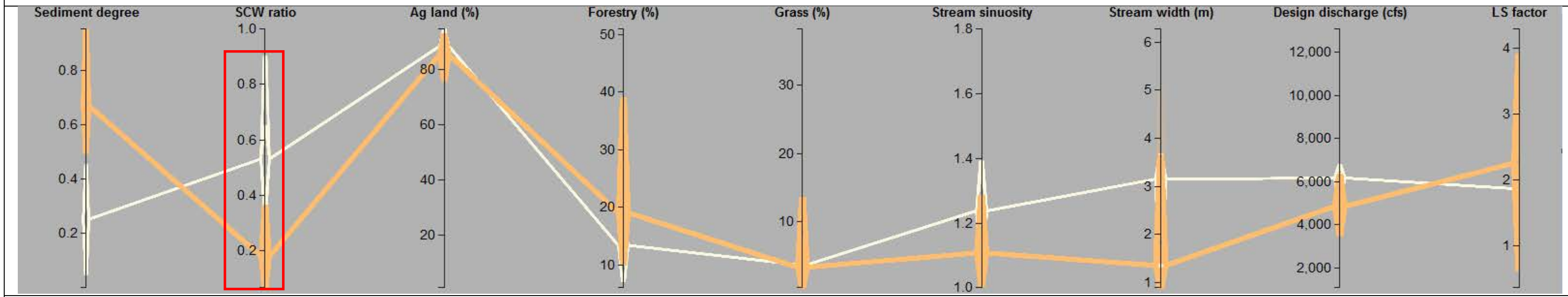




(Table A.5 Continued)

Design suggestions:

- Land managers should avoid channelization in the culvert upstream by attempting to attain a sinuosity larger than 1.3. *Re-meandering* channelized and *straightened streams could be another suggestion for preventing culvert sedimentation.*
- Culvert designers should, through land management practices, increase the percentage of forestry in culverts' stream corridor (represented by the blue cluster) to reduce the risk of culvert sedimentation.



Design suggestions:

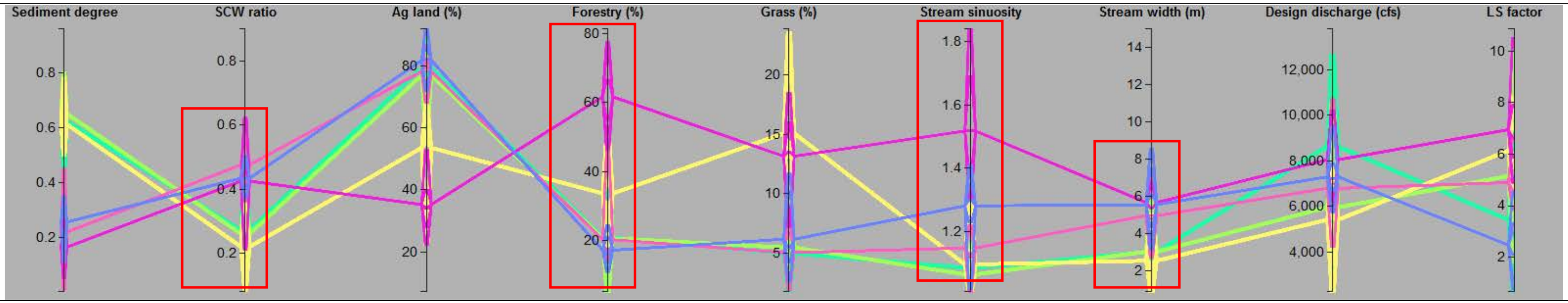
- Structural engineers should *avoid placing an over-sized channel expansion area* near culvert inlet by attempting to attain a stream-culvert width ratio above 0.40.

Table A.6 Multivariate dependency between the culvert sedimentation degree and drivers in the Iowan Surface region

Sediment degree	**SCW ratio**	**Ag land (%)**	**Forestry (%)**	**Grass (%)**	**Stream sinuosity**	**Stream width (m)**	**Design discharge (cfs)**	**LS factor**
0.50-0.95	0.07-0.36	77-93	10-38	0-13	1.01-1.28	0.9-5	3594-6272	0.65-3.88
0.50-0.95	0.09-0.34	81-95	6-18	0-11	1.00-1.30	0.9-5	3825-8396	0.56-2.43
0.40-0.79	0.13-0.35	57-85	20-37	3-14	1.01-1.42	1.1-3.5	5020-7513	2.56-4.24
0.50-0.66	0.15-0.30	31-88	20-41	1-18	1.06-1.27	1.1-2.1	1090-7974	0.38-2.28
0.25-0.50	0.32-0.41	81-89	20-25	1-7	1.23-1.33	1.2-1.5	8075-9789	0.78-2.03
0.13-0.44	0.37-0.59	78-91	12-23	2-7	1.23-1.58	1.3-1.6	5951-8581	1.27-2.50
0.05-0.45	0.26-0.90	80-95	7-16	1-11	1.03-1.39	1.3-5	5196-6762	0.97-2.76
0.15-0.45	0.14-0.33	72-90	10-31	0-11	1.28-1.62	1.1-3.5	3423-8608	1.23-2.19

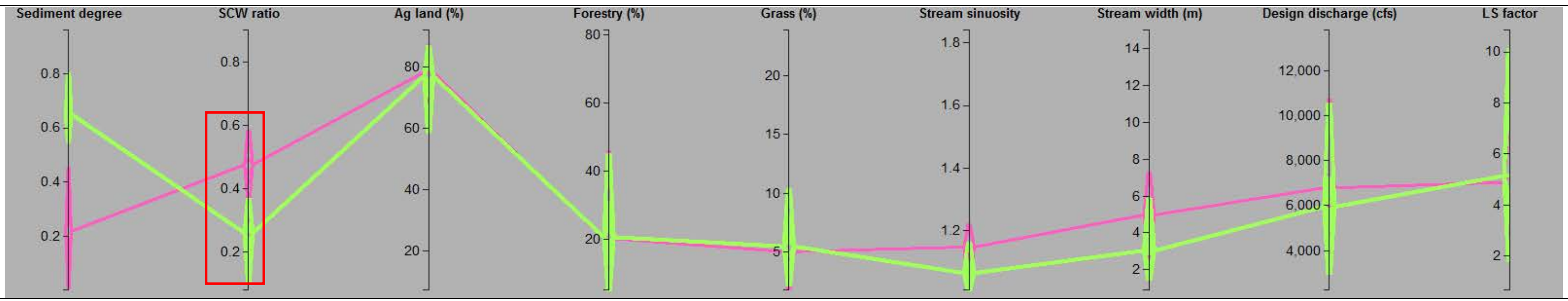
Table A.7 Data-driven insights for designing culverts in the Southern Iowa Drift Plain region

Southern Iowa Drift Plain



Key drivers:

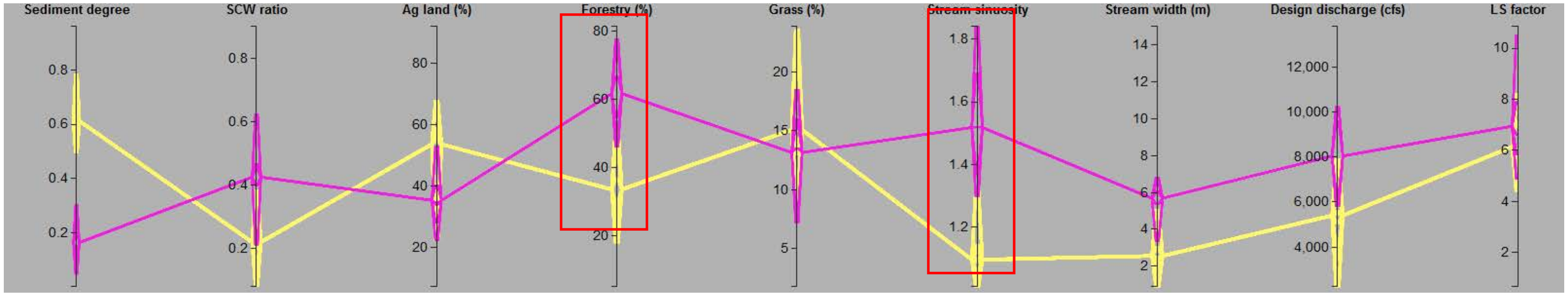
The overview of the multivariate patterns highlights group drivers that are sensitive to culvert sedimentation in the Southern Iowa Drift Plain. These factors include (1) stream-culvert width ratio, (2) forestry, (3) sinuosity, (4) design discharge, and (5) stream width. Among these factors, stream-culvert width ratio, forestry, and sinuosity are drivers that can be modified through culvert structural design and land management.



Design suggestions:

- Structural engineers should *avoid placing an over-sized channel expansion area* near culvert inlet by attempting to attain a stream-culvert width ratio above 0.40.

(Table A.7 Continued)



Design suggestions:

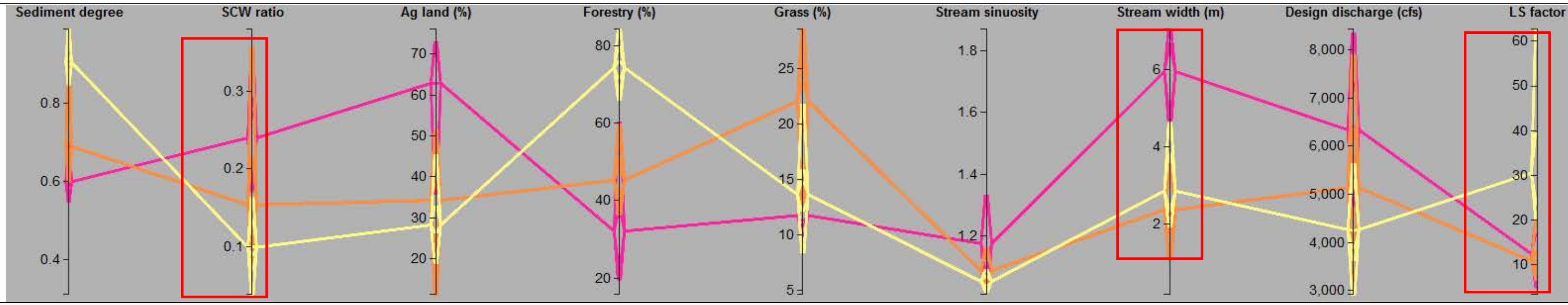
- Land managers should avoid channelization in the culvert upstream by attempting to attain a sinuosity larger than 1.3. *Re-meandering* channelized and *straightened streams could be another suggestion for preventing culvert sedimentation.*
- Culvert designers should, through land management practices, increase the percentage of forestry in culverts' stream corridor (represented by the blue cluster) to reduce the risk of culvert sedimentation.

Table A.8 Multivariate dependency between the culvert sedimentation degree and drivers in the Southern Iowa Drift Plain region

Sediment degree	**SCW ratio**	**Ag land (%)**	**Forestry (%)**	**Grass (%)**	**Stream sinuosity**	**Stream width (m)**	**Design discharge (cfs)**	**LS factor**
0.55-0.80	0.11-0.37	58-86	5-45	2-10	1.01-1.16	1.5-5.8	3004-10524	1.83-10.07
0.50-0.78	0.09-0.38	22-67	18-52	9-23	1.01-1.32	0.9-5	2245-7867	4.43-8.17
0.45-0.75	0.11-0.47	60-89	11-40	2-10	1.03-1.15	1.7-5.6	5365-12641	0.66-6.93
0.01-0.45	0.36-0.58	68-87	14-46	2-8	1.01-1.22	2.8-7.3	4234-10724	3.86-6.71
0.11-0.35	0.37-0.50	72-92	11-24	2-12	1.02-1.41	3.2-8.6	5768-9497	0.80-4.53
0.05-0.30	0.21-0.62	22-53	46-78	7-18	1.30-1.84	3.4-6.8	5833-10217	4.91-10.50
0.01-0.20	0.36-0.88	37-70	50-64	7-14	1.24-1.52	3.9-7.7	5473-11136	2.05-5.43

Table A.9 Data-driven insights for designing culverts in the Paleozoic Plateau region

Paleozoic Plateau



Key drivers:

In the Paleozoic Plateau region, the three major clusters exhibit a very distinct multivariate pattern. Compared with other regions, the Paleozoic Plateau region has significantly larger LS factor (from 10-60). The yellow cluster represents culverts in the Coon-Yellow watersheds (steep and mountainous areas) in the north-eastern corner of the state. The culverts represented by orange and pink clusters are located in watersheds with mild slopes. According to the field observation (illustrated in Table 5.7), this culvert sedimentation in this region is dominated by the mixed loads (fine-grained and coarse particles). Multiple culvert sites are observed to have deposits made of gravel and cobble.

In the Paleozoic Plateau region, the slope length-steepness factor (LS-factor) is much higher than that of other regions and is the only sensitive environmental factor to the culvert sediment degrees. Therefore, environmental variables (e.g., agriculture, forestry, and grassland use) that are pre-selected based on suspended load-related processes are no longer valid. This circumstance can be justified by both field inspections (observations of cobbles) and the domain knowledge that many streams in the mountainous (the large LS-factor) are bedrock-controlled or semi-controlled, therefore their sedimentation responses are very different from the alluvium channels in other regions (IGS, 2017; Schumm, 1985).

Design suggestions:

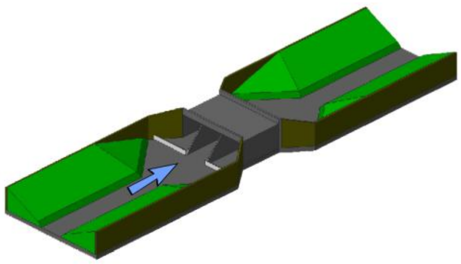
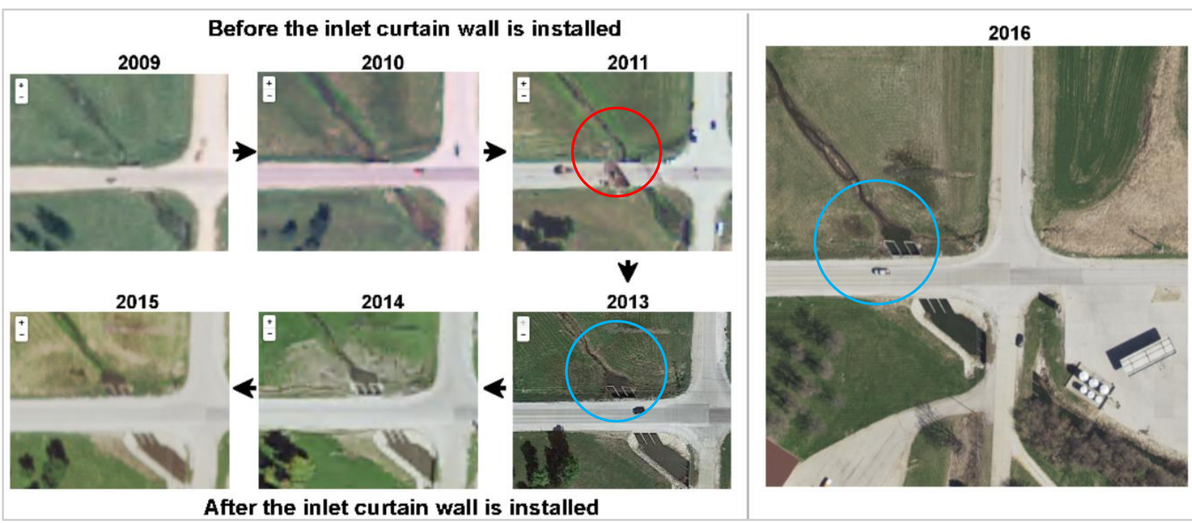
- Given the unique characteristics of this region, this study suggests that culvert designers should adopt the *pipe-arch culvert design* (Kosicki and Davis, 2001) proposed by the Maryland State Highway Administration (MDSHA). Conceptually, this structural design minimizes the change of natural channel geometry at culvert location, thus can prevent the deposition of coarse particles near culvert inlets.

Table A.10 Multivariate dependency between the culvert sedimentation degree and drivers in the Paleozoic Plateau region

Sediment degree	SCW ratio	Ag land (%)	Forestry (%)	Grass (%)	Stream sinuosity	Stream width (m)	Design discharge (cfs)	LS factor
0.85-0.99	0.04-0.16	19-45	66-85	9-22	1.02-1.09	2-4.6	2914-5626	20.23-62.69
0.60-0.91	0.09-0.36	11-51	36-60	12-28	1.03-1.16	1.1-3.8	3038-7893	6.54-19.68
0.62-0.65	0.13-0.17	50-70	24-55	5-9	1.07-1.36	0.2-3.3	3500-5265	4.22-12.16
0.55-0.70	0.18-0.35	36-73	20-60	10-13	1.02-1.33	4.6-7	4772-8336	4.92-18.04

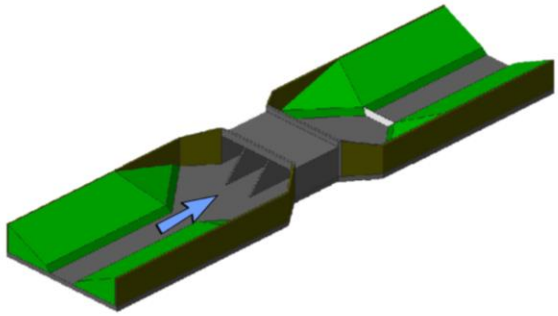
APPENDIX B

Table B.1 Culverts with inlet curtain wall

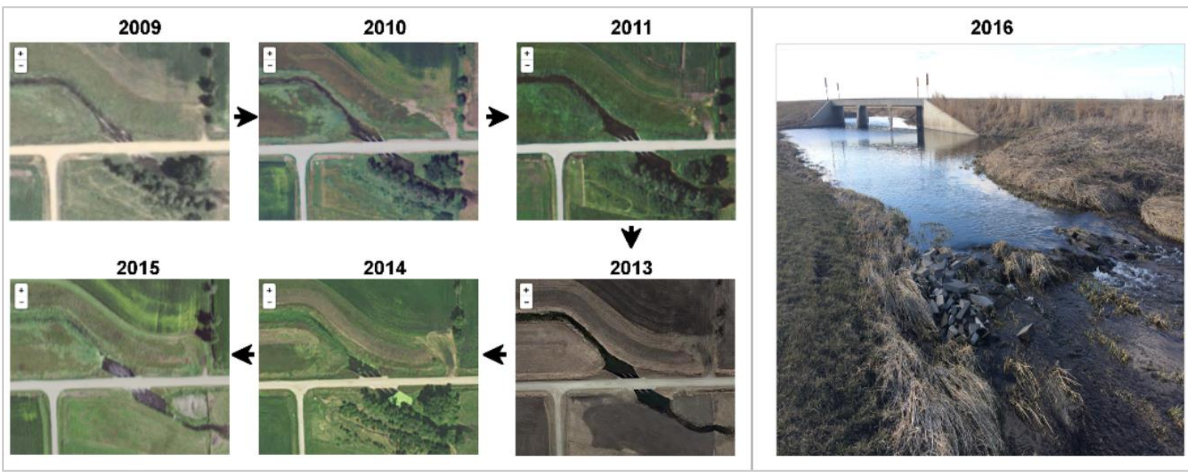
<table>
<tr><th colspan="2">Structure overview</th><th colspan="2">Field observations</th></tr>
<tr><td colspan="2"></td><td colspan="2"></td></tr>
<tr><td colspan="4">Effects on the culvert sedimentation</td></tr>
<tr><td colspan="4">

This observation is made at a culvert in the Southern Iowa Drift Plain (Structure Code: 363500). Before the installation of the inlet curtain wall (before 2011), the culvert was severely silted. After the installation of this structure (from 2013 on), sediments stopped accumulating at the culvert site.
It is also observed that several other culverts with this structural feature are likely to have medium-low culvert sedimentation, even these sites are located in sedimentation-prone areas. </td></tr>
<tr><td colspan="4">Culverts with the inlet curtain wall</td></tr>
<tr><td>363500 (Southern Iowa Drift Plain)</td><td>Low sediment</td><td>337281 (Iowan Surface)</td><td>Medium sediment</td></tr>
<tr><td>608945 (Northwest Iowa Plains)</td><td>Medium sediment</td><td>337301 (Iowan Surface)</td><td>Medium-low sediment</td></tr>
</table>

Table B.2 Culverts with downstream weir

Structure overview	Field observations
	

Effects on the culvert sedimentation



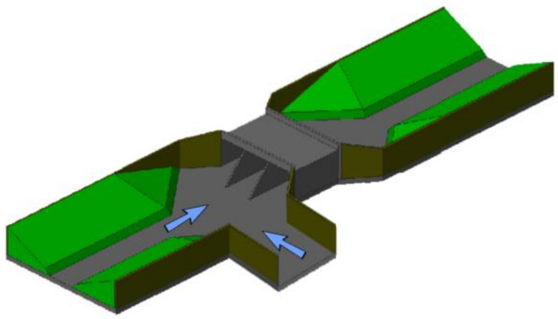
The downstream weir is able to decrease the flow velocities in the expansion-contraction area, and expand inundated areas at the culvert inlet, which encourage sedimentation but prevent vegetation growth. Without the stabilization from the vegetation, sediment deposition can be readily flushed away during high-flow events.

The field inspection shown that most of the culverts with the downstream weir do not have sedimentation problems.

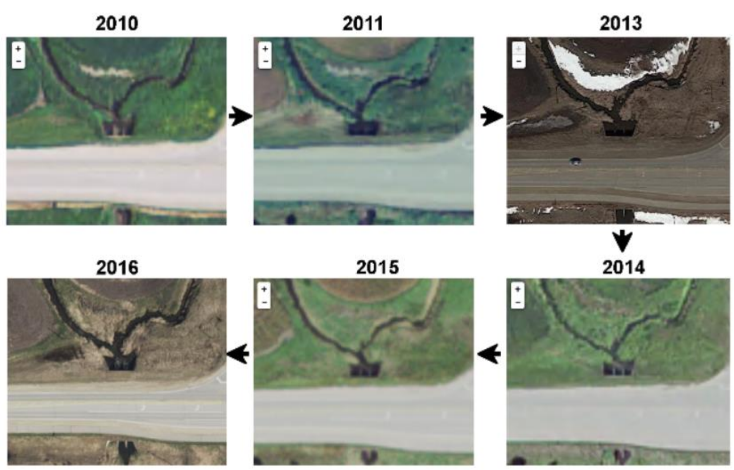
Culverts with downstream weir

18791 (Iowan Surface)	Low sediment	105161 (Iowan Surface)	Low sediment
365740 (Des Moines Lobe)	Low sediment	24931 (Iowan Surface)	Low sediment
65051 (Southern Iowa Drift Plain)	Low sediment		

Table B.3 Culverts at stream confluence

Site condition	Field observations
	

Effects on the culvert sedimentation

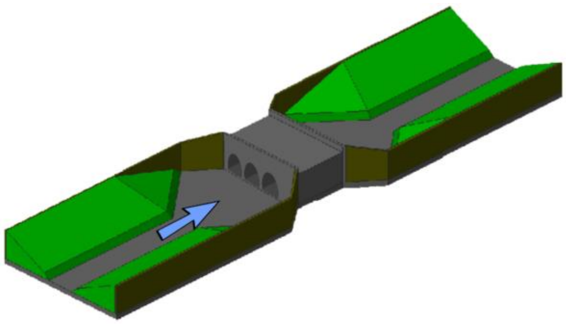


Many culverts grouped into the minor clusters are constructed near stream confluences. These culverts are observed to have low sedimentation potential. It is assumed that the streamwise-oriented vortical structures formed in the mixing interface limits the sediment entrainment (Miyawaki et al., 2009, 2010) near the culvert expansion area.

Culverts locatated at stream confluence

25891 (Iowan Surface)	Low sediment	339671 (Southern Iowa Drift Plain)	Low sediment
122931 (Iowan Surface)	Low sediment	58342 (Southern Iowa Drift Plain)	Low sediment
33321 (Southern Iowa Drift Plain)	Low sediment	29390 (Southern Iowa Drift Plain)	Medium-low sediment

Table B.4 Pipe culverts

<table>
<tr><th colspan="2">Site condition</th><th colspan="2">Field observations</th></tr>
<tr><td colspan="2"></td><td colspan="2"></td></tr>
<tr><td colspan="4">Effects on the culvert sedimentation</td></tr>
<tr><td colspan="4">

Few pipe culverts are observed with low sediment deposition in the Northwest Iowa Plains and Des Moines Lobe region.</td></tr>
<tr><td colspan="4">Pipe culverts</td></tr>
<tr><td>261641 (Northwest Iowa Plains)</td><td>Low sediment</td><td>504405 (Des Moines Lobe)</td><td>Low sediment</td></tr>
<tr><td>365000 (Northwest Iowa Plains)</td><td>Low sediment</td><td></td><td></td></tr>
</table>