SEVENTH FRAMEWORK PROGRAMME

THEME ICT-2013.5.4

"ICT for Governance and Policy Modelling"

D4.2.1
Optimization and Visual Analytics Report

Project acronym: Consensus

Project full title: *Multi-Objective Decision Making Tools through Citizen Engagement*

Contract no.: 611688

Workpackage:	4	Optimization & Visual Analysis	
Editors:	M. Gavish, D. Baras, A. Ronen		IBM
Author(s):	M. Gavish, D. Baras, A. Ronen K. Tserpes, A. Xenaki S. Frank, P. Havlik L. Kallipolitis J. Fuchs, T. Schreck, D. Keim G.Ceccarelli L.Mathe A. Kopsacheili, G. Yannis, K. Diamandouros		IBM NTUA IIASA ATC UKON OXFAM WWF ERF
Authorized by	K. Tserpes		NTUA
Doc Ref:	D4.2.1		
Reviewer(s):	K. Tserpes A. Ronen		NTUA IBM
Dissemination Level	PU		

Consensus Consortium

No	Name	Short name	Country
1	Institute of Communication and Computer Systems/National Technical University of Athens	NTUA	Greece
2	IBM Israel Science and Technology Ltd.	IBM	Israel
3	International Institute for Applied Systems Analysis	IIASA	Austria
4	Athens Technology Center	ATC	Greece
5	University of Konstanz	UKON	Germany
6	OXFAM Italia ONLUS	OXFAM	Italy
7	WWF - World Wide Fund for Nature	WWF	Switzerland
8	European Union Road Federation	ERF	Belgium

Document History

Version	Date	Changes	Author/Affiliation
v.0.1	19-06-2014	TOC	M. Gavish/IBM
v.0.2	19-06-2014	TOC Review	K. Tserpes /NTUA
v.0.31	05-08-2014	Section 2.5, Chapter 6	J. Fuchs/UKON T. Schreck/UKON D. Keim/UKON
v.0.32	28-08-2014	Section 2.4	A. Kopsacheili/ERF G. Yannis/ERF K. Diamandouros/ERF
v.0.33	29-08-2014	Section 2.3	L.Mathe/WWF, G.Ceccarelli/OXFAM
v.0.34	01-09-2014	Chapter 3	S. Frank/IIASA P. Havlik/IIASA
v.0.35	03-09-2014	Section 2.6, Chapter 5	A. Xenaki/ NTUA N. Dimakopoulos/ATC L. Kallipolitis/ATC
v.0.36	05-09-2014	Section 2.2, Chapter 4	M. Gavish/IBM D. Baras/IBM
v.0.4	14-09-2014	1st Integrated Version	M. Gavish/IBM D. Baras/IBM
v.0.5	16-09-2014	Review of Individual Chapters by Authors	M. Gavish/IBM D. Baras/IBM
v.0.6	24-09-2014	Revision for Review	M. Gavish/IBM D. Baras/IBM
v0.61	28-09-2014	Review Comments	K. Tserpes /NTUA
v.1.0	29-09-2014	Final Revision	M. Gavish/IBM D. Baras/IBM

Executive Summary

This paper, Deliverable D4.2.1 Optimization and Visual Analytics Report, is an official deliverable that accompanies Deliverable D4.1.1 Optimization and Visual Analytics Prototypes.

In this paper, we document our research challenges and findings and describe the models and components we developed within WP4. We also document the Consensus software components, including their design, deployment, and use.

Specifically, the major topics of this report are multi-objective optimization, visual-interactive aids, conflict analysis, and crowdsourcing validation.

This report is the first of three revisions of the Optimization and Visual analytics report and is submitted in Month 12 of the project. The next revision will be submitted in Month 24 and the final one in Month 30.

This deliverable is organized into six chapters:

Chapter 1 is an introductory chapter that provides more details about this document and its methodology and scope.

Scientific background is presented in **Chapter 2**.

Chapter 3 deals with the GLOBIOM optimization model.

In **Chapter 4**, we present the Consensus Multi-Objective Optimization and Visualization Tool (MOOViz). This is a major prototype within this work package that is intended for policy decision makers to assist them in the overall process of decision making.

Chapter 5 is dedicated to Consensus Game—a web tool intended for the public and aimed at education, collaboration, and communicating policy decision conflicts to the citizens as well as for enabling citizens to express their policy preferences.

Finally, **Chapter 6**, Visual Analytics, focuses on visual support, interaction possibilities, and automatic algorithms that are essential for augmenting the capabilities in the decision cycle.

Contents

Executive Summary 3

Contents 4

Table of Figures 8

List of Tables 14

Referenced project deliverables 14

1 Introduction 15

1.1 Objectives 15

1.2 Scope 16

1.3 Methodology 16

1.4 Structure 17

1.5 Quality Management 17

2 Background 19

2.1 Introduction 19

2.2 Multi-Objective Optimization 19

2.2.1 Introduction 19

2.2.2 Formulation 20

2.2.3 Approaches 21

2.3 Multi-Criteria Decision Making in the Environmental Sector 26

2.3.1 Introduction 26

2.3.2 Arguments for Using MCA 27

2.3.3 Types of MCAs Applied in the Environmental Sector: 27

2.3.4 Common Stages in Applying MCA 28

2.3.5 Multi-Criteria Analysis and Biofuels 28

2.3.6 Further Recommendations 30

2.4 Multi-Criteria Decision Making in the Transport Sector 31

2.4.1 Introduction 31

2.4.2 General Procedure of Multi-Criteria Decision Making in Transport Sector 31

2.4.3 Decision Making Approaches in Transport Sector 32

2.4.4 Decision Making Subjects in Transport Sector 33

2.4.5 Role of Multi-Criteria Decision Making in Transport Sector........................... 34

2.4.6 Multi-Criteria Decision Making Methods Used in Transport Sector 35

2.4.7 Multi-criteria Decision Making (*Evaluation*) Parameters Commonly Used in Transport Sector 36

2.4.8 Participation of Stakeholders in Multi-Criteria Decision Making in Transport Sector 39

2.4.9 Multi-Criteria Decision Making in Transport Policy Scenario of Consensus.... 40

2.5 Visual Analytics 41

2.5.1 Introduction 41

2.5.2 Glyph-Based Evaluation 41

2.6 Gamification and Crowdsourcing 42

2.6.1 Introduction 42

2.6.2 Gamification 43

2.6.3 Croudsourcing 48

2.6.4 Serious Games 48

3 The GLOBIOM Optimization Model.. 50

4 Multi-Objective Optimization and Visualization Tool (MOOViz).................................. 55

4.1 Introduction to the MOOViz Tool 55

4.2 Introduction to Multi Objective Optimization Problems 56

4.2.1 Mathematical Background 57

4.3 Research Overview 58

4.4 MOOViz technical Model Specification 58

4.4.1 Domain Definition for MOOViz Tool.. 58

4.4.2 Policy Alternatives Data... 62

4.4.3 Decision Scenario Data 62

4.5 Using MOOViz Tool for Bio-Fuel Scenario 63

4.5.1 Introduction 63
4.5.2 The Polygon 64
4.5.3 Policy Glyphs 65
4.5.4 Range sliders 68
4.5.5 Optimal and Auto-excluded policies 69
4.5.6 Top left panel 71
4.5.7 Lines view 72
4.5.8 Favorites pane 77
4.5.9 Tradeoff analyzer 78
4.5.10 Zoom in on views 80
4.6 Using MOOViz Tool for Transportation Scenario 83
4.6.1 Introduction 83
4.6.2 Lines view 84
4.6.3 The policy line 87
4.6.4 Range sliders 90
4.6.5 Optimal and Auto-excluded policies 91
4.6.6 Top left panel 94
4.6.7 The Polygon 94
4.6.8 Policy Glyphs 95
4.6.9 Favorites pane 97
4.6.10 Tradeoff analyzer 98
4.6.11 Zoom in on views 99
5 Consensus Game 101
5.1 Introduction 101
5.2 Research Challenges 101
5.3 Consensus Game 102
5.3.1 Game Scenario 103

5.3.2 Implementation Details 103

5.3.3 Prototype Access Details 105

5.4 Reward Model 105

5.4.1 Mathematical Model 106

5.4.2 Other Forms of Reward Models 107

5.5 Game Evaluation 107

5.5.1 Learning Curve 108

5.5.2 Participation 108

5.5.3 Implementation details 108

5.5.4 Prototype access details 109

6 Visual Analytics 110

6.1 Introduction 110

6.2 Visual Support for Similarity Perception 110

6.2.1 Experiment 1: Contours for Novices vs. Experts 112

6.2.2 Experiment 2: Perception of Similarity 118

6.2.3 Experiment 3: Improvements for Star Glyph 122

6.2.4 Design Considerations 127

6.2.5 Conclusion 128

6.3 Visual Alignment: A Technique to Facilitate the Comparison of Multi-Dimensional Data Points 128

6.3.1 Visual Alignment for Scatterplot Matrices 129

6.3.2 Visual Alignment for Parallel Coordinate Plots 129

6.3.3 Visual Alignment for Data Glyphs 130

6.3.4 Conclusion and Future Work 130

6.4 Visual Analytics Prototype 131

6.5 Future Work 133

6.6 Workflow 133

Table of Figures

Figure 1: Consensus components' diagram 16

Figure 2: Procedure of MCDM in transport sector 32

Figure 3: High-level view of MOOViz workflow 55

Figure 4: Biofuel Scenario Domain Definition 60

Figure 5: Transportation Scenario Domain Definition 61

Figure 6: MOOViz tool in bio-fuel scenario – Snapshot of application 63

Figure 7: : MOOViz tool in bio-fuel scenario – main components 64

Figure 8: MOOViz tool in bio-fuel scenario – objective visualization 64

Figure 9: MOOViz tool in bio-fuel scenario – optimization direction 64

Figure 10: MOOViz tool in bio-fuel scenario –disabled objective 65

Figure 11: MOOViz tool in bio-fuel scenario – policy glyph 65

Figure 12: MOOViz tool in bio-fuel scenario – referring slice to objective 65

Figure 13: MOOViz tool in bio-fuel scenario – details of policy 66

Figure 14: MOOViz tool in bio-fuel scenario – tooltip window details 66

Figure 15: MOOViz tool in bio-fuel scenario – policy added to favorites 66

Figure 16: MOOViz tool in bio-fuel scenario – policy highlighted 67

Figure 17: MOOViz tool in bio-fuel scenario – policy marked as decision 67

Figure 18: MOOViz tool in bio-fuel scenario – example of more appealing policies in window tooltip 67

Figure 19: MOOViz tool in bio-fuel scenario – decision panel 67

Figure 20: MOOViz tool in bio-fuel scenario – delete decision example 68

Figure 21: MOOViz tool in bio-fuel scenario – policy glyph with a disabled objective 68

Figure 22: MOOViz tool in bio-fuel scenario – range sliders 68

Figure 23: MOOViz tool in bio-fuel scenario – optimal and auto-excluded tabs 69

Figure 24: MOOViz tool in bio-fuel scenario – optimal policies tab 69

Figure 25: MOOViz tool in bio-fuel scenario – example of gray policy glyph when mouse is placed over an optimal policy name 70

Figure 26: MOOViz tool in bio-fuel scenario – adding policy to favorites from optimal policies tab 70

Figure 27: MOOViz tool in bio-fuel scenario – example of notification on existing appealing policies 70

Figure 28: MOOViz tool in bio-fuel scenario – highlighted policy in the optimal policies tab 70

Figure 29: MOOViz tool in bio-fuel scenario – excluded policy in the optimal policies tab.... 70

Figure 30: MOOViz tool in bio-fuel scenario – decided policy in the optimal policies tab 71

Figure 31: MOOViz tool in bio-fuel scenario – combination of excluded, highlighted, decided and favorite policy 71

Figure 32: MOOViz tool in bio-fuel scenario – top left panel 71

Figure 33: MOOViz tool in bio-fuel scenario – moving from polygon to lines view 72

Figure 34: MOOViz tool in bio-fuel scenario – lines view 72

Figure 35: MOOViz tool in bio-fuel scenario – lines view, zoom on lines area 73

Figure 36: MOOViz tool in bio-fuel scenario – information on policy when mouse placed over corresponding lines 74

Figure 37: MOOViz tool in bio-fuel scenario – tooltip window in lines view 74

Figure 38: MOOViz tool in bio-fuel scenario – examples of highlighted, favorite and decided policy in lines view 75

Figure 39: MOOViz tool in bio-fuel scenario – rearranging order of objectives – in process . 76

Figure 40: MOOViz tool in bio-fuel scenario – rearranging order of objectives – after reorder 77

Figure 41: MOOViz tool in bio-fuel scenario – favorites pane 77

Figure 42: MOOViz tool in bio-fuel scenario – favorites view 78

Figure 43: MOOViz tool in bio-fuel scenario – examples of more appealing policies 78

Figure 44: MOOViz tool in bio-fuel scenario – tradeoff analyzer 79

Figure 45: MOOViz tool in bio-fuel scenario – tradeoff analyzer – adding policy to favorites 80

Figure 46: MOOViz tool in bio-fuel scenario – zoom in button 80

Figure 47: MOOViz tool in bio-fuel scenario – snapshot before zoom in 81

Figure 48: MOOViz tool in bio-fuel scenario – snapshot after zoom in 81

Figure 49: MOOViz tool in bio-fuel scenario – link to zoom in view 82

Figure 50: MOOViz tool in bio-fuel scenario – link to nested zoom in.................................... 82

Figure 51: MOOViz tool in transportation scenario – snapshot of application....................... 83

Figure 52: MOOViz tool in transportation scenario – main components 83

Figure 53: MOOViz tool in transportation scenario – lines view 84

Figure 54: MOOViz tool in transportation scenario – lines view, zoom on lines area 85

Figure 55: MOOViz tool in transportation scenario – disabled objective 85

Figure 56: MOOViz tool in transportation scenario – rearranging order of objectives – in process.. 86

Figure 57: MOOViz tool in transportation scenario – rearranging order of objectives – after reorder.. 86

Figure 58: MOOViz tool in transportation scenario – details of policy 87

Figure 59: MOOViz tool in transportation scenario – tooltip window details 87

Figure 60: MOOViz tool in transportation scenario – policy added to favorites 88

Figure 61: MOOViz tool in transportation scenario – policy highlighted............................... 89

Figure 62: MOOViz tool in transportation scenario – policy marked as decision 89

Figure 63: MOOViz tool in transportation scenario – decision panel 90

Figure 64: MOOViz tool in transportation scenario – delete decision example 90

Figure 65: MOOViz tool in transportation scenario – range sliders....................................... 90

Figure 66: MOOViz tool in transportation scenario – optimal and auto-excluded tabs 91

Figure 67: MOOViz tool in transportation scenario – optimal policies tab............................ 91

Figure 68: MOOViz tool in transportation scenario – example of bold policy line when mouse is placed over an optimal policy name 92

Figure 69: MOOViz tool in transportation scenario – adding policy to favorites from optimal policies tab.. 92

Figure 70: MOOViz tool in transportation scenario – example of notification on existing appealing policies 93

Figure 71: MOOViz tool in transportation scenario – highlighted policy in the optimal policies tab.. 93

Figure 72: MOOViz tool in transportation scenario – excluded policy in the optimal policies tab.. 93

Figure 73: MOOViz tool in transportation scenario – decided policy in the optimal policies tab ... 93

Figure 74: MOOViz tool in transportation scenario – combination of excluded, highlighted, decided and favorite policy ... 93

Figure 75 MOOViz tool in transportation scenario – top left panel ... 94

Figure 76: MOOViz tool in transportation scenario – moving from lines view to polygon ... 94

Figure 77: MOOViz tool in transportation scenario – objective visualization ... 95

Figure 78: MOOViz tool in transportation scenario – optimization direction ... 95

Figure 79: MOOViz tool in transportation scenario – disabled objective ... 95

Figure 80: MOOViz tool in transportation scenario – policy glyph ... 95

Figure 81: MOOViz tool in transportation scenario – referring slice to objective ... 96

Figure 82: MOOViz tool in transportation scenario – details on policy ... 96

Figure 83: MOOViz tool in transportation scenario – tooltip window details ... 96

Figure 84: MOOViz tool in transportation scenario – favorites pane ... 97

Figure 85: MOOViz tool in transportation scenario – favorites view ... 97

Figure 86: MOOViz tool in transportation scenario – example for more appealing policies .. 98

Figure 87: MOOViz tool in transportation scenario – tradeoff analyzer ... 98

Figure 88: MOOViz tool in transportation scenario – zoom in button ... 99

Figure 89: MOOViz tool in transportation scenario – snapshot before zoom in ... 99

Figure 90: MOOViz tool in transportation scenario – snapshot after zoom in ... 100

Figure 91: MOOViz tool in transportation scenario – link to zoom in view ... 100

Figure 92: MOOViz tool in transportation scenario – link to nested zoom in ... 100

Figure 93: Consensus Game Architecture ... 104

Figure 94: Experiment 1 Contour Variations: (from left to right) star glyph with rays and no contour (D); common star glyph (D+C); only the contour line of the star glyph (C) [145]. .. 112

Figure 95: Experiment Setting: The participant was seated in front of a 24" screen with a resolution of 1920x1200. The only input device was a computer mouse. ... 113

Figure 96: Experiment Setting: For each trial glyphs were arranged in a 3x3 matrix. The stimulus is highlighted and positioned in the middle to assure an equal distance to the other glyphs. This setting is used in all three experiments. ... 115

Figure 97: Experiment 1 Summary: The bar charts illustrated the percentage of correct answers and the standard deviation. 116

Figure 98: Experiment 1 Results: The bar charts illustrate the percentage of selections and the standard deviation for each factor. In the high dimensional condition experts using variations D+C, C were lead to judge shape similarity rather than data similarity whereas the accuracy of novices was low for all three variations. 117

Figure 99: Experiment 2 design space: We enriched the design space from our previous study by adding a “fill” version of the star glyph. The design variations of the first study (i.e., D, D+C, C) are applied to both Fill and No-Fill. 118

Figure 100: Experiment 2 results: The bar charts illustrate the percentage of selections and the standard deviation for each factor. The left chart represents low dimensionality, the right one the high dimension condition. Even without training or explaining the visual encoding participants using variation D judged data similarity rather than shape similarity. 120

Figure 101: Experiment 3 design space: We have chosen the star glyph only with data whiskers (D) and with an additional contour line (D+C) and applied tickmarks (T) and gridlines (G) to these designs. 123

Figure 102: Experiment 3 results of the percentage of selections and the standard deviation for each factor. Design improvements (T, G) do not significantly increase the accuracy of the two star glyph variations (D+C, C). 126

Figure 103: Visual Alignment for scatterplot matrices: Common scatterplot with one data point highlighted as point of interest (left). After selecting this data point it is considered the new baseline and smoothly moved to the center. All other elements are repositioned accordingly (right) 129

Figure 104: Visual Alignment for glyph designs: Color saturation is used to encode the data value for each dimension (top). After selecting a point of interest the values of the glyph are considered as new baseline, thus, all elements change their coloring to fit the new baseline (bottom) 130

Figure 105: Inverse Functionality: The scatterplot matrix visualizes the three dimensions Bio Diversity, CO2 and Cost Food (left). CO2 and Cost Food are marked as inverse since a higher value has a negative meaning. Therefore, the visualization changes showing this fact (right). 131

Figure 106: Inferior Function: Each glyph represents a multi-dimensional data point (left). Data points, which have lower data values for each dimension compared to the others are considered as inferior. These alternatives need not be considered in the final analysis and are, therefore, hidden (right). 132

Figure 107: As a default setting glyphs are positioned in a matrix (left). By applying a PCA the analyst can detect two outliers and two main clusters (right). 133

Figure 108: Visual Analytics Prototype: Default setting showing a scatterplot matrix with no further algorithms applied. 134

Figure 109: Dimension Swapping: Several dimensions can be selected via checkboxes to swap their values.. 134

Figure 110: Visual Alignment: After selecting the upper right data point in the left scatterplot the visual alignment technique is applied arranging all data points according to the new baseline. 135

List of Tables

Table 1: Comparison of MODM and MADM 27

Table 2: Biofuel criteria and importance rank. Adapted from Buchholz, Luzadis and Volk (2009) 29

Table 3: BFD criteria and sustainability condition. Adapted from Hayashe, Ierland and Zhu[62] 30

Table 4: MOOViz inputs – a domain-definition containing three objectives and a corresponding scenario containing four policies 56

Referenced project deliverables

D2.1.1: User requirements, ERF, Consensus Project Report, Confidential, available only to project members and review panel

D4.1.1: Optimization and Visual analytics Prototypes, IBM, Consensus Project Prototypes, Confidential, available only to project members and review panel

D2.2: State of the art report, ERF, Consensus Project Report, Public, available at: http://consensus-project.eu

D2.3: Domain Data Sources, ICCS/NTUA, Confidential, available only to project members and review panel

D3.2.1: Models and Simulators Report, IIASA, Consensus Project Report, Public, available at http://www.consensus-project.eu/deliverables

D2.4.1: System Architecture, NTUA, Consensus Project Report, Confidential, available only to project members and review panel

1 Introduction

Almost every real-life policy determination problem encountered is actually a multi-objective problem. Multi-objective decisions are made implicitly and in most cases, people are not specifically aware they are solving a multi-objective problem. Some decision situations, however, cannot be solved on the basis of casual intuition for a variety reasons. For example, some cases might involve substantial consequences, long-term impacts affecting many people, irreversibility, uncorrected mistakes, or a large number of alternatives. In such cases, a policy decision support framework is necessary. Note that such problems exist in almost any policy implementation sector.

Policy decision makers are faced daily with different policy choices and objectives that, more often than not, are subject to inherent conflicts, implying underlying trade-offs that must be taken into account. Under these circumstances, some form of decision-making aid is required to help decision makers in preparing and making their decisions and to study decision problems in which more than one point of view must be considered.

The Consensus project strives to support policy decision makers throughout the steps of the policy decision-making lifecycle, through a multidisciplinary partnership among experts from the fields of operational research, decision science, social technologies (gamification, crowdsourcing, and social analytics), applied system analysis, and visual analytics.

The developed framework will be validated through the modelling and evaluation of two real-world (complex) policy decision scenarios—biofuel and transport.

1.1 Objectives

The objective of D4.2.1 is to report the research and technical development (RTD) work towards the development of the Consensus tools that was done within WP4: Optimization and Visual Analytics. This report accompanies the software prototypes delivered within D4.1.1.

This report presents the scientific challenges, research, and innovations, as well as the technical implementation notes of the prototypes developed.

Note that pilot testing of the prototypes for evaluating the tools' capabilities will be executed within WP5. Therefore, this deliverable will be evaluated during the first project iteration, and an updated revision will be delivered (i.e. D4.2.2) during the second year of the project to develop the final tools and technologies that, once integrated, will implement the Consensus vision.

1.2 Scope

The scope of this report is demonstrated using the following Consensus components' diagram:

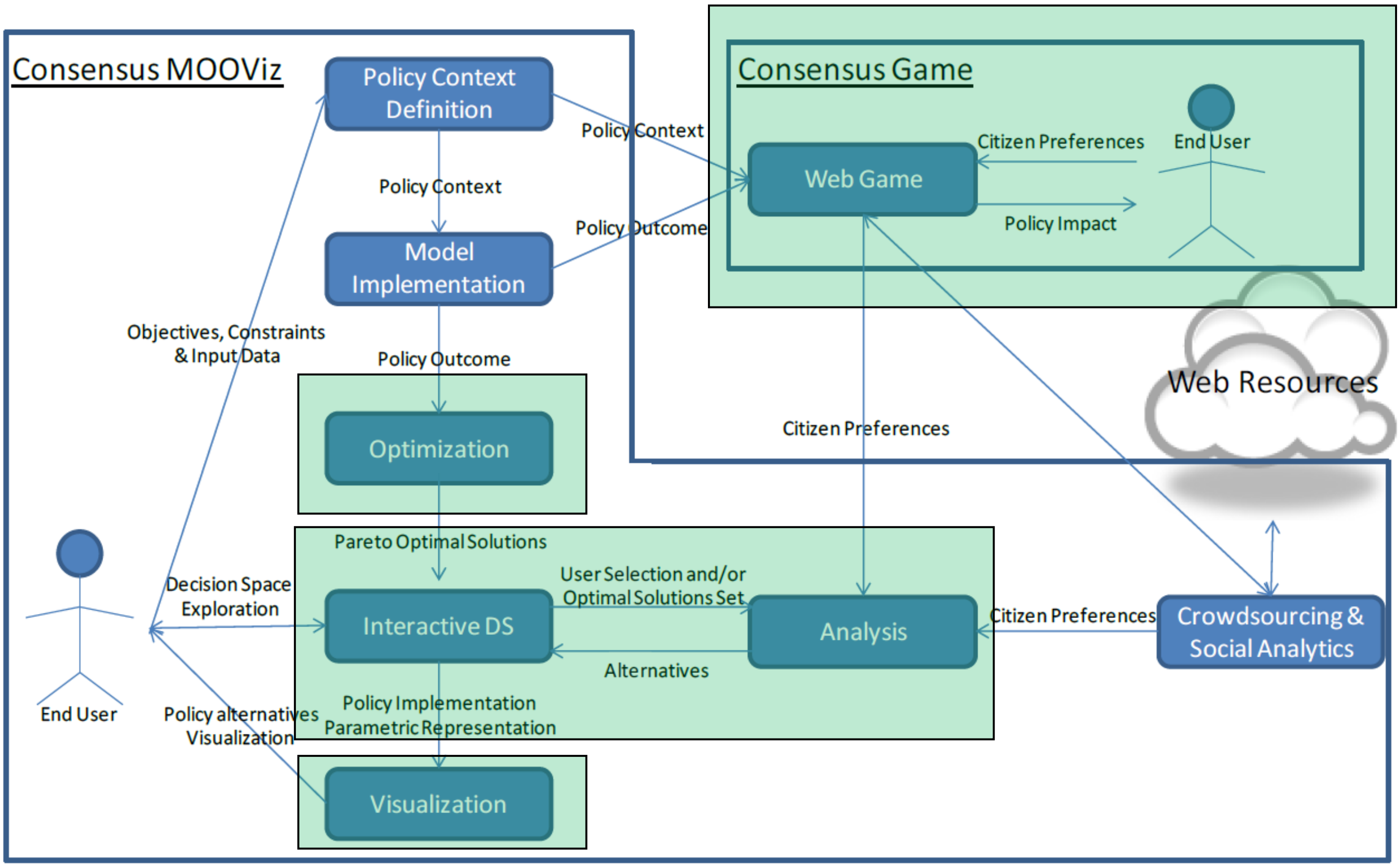


Figure 1: Consensus components' diagram

The highlighted green boxes represent components in the policy decision framework that are being researched and developed as part of WP4 and are within the scope of this report: optimization, interactive decision support, trade-off analysis, visualization, and Consensus Game. The internal focus and effort among the various components were set according to the project needs as derived from the D2.1.1: User Requirements report .

1.3 Methodology

This is the first report of three revisions of the Optimization and Visual analytics report. In this paper, we present and summarize the RTD work conducted during project months M1 to M12 (October 2013-September 2014) under WP4 of the Consensus project.

This report accompanies the respective prototypes deliverable and presents the scientific research, technical development, and implementations details that took place along the path for constructing the prototypes. The theoretical background for this work is based on the State of the Art Report, which details the previous scientific basis for the research work done. The framework is addressing policy decision makers' requirements as reflected in the User Requirements report.

The data sources used for the algorithms and tools are specified in the Domain Data sources report. Note that the major data sources for the optimization and visualization components are based on output produced by WP3 and are fully described in the Models and Simulators Report.

System architecture considerations that were applied for constructing the prototypes are specified in detail within the System Architecture report.

The major effort of WP4 is associated with RTD challenges, as reflected in the work performed for creating the Optimization and Visual Analytics Prototypes and reports (specifically, NTUA, IBM, IIASA, ATC, and UKON). The end user partners' contributions (ERF, OXFAM, WWF) will close the development cycle in WP5 Evaluation, in which valuable input will be gathered for improving the systems towards the next revisions.

1.4 Structure

This report is organized into six chapters.

Chapter 1 is a general introduction of the project.

Chapter 2 describes the scientific background. It addresses various aspects of multi-objective optimization; decision support, both general and specifically for the Consensus project domains (Environmental and Transport); visual analytics; gamification (with an emphasis on reward models); and crowdsourcing.

Chapter 3 deals with the GLOBIOM optimization model, in which the global forestry and agriculture market equilibrium is determined by choosing economic activities subject to resource, technological demand, and policy constraints to maximize social welfare.

In **Chapter 4**, we present the Consensus Multi-Objective Optimization and Visualization Tool (MOOViz). This is a framework intended for policy decision makers to assist in exploring alternative policy implementations, understanding trade-offs, and consciously determining the optimal policy.

Chapter 5 is dedicated to the Consensus Game—a web tool intended for the public and aimed at education, collaboration, and communicating policy decision conflicts to citizens as well as enabling citizens to express their preferences with regards to the questioned policy domains. The acquired citizen's feedback will be taken into account as an additional decision input.

The final chapter, **Chapter 6** Visual Analytics, focuses on visual support, interaction possibilities, and automatic algorithms that are essential for augmenting the capabilities in the decision cycle. The scientific approaches developed are accompanied by an online prototype for supporting and demonstrating the concepts.

1.5 Quality Management

The D4.2.1 document has been structured, compiled, and edited by the WP4 leader IBM to ensure the compliance of the document to the Consensus project's particular deliverable format. The content provider partners have sent the sections relevant to their

responsibilities to the editor and the documents have been merged.
The review process was conducted in three steps. The first step was to provide feedback to the general structure and to the draft content of the document. The second step of the review was to edit and to give feedback for the core parts of the document (scientific background, research, and technical details). The third and final step was achieved by the assigned reviewers of D4.2.1. This step was to provide feedback on the structure, clarity, and consistency of the document.

urthermore, the designed and implemented concepts and prototypes will be evaluated and further improved as part of the evaluation process conducted in WP5.

2 Background

2.1 Introduction

Chapter 2 provides scientific background on the relevant aspects of multi-objective optimization, decision-making, visual analytics and gamification, and on the research done on implementing the Consensus framework for tackling policy decision-making challenges.

Section 2.2 presents a variety of approaches for multi-objective optimization. Sections 2.3 and 2.4 dive into the relevant Consensus use-case domains, Multi-Criteria Decision Making in the Environmental Sector and Multi-Criteria Decision Making in the Transport Sector.

Section 2.5 describes visual analytics concepts that are tightly coupled with data mining and visualization approaches, helping to make sense of data and find appropriate decisions.

Section 2.6 presents the concepts that are used within this project to bring policy decision-making to the citizens' level through gamification and to capture the public preferences and incorporate them as additional decision input (crowdsourcing).

2.2 Multi-Objective Optimization

2.2.1 Introduction

Multi-objective optimization plays a major role in real-world decision problems. It aims at simultaneously optimizing a number of conflicting objectives, thereby explicitly considering multiple criteria in the decision-making process. One such example could be selecting a public policy that maximizes efficiency in achieving its goals while minimizing tax-payers' expenditures and negative environmental effects. This case represents a nontrivial multi-objective optimization problem in which no single solution can simultaneously optimize all the objectives. In such cases, the objective functions are said to be conflicting, and a (possibly infinite number of) Pareto optimal solutions exist. These solutions are called non-dominated, Pareto optimal, Pareto efficient, or non-inferior. Without additional subjective preference information, all the Pareto optimal solutions are considered equivalent.

In the context of Consensus, several challenges exist. The first is finding a diverse set of alternative efficient policies, providing a means for decision makers and for the public to understand the trade-offs among the variety of alternatives and different objectives (aligned or conflicting, dependent and independent of one another, etc.). Other challenges include how to elicit preferences, advise recommendations, and measure and integrate crowd opinions regarding alternative plans.

A major challenge under this framework is revealing the Pareto optimal set or a region of interest in the trade-off surface among the objectives. This framework is not limited to traditional optimization approaches that consider multi-objective problems by posing a weighted sum of its objectives and employing single-objective optimization to solve them[1].

Common approach for solving the multi-objective optimization problem are methods which applying several scalarizations; the solution to each scalarization yields a Pareto optimal solution, whether locally or globally. The scalarizations are constructed with the target of obtaining evenly distributed Pareto points that give a diverse, evenly distributed approximation of the real set of Pareto points. Examples are the Normal Boundary Intersection (NBI)http://en.wikipedia.org/wiki/Multi-objective_optimization - cite_note-16[2], Modified Normal Boundary Intersection (NBIm)[3], Normal Constraint (NC)[4][5] Successive Pareto Optimization (SPO)[6] and Directed Search Domain (DSD)[7].

Evolutionary Algorithms (EAs)[8], powerful stochastic global search methods gleaned from the model of organic evolution, have been successful in treating high dimensional optimization problems for several decades. They especially excel in scenarios where quality evaluation provided by computer-based simulation constitutes the objective function, also referred to as simulation-based optimization[9]. Their broad success in this domain is primarily attributed to two factors - first, the fact that they constitute direct search methods, i.e., do not require derivatives determination, and second, their inherent robustness to noise [10]. In the last two decades evolutionary multi-objective optimization algorithms (EMOA) have undergone considerable development [11][12].

Most evolutionary multi-objective optimization algorithms apply Pareto-based ranking schemes. The main advantage of evolutionary algorithms, when applied to solve multi-objective optimization problems, is the fact that they typically generate sets of solutions, allowing computation of an approximation of the entire Pareto front at once. The main disadvantage of evolutionary algorithms is their lower speed and the fact that Pareto optimality of the solutions cannot be guaranteed. Examples for EMO methods are Non-dominated Sorting Genetic Algorithm-II (NSGA-II), Strength Pareto Evolutionary Algorithm 2 (SPEA-2) and methods based on particle swarm optimization and simulated annealing[13]. Among other a-posteriori methods include: PGEN (Pareto surface generation for convex multi-objective instances)[14], IOSO (Indirect Optimization on the basis of Self-Organization), SMS-EMOA (S-metric selection evolutionary multi-objective algorithm)[15], Reactive Search Optimization (using machine learning for adapting strategies and objectives),[16][17], Benson's algorithm for linear vector optimization problems.

2.2.2 Formulation

Let a vector of objective functions in $R^m, \bar{f}_{(\bar{x})} = (f_{1(\bar{x})}, f_{2(\bar{x})}, \cdots, f_{m(\bar{x})})^T$, be subject to minimization, and let a partial order be defined in the following manner. Given any $f^{(1)} \in R^m$, and $f^{(2)} \in R^m$, we state that $f^{(1)}$ strictly Pareto dominates $f^{(2)}$, which is denoted as $f^{(1)}$ $f^{(1)} \prec f^{(2)}$ $f^{(2)}$, if and only if $\forall i \in \{1, \dots, m\}: f_i^{(1)} \leq f_i^{(2)} \wedge \exists i \in \{1, \dots, m\}: f_i^{(1)} < f_i^{(2)}$. The individual Pareto-ranking of a given candidate solution is defined as the number of other solutions dominating it. The crucial claim is that for any compact subset of R^m, there exists a non-empty set of minimal elements with respect to the partial order ≼ (see, e.g.,[18]). Non-dominated points are then defined as the set of minimal elements with respect to the partial order ≼, and by definition their Pareto-ranking is zero. The goal of Pareto optimization is thus to obtain the non-dominated set and its pre-image in the design space, the so-called Pareto optimal set, also referred to as the efficient set.

The Efficient (Pareto) Frontier F is defined as the set of all points in the objective space that correspond to the solutions in the Pareto optimal set. The set that is jointly dominated by F

but is not dominated by any other solution has Pareto-ranking 1, and so goes the ranking for subsequently dominated sets; following this notion the ranking of each solution can be defined (see, e.g.,[19]).

The computational problem of attaining the Pareto Frontier of a multi-objective optimization problem[20] can be treated by means of algorithms utilizing mathematical programming solvers (e.g., the so-called Diversity Maximization Approach[21] employing IBM's ILOG-CPLEX [22]), or alternatively, approximated by population-based heuristics. The wide applicability of Pareto-driven optimization is evident in the vast number of published work - see, e.g., [23], [24]. The crucial claim is that many real-world problems are inherently multi-objective in nature. This concept ranges from Combustion Processes[25], Yeast Fermentations[26] and Photo induced Processes[27] to potentially as far as to Theory Choice (see[28] for the broad overview, and[29] for the explicit multi-criterion perspective).

2.2.3 Approaches

Among the goals of Consensus, is to develop a decision aiding system which through the use of models, helps obtain elements of responses to the questions posed by a stakeholder of the policy decision process. The system shall work towards clarifying the decision and towards recommending, or simply favoring, a policy that is Pareto efficient and will increase the consistency between the selected policy and the stakeholder's objectives and value system[30].

The system shall analyze the multi-objective policy decision problems from different viewpoints; apply different solution philosophies and aims at setting and solving the decision problem.

The goals of the system are:

- To find a representative set of Pareto optimal policies
- Quantify/Visualize the trade-offs in satisfying the different objectives
- Finding a single policy (or a subset of policies) that satisfies the subjective preferences of a human decision maker (DM) or satisfying the aggregated preferences of decision makers group

Amongst the approaches being examined are **Scalarizing methods** that convert the original problem with multiple objectives into a single-objective decision problem, **no-preference methods** that requires no preference information to be articulated by the decision maker, **A priori methods** that require sufficient decision maker preference information to be expressed before the solution process, **A posteriori methods** that aim at producing all the Pareto optimal solutions and **interactive methods**, in which the decision maker iteratively interacts with the system during the solution process[31]. Hybrid methods combine more than a single approach.

There are four classes of multi-objective optimization approaches - Each class of methods involves DM preference information in different ways (No Preference/A-priori/A posteriori/Interactive)

In no preference methods, no decision maker (DM) is expected to be available, but a neutral compromise solution is identified without preference information. In a priori methods, preference information is first asked from the DM and then a solution best satisfying these preferences is found. In a posteriori methods, a representative set of Pareto optimal solutions is first found and then the DM must choose one of them. In interactive methods, the decision maker is allowed to iteratively search for the most preferred solution. In each iteration of the interactive method, the DM is shown Pareto optimal solution(s) and describes how the solution(s) could be improved. The information given by the decision maker is then taken into account while generating new Pareto optimal solution(s) for the DM to study in the next iteration. In this way, the DM learns about the feasibility of her wishes and can concentrate on solutions that are interesting to her. The DM may stop the search whenever he/she wants to.

2.2.3.1 Scalarizing Multi-Objective Optimization Problems

One approach considered under this framework, is the approach of applying **Scalarizing methods**, in which we convert the original problem with multiple objectives into a single-objective optimization problem. This means formulating a single-objective optimization problem such that optimal solutions to the single-objective optimization problem are Pareto optimal solutions to the multi-objective optimization problem[31]. Applying this formulation, it is often required that every Pareto optimal solution can be reached with some parameters of the scalarization[31]. And naturally, With different parameters for the scalarization, different Pareto optimal solutions are selected.

A well-known example is **linear scalarization** (also known as weighted sum)

$$\min_{x\in X} \sum_{i=1}^{k} w_i f_i(x),$$

Where the weights of the objectives $w_i > 0$ are the parameters of the scalarization.

Pay attention that the weights in this representation may be used for both to representing the DM preferences as well as for scaling the dimensions of different objectives.

And the **ϵ-constraint method** (see, e.g.[32])

$$\begin{array}{ll} \min & f_j(x) \\ \text{s.t.} & x \in X \\ & f_i(x) \le \epsilon_j \text{ for } i \in \{1,\ldots,k\} \setminus \{j\}, \end{array}$$

where upper bounds ϵ_j are parameters as above and f_j is the objective to be minimized.

Another examples are Goal Programming and **Achievement scalarizing problems**[33]. They can be formulated as

$$\begin{array}{ll} \min & \max_{i=1,\ldots,k} \left[\frac{f_i(x)-\bar{z}_i}{z_i^{\text{nad}}-z_i^{\text{utopia}}} \right] + \rho \sum_{i=1}^{k} \frac{f_i(x)}{z_i^{nad}-z_i^{\text{utopian}}} \\ \text{subject to} & x \in S, \end{array}$$

where the term $\rho\sum_{i=1}^{k}\frac{f_i(x)}{z_i^{nad}-z_i^{\text{utopia}}}$ is called the augmentation term, $\rho>0$ is a small constant, and z^{nad} and z^{utopian} are the nadir vector (the upper bound of the Pareto optimal set) and an utopian (an infeasible objective vectors which is ideal across all objectives) vectors, respectively. In the above problem, the parameter is the so-called reference point $\bar{z}$ which represents objective function values preferred by the decision maker.

2.2.3.2 No Preference Methods

Another approach is using multi-objective optimization methods that do not require any preference information to be explicitly articulated by a decision maker. Those methods can be classified as ***no-preference methods***[31]. A well-known example is the method of global criterion[34], in which a scalarized problem of the form

$$\begin{aligned}&\min\|f(x)-z^{ideal}\|\\&\text{s.t. } x\in X\end{aligned}$$

is solved. $\|\cdot\|$ can be any L_p norm, with common choices including L_1, L_2 and L_∞ [32]. The method of global criterion is sensitive to the scaling of the objective functions, and thus, it is recommended that the objectives are normalized into a uniform, dimensionless scale

2.2.3.3 A priori Methods

A priori methods require that sufficient preference information is expressed before the solution process[31]. Well-known examples of a priori methods include the **utility function method**, lexicographic method, and goal programming.

In the utility function method, it is assumed that the decision maker's utility function is available. A mapping $u:Y\rightarrow\mathbb{R}$ is a utility function if for all $\mathbf{y}^1,\mathbf{y}^2\in Y$ it holds that $u(\mathbf{y}^1)>u(\mathbf{y}^2)$ if the decision maker prefers $\mathbf{y}^1$ to $\mathbf{y}^2$, and $u(\mathbf{y}^1)=u(\mathbf{y}^2)$ if the decision maker is indifferent between $\mathbf{y}^1$ and $\mathbf{y}^2$. The utility function specifies an ordering of the decision vectors (recall that vectors can be ordered in many different ways). Once u is obtained, it suffices to solve $\max\ u(\mathbf{f}(\mathbf{x}))\text{ subject to }\mathbf{x}\in X,$

But in practice it is very difficult to construct a utility function that would accurately represent the decision maker's preferences[32] - particularly since the Pareto front is unknown before the optimization begins. Lexicographic method assumes that the objectives can be ranked in the order of importance. We can assume, without loss of generality, that the objective functions are in the order of importance so that f_1 is the most important and f_k is the least important to the decision maker. The lexicographic method consists of solving a sequence of single-objective optimization problems of the form

$$\begin{aligned}&\min f_l(\mathbf{x})\\&\text{s.t. } f_j(\mathbf{x})\le\mathbf{y}_j^*,\ j=1,\ldots,l-1,\\&\qquad\mathbf{x}\in X,\end{aligned}$$

Where $\mathbf{y}_j^*$ is the optimal value of the above problem with $l=j$. Thus $\mathbf{y}_1^* := \min\{f_1(\mathbf{x}) \mid \mathbf{x} \in X\}$, and each new problem of the form in the above problem in the sequence adds one new constraint as l goes from 1 to k.

2.2.3.4 A Posteriori Methods

A posteriori methods aims at producing all the Pareto optimal solutions (known as the "Pareto Frontier") or a representative subset of the Pareto Frontier. Then, applying preferences to select a solution from the resulted set. The posteriori preferences techniques implemented in this project include three steps:

1. Computer approximates the Pareto front (i.e. the Pareto optimal set in the objective space)

2. The decision maker explores and studies the Pareto front approximation

3. The decision maker identifies the preferred point (or the preferred regions) at the Pareto front

From the point of view of the decision maker, the step of exploring and understanding the Pareto front is the most complicated one.

In the case of bi-objective problems, the Pareto front, (also named the "Tradeoff Curve" in this case), can be drawn at the objective plane. It gives the decision maker full information on objective values and on objective tradeoffs, which inform how improving one objective is related to deteriorating the second one while moving along the tradeoff curve. The decision maker takes this information into account while specifying the preferred Pareto optimal objective point[35]. Bi-objective problems are well studied but in this project we were focusing on decision problems comprise of three or more objectives, for which a simple visual representation of the Pareto front cannot be provided to the user. Exploration of the Pareto front in higher dimensions is a non-trivial task and is a major challenge of this project.

2.2.3.5 Interactive Methods

When applying interactive methods, the decision making process is iterative and the decision maker continuously interacts with the method while searching for the most preferred policy (see e.g. [32][36]). Practically, the decision maker express preferences at each iteration in order to get Pareto optimal solutions that are of interest to her and learn the trade-offs between attainable solutions. The following steps are commonly present in interactive methods:[36]

1. Initialize
2. Generate a Pareto optimal starting point (by using e.g. some no-preference method or solution given by the decision maker)
3. Ask for preference information from the decision maker
4. Generate new Pareto optimal solution(s) according to the preferences and show it/them and possibly some other information about the problem to the decision maker
5. If several solutions were generated, ask the decision maker to select the best solution so far
6. Stop, if the decision maker wants to; otherwise, go to step 3

Instead of mathematical convergence that is often used as a stopping criterion in mathematical optimization methods, a psychological convergence is emphasized in interactive methods. Generally speaking, a method is terminated when the decision maker is confident that she has found the most preferred solution available.

Different interactive methods involve different types of preference information. For example, three types of methods can be identified; based on:

- **trade-off information**: the decision maker is shown several objective trade-offs at each iteration, and she is expected to say whether she likes, dislikes or is indifferent with respect to each trade-off (e.g. the Zionts-Wallenius method,[37]).
- **reference points**: the decision maker is expected at each iteration to specify a reference point consisting of desired values for each objective and a corresponding Pareto optimal solution(s) is then computed and shown to her for analysis. (see e.g., [1],[38]).
- **classification of objective functions**[36]: the decision maker is assumed to give preferences in the form of classifying objectives' values at the current Pareto optimal solution into different classes indicating how the values of the objectives should be changed to get a more preferred solution - for example objectives whose values a) should be improved, b) can be relaxed, and c) are acceptable as such. Then, the classification information given is taken into account when new (more preferred) Pareto optimal solution(s) are computed (see e.g. satisfying trade-off method (STOM) [39]and the NIMBUS method,[40][41]).
- **Selection between a small sample of solutions**[42][43].

2.2.3.6 Preference Elicitation

Another major challenge within the decision process is the elicitation of the preferences or in other words, the utility embedded in each of the alternatives, note that this is a more general approach than ranking or weighting the different criteria, as the tradeoffs and constraints between different objectives may vary across the manifold.

For example, Conjoint analysis is a statistical technique used to determine how people value different features that make up an individual alternative. The objective of conjoint analysis is to determine what combination of a limited number of attributes is most influential on respondent choice or decision making. A controlled set of potential alternatives is shown to respondents and by analyzing how they make preferences between these alternatives, the implicit valuation of the individual elements making up an alternative can be determined. These implicit valuations (utilities or part-worths) can be used to create models for trade-off elicitation.

Conjoint originated in mathematical psychology and was developed by marketing professor Paul Green at the University of Pennsylvania and Data Chan. Other prominent conjoint analysis pioneers include professor V. “Seenu” Srinivasan of Stanford University who developed a linear programming (LINMAP) procedure for rank ordered data as well as a self-explicated approach, Richard Johnson (founder of Sawtooth Software) who developed the Adaptive Conjoint Analysis technique in the 1980s and Jordan Louviere (University of Iowa) who invented and developed Choice-based approaches to conjoint analysis and related techniques such as MaxDiff. Conjoint analysis techniques may also be referred to as multiattribute compositional modelling, discrete choice modelling, or stated preference research[44].

Peter Fishburn is another fundamental contributor to this area in the context of the theory of social choice and utility[45][46]. In many circumstances when trying to analyze decision maker’s preferences, a political challenge exists as well, such example is an interview technique to elicit person's utility function which was developed by Ragner Frisch. An attempt to apply this method to the Norwegian Parliament failed, due to the reluctant of the Parliament members to make their utility function explicit[47].

2.3 Multi-Criteria Decision Making in the Environmental Sector

2.3.1 Introduction

Multi-criteria analysis (MCA) also called multi-criteria decision making (MCDM) or multi-criteria decision analysis (MCDA) has been increasingly used in various sectors, including environment in the recent years (Steele et al., 2009[48]). The group of methods described by MCA can be defined as ‘formal approaches which seek to take explicit account of multiple criteria in helping individuals and groups explore decisions that matter’ (Belton and Stewart 2002[49]). Its merits have been recognised by those individuals, companies or decision makers that are facing complex decisions with multiple variables. The UK Government has recognised its usefulness by issued a manual specifically designed for institutions belonging to the local government.

The environmental sector has also embraced MCA, mainly because there is still a lack of guidance on aiding environmental decision making (Omman, 2000[50]). Balasubromiam and Voulvoulis[51] note that “MCA can be particularly appropriate when the decision-making context is characterized by multiple objectives and multiple criteria, incommensurable criteria, mixed data and the need for ease of use, and the analysis context is characterized by multiple participants.”

Conflicts related to land use and land management are getting more frequent and more serious (Joerin and Musy 2000[52]). Demand for natural resources, food and fibre has been steadily growing in line with population growth and increased purchase power in developing countries. The European biofuel legislation is a typical example for a complex policy with various objectives that can be in direct competition with other objectives of the European community. Stakeholders involved in the discussions related to biofuel sustainability were unable to negotiate a compromise solution for addressing Indirect land use change (ILUC). In this context the topic lends itself to be analysed through an MCA approach.

Studies such as that by Mendoza and Martins[53] show that MCA offers a sound and robust approach to planning and decision-making for natural resources management by developing a clear set of criteria, balancing social, economic and environmental aspects of complex problems. Further arguments are presented below.

2.3.2 Arguments for Using MCA

Users of the MCA approach list a number of well grounded arguments in support of MCA especially when considering alternatives such as cost benefit analysis (CBA). These include:

- Non-market valuation data (revealed and stated preference) may not be readily available or expensive to collect
- It may not be able to present some impacts of policy in a way that can be traded-off for money - practical or moral reasons
- It may not be able to quantify impacts, e.g. diffuse social impacts such as social cohesion
- CBA may not account for interactions of impacts, e.g. synergy

Users of MCA implement various techniques (particularly mathematical programming techniques) but all have common thread: they recognise the existence of multiple judgement or evaluation criteria since any plan, policy or project is likely to have different but simultaneous impacts, their evaluation requires simultaneous assessment from different perspectives (Zhang et al, 2012[54]).

2.3.3 Types of MCAs Applied in the Environmental Sector:

Multi-criteria methods can essentially be split up into two broad categories:

- Discrete Multi Criteria methods (DMCM) - consider a finite number of feasible choice possibilities (alternative plans of action, alternative objectives/decision criteria), also known as Multi Objective Decision Making (MODM).

- Continuous Multi Criteria methods (CMCM) - consider an infinite number of feasible choice possibilities, also known as Multi Attribute Decision Making (MADM)

Continuous MC methods lend themselves more to economic evaluation where financial measures can be broken down ad infinitum to represent alternative strategies.

A summary of the various approaches to MCA can be found in Table 1.

Table 1: Comparison of MODM and MADM

Criteria for comparison	MODM	MADM
Criteria defined by	Objective	Attributes
Objective defined	Explicitly	Implicitly
Attributes defined	Implicitly	Explicitly
Constraints defined	Explicitly	Implicitly
Alternatives defined	Implicitly	Explicitly
Number of alternatives	Infinite (large)	Finite (small)
Decision makers control	Significant	Limited
Decision modelling paradigm	Process-oriented	Outcome-oriented
Relevant to	Design/search	Evaluation/choice

*Adapted from Mendoza and Martins 2006[53]

Discrete MC methods are of more use when we are trying to decide between a fixed number of specific plans/policies. They allow us to focus more closely on the pertinent issues. DMCM allow us to classify, rank and thus decide between alternative choices or strategies which have multiple impacting factors (criteria).

2.3.4 Common Stages in Applying MCA

The following stages are common when applying weighted sum MCA approach:

- Establish the decision context
- Identify the options to be appraised
- Identify objectives and criteria
- "Scoring": Assess the expected performance of each option against the criteria. Then assess the value associated with the consequences of each option for each criterion
- "Weighting": Assign weights for each of the criteria to reflect their relative importance to the decision
- Combine the weights and scores for each option to derive an overall value
- Examine the results
- Sensitivity analysis

2.3.5 Multi-Criteria Analysis and Biofuels

Over the past years several researchers have used the multi-criteria analysis framework to assess various aspects of the ongoing bioenergy debate. Studies looking at applying MCA to decisions around bioenergy systems point out that not only does it help to create a broad criteria for analysing sustainable attributes – largely missing from this arena (Buchholz, Luzadis and Volk 2009[55]), but it also helps with stakeholder integration (Buchholz et al 2009) and its participatory nature can increase the legitimacy of decisions (Ziolkowska 2013[56]). Other benefits include findings that viable bioenergy systems often rely on sound social criteria being considered at the conceptual stage (Buchholz et al 2009[55]).

Turcsin et al.[57] used the framework to assess stakeholder support for various biodiesel options in Belgium. The Consensus project will develop the ConsensusGame that is specifically focussing on exploring stakeholder support for various options. Perimenis et al.[58] used the MCA to develop a framework for decision makers. While Mohamadabadi et. al.[59] used this framework to rank various renewable and non-renewable energy sources. Buchholz et. al.[55] conducted a review of various MCA studies focussing on bioenergy and concluded that *"MCA tools should also be applied to more sophisticated case studies, with more scenarios, a larger scale, and more stakeholders."*

When conducting an MCA analysis, the selection of criteria is crucial to the robustness of the assessment. Key debates such as that of weak vs strong sustainability must be addressed during the selection process, but can be problematic (Myllyviita *et al* 2013[60]. A number of projects have developed criteria, and some have ranked these according to relative importance assigned by experts (Buchholz, Luzadis and Volk 2009[55]).

The selection of criteria can vary according to the specific biofuel system in question, as well as the region, and expertise represented within the stakeholder group (Buchholz, Luzadis and Volk 2009, Myllyviita *et al*[61]).

The following are a selection of criteria presented in two studies specifically designed for biofuel systems. In the case of Buchholz, Luzadis and Volk[55] identified 35 sustainability criteria regularly related to bioenergy sustainability and asked 137 experts to rank them.

Table 2: Biofuel criteria and importance rank. Adapted from Buchholz, Luzadis and Volk (2009)

Criteria	Environment/social/economic	Importance rank
Green house gas balance	Environmental	3.55
Energy balance	Environmental	3.44
Soil protection	Environmental	3.27
Participation	Social	3.16
Water management	Environmental	3.14
Natural resource efficiency	Environmental	3.11
Microeconomic sustainability	Economic	3.10
Compliance with laws	Social	3.09
Ecosystems protection	Environmental	3.07
Monitoring of criterial performance	Social	3.02
Food security	Social	2.95
Waste management	Environmental	2.93
Adaptation capacity to environmental hazards and climate change	Environmental	2.90
Crop diversity	Environmental	2.86
Working conditions of workers	Social	2.83
Planning	Social	2.79
Economic stability	Economic	2.79
Species protection	Environmental	2.76
Use of chemicals, pest control and fertilizer	Environmental	2.72
Potentially hazardous atmospheric emissions other than GHGs	Environmental	2.72
Employment generation	Economic	2.69
Property rights and rights of use	Social	2.68
Land use change	Environmental	2.68
Use of genetically modified organisms	Environmental	2.64
Ecosystem connectivity	Environmental	2.57
Respect for human rights	Social	2.48
Macroeconomic sustainability	Economic	2.39
Cultural acceptability	Social	2.37
Respecting minorities	Social	2.35
Exotic species applications	Environmental	2.33
Social cohesion	Social	2.26

Land availability for other human activities than food production	Social	2.25
Standard of living	Social	2.14
Noise impacts	Social	2.10
Visual impacts	Social	1.98

***A higher importance rank indicates experts feel this criteria is more relevant, practical, reliable or important than those with a lower score.**

Table 3: BFD criteria and sustainability condition. Adapted from Hayashe, Ierland and Zhu[62]

Criteria	Environment/social/economic	Sustainable/unsustainable
GHG emission	Environment	Sustainable
NOx emission	Environment	Sustainable
SOx emission	Environment	Sustainable
Wage for employment	Social	Unsustainable
Injury, illness fatality	Social	Sustainable
Production cost	Economic	Unsustainable
Gross value added	Economic	Unsustainable
Energy diversity	Economic	Sustainable

The Consensus project will be using a selection of these criteria, depending on the limitations of the land use models.

2.3.6 Further Recommendations

The Consensus project will combine an advanced multi criteria analysis framework, with state of the art land use modelling and the latest visualisation technology. The deliverables will enable decision makers to improve legislation, but also inform themselves about the citizen's view on various objectives and trade-offs related to the bioenergy.

While it is clear from the studies presented here that MCA is a robust and useful method to apply to decisions surrounding bioenergy systems, there are also methodological factors to be taken into consideration, and areas where further research is required.

Myllyviita et al[60] found, based on a Finnish case study, that the selection of criteria must take into account the specific system being assessed both in terms of the bioenergy system and the regional context. They also point out that the availability of relevant data can be limited and collection costly and time consuming. Data availability is a key concern however we are confident that the IIASA Globiom model will provide adequate quantity and quality of information.

The criteria selection was also discussed by Sliogerience et al[63] in relation to MCA of bioenergy systems in Lithuania. They point out that the relationship between criteria and the relative importance assigned must be considered at the outset to give a reliable assessment.

Further areas which must be more fully understood and the Consensus project will contribute include (Montibeller and Franco 2010[64]): the role of the decision analyst/ facilitator in balancing the view and objectives of stakeholders; using this approach to develop complex policies; and finally the long term consequences of this decision making approach.

2.4 Multi-Criteria Decision Making in the Transport Sector

2.4.1 Introduction

Transport Sector decisions affect almost all aspects of human life (mobility, health, safety, living costs, economic opportunities, conditions for work and leisure etc.); additionally, decision making is constantly required in the transport sector, from the strategic planning of projects and policies, the design of infrastructure works and the selection of alternatives, to the application of specific policy measures.

Thus, decision-making is an integral part of the management of transportation systems, that generally includes: identification of existing problems; problem definition (objectives, criteria, measures, constraints, etc.); generation of alternative solutions (options/ alternatives) for the problem (e.g. building new infrastructure, rehabilitating existing infrastructure, improving its management, applying policy measures etc.); and evaluation and selection of the best solution[65].

For years, the most common forms of evaluation in transport related decisions were cost-effectiveness analysis (CEA) and/or cost benefit analysis (CBA)[66]. However, both methods have certain limitations, which are primarily related to the difficulty to objectively and adequately value all the costs and impacts of the examined alternatives; additionally, in transportation projects the multiplicity of objectives lead most of the times in disagreements among the different involved actors about the scope of the project or the procedure to be followed[67].

To this end, Multi-Criteria Decision Making (MCDM) techniques seem to provide a more flexible and transparent way to find solutions to complex problems with various actors (stakeholders) and as such nowadays are broadly used in transport related decision-making.

2.4.2 General Procedure of Multi-Criteria Decision Making in Transport Sector

Despite the fact that every decision problem is different and that the detailed procedure for MCDM in transport sector can vary according to the characteristics of each problem, a general procedure for MCDM in transport is identified in relevant literature[66][68][69][70][71].

This general procedure is presented in Figure 2 below and it can be easily adapted to the requirements of each specific transport problem.

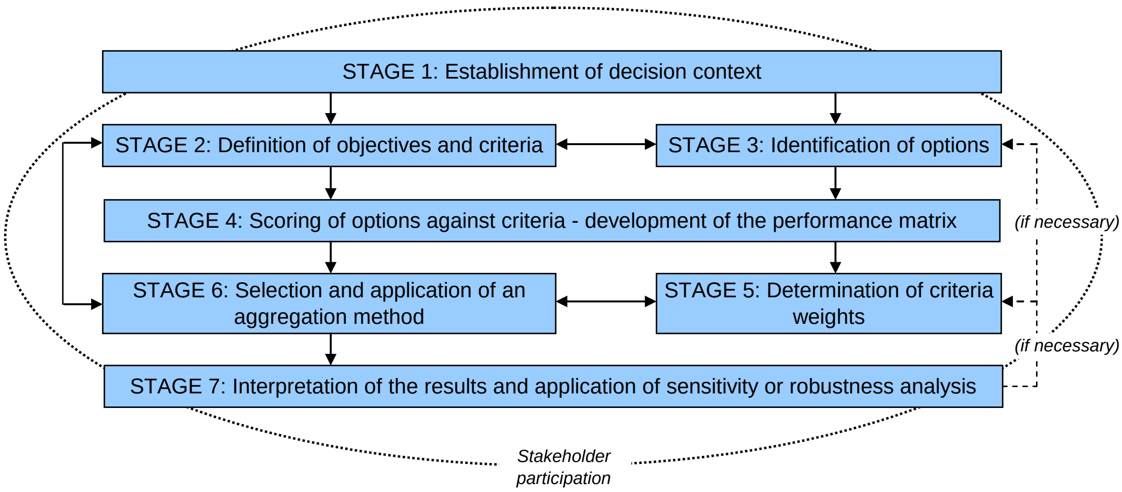


Figure 2: Procedure of MCDM in transport sector

The stages of the procedure presented above are not separate features but have linkages and effects upon each other. They do not necessarily follow a linear pattern, instead they sometimes run in parallel or it may be required to step back again (e.g. new criteria come up and have to be integrated into the analysis).

2.4.3 Decision Making Approaches in Transport Sector

Over time, three broad approaches have been developed in transport sector decision making[72]: Vision-led, Plan-led and Consensus-led.

Vision-led approaches usually involve an individual having a clear view of the future form of the transport system that is required, and the policy instruments needed to achieve that vision. The focus then is on implementing them as effectively as possible. It is obvious that this approach is critically dependent on the individual with the vision, and, most probably, if he/she leaves office, the strategy will be abandoned.

Plan-led approaches involve specifying objectives and problems, sometimes in the context of a vision statement, and following a certain procedure to identify possible solutions and select those that perform best. Problems are highlighted as failure of current or predicted future conditions to meet the objectives. This list of problems can then be discussed with stakeholders to see whether they have different perceptions of the problems. If they do, objectives are redefined accordingly. The main drawback with this approach is that many politicians and members of the public are less familiar with the abstract concept of objectives (e.g. improving accessibility) than they are with concrete problems (e.g. the nearest job centre being 50 minutes away). Also, a plan-led approach can become excessively dependent on professional planners / analysts, who may lose sight of the needs of decision makers and stakeholders.

Finally, **Consensus-led** approaches involve discussions between the stakeholders to try to reach agreement on each of the stages in the decision making process. Ideally agreement is needed on the objectives to be pursued and their relative importance, the problems to be tackled and their seriousness, the options (projects, policies or policy instruments) to be considered and their appropriateness, the selection of options which best meet the objective and the way in which they should be combined into an overall strategy and implemented. In practice much consensus-building focuses on the choice of options, but it can be considerably enhanced by considering objectives and problems as well. The main concern with the consensus-led approach is that, unless agreement can be quickly reached and sustained, it may result in serious delays or even inaction.

Since each of the above approaches has its advantages and drawbacks, in most cases **a mixed approach** is adopted, with most common a mix of plan-led and consensus-led decision-making[72].

2.4.4 Decision Making Subjects in Transport Sector

Several categorizations exist in pertinent literature regarding the subjects or kind of decisions that are usually studied in Transport Sector Decision-Making[73][74][75]. Nonetheless, for the purposes of the CONSENSUS project, probably the most useful classification regarding the subjects or kind of decisions that are usually studied in Transportation Policy Decision-Making is according to the nature of the subject:

- **Alternative design solutions** of an infrastructure transportation project: they can include alternative alignments/paths for roads or rail projects, alternative locations for ports, airport terminals and garages or their concepts or forms, different designs for public transport lines in urban areas etc.

- **Alternative infrastructure transportation projects**, to give priorities in the construction of different transport infrastructure projects, taking into account the availability of funds.

- **Alternative transport options**, such as alternative freight transportation routes (for multimodal freight transport) etc.

- **Alternative transport policies** or **transport policy measures**, such as transport pricing alternatives, application of transport demand management etc.

Especially for decisions regarding transport policies or transport policy measures, an important element of the decision making process are the available **policy instruments**, i.e. the tools which can be used to overcome the identified problems and achieve the desired objectives. A common classification of the available policy instruments is according to the type of intervention[76][77]:

- **Infrastructure provision** refers to additions or enhancements to the existing transportation infrastructure.

- **Management measures** involve changes in the way existing transportation infrastructure is used. They include a wide range of approaches, including increases and reductions in road capacity, reallocations of that capacity, and changes in the operation of public transportation.

- **Information provision** refers to improvements in the information available to transportation users and operators. Some are traditional fixed information systems; others draw on real time applications of information technology.

- **Pricing measures** refer to changes in the cost of transportation use for both private vehicles and public transportation.
- **Land use measures**: these measures focus on the land use patterns, which generate the demand for transportation and not on the transportation system as such. The overall emphasis is placed in identifying ways for the reduction of travel demand, or in alleviating its impact.
- **Behavioral/ attitudinal measures** aim to change users' understanding of transportation problems and hence induce changes in travel patterns.

Unfortunately the evidence which is available on the performance of many of these policy instruments is generally very incomplete. In some cases this is because the policy instruments are novel, and experience is still limited; in others the information gained, especially by unsuccessful implementation of measures is not made publicly available. Even where experience is available it may not be directly relevant in another context. For all of these reasons it can be difficult to judge how transferable experience with successful policy instruments will be[72].

It should be mentioned also that, typically, MCDM methods are being applied for the evaluation of transport projects (alternative solutions or different infrastructure projects) rather than transport policies or programs[78].

2.4.5 Role of Multi-Criteria Decision Making in Transport Sector

Since many diverse forms of decision problems in transport sector exist, it is obvious that multi-criteria decision making can assist in different ways and produce various kinds of results. According to relevant research literature and case studies[1], application of MCDM in transport sector problems, can result in the following general forms of solutions:

- **Ranking of examined options** is probably the most common form of solution from the application of MCDM in transport sector problems. In such cases, the analysis concludes that, according to the objectives and criteria established, option A is "better" at fulfilling the assumed goal than option B, which is "better" than option C etc.
- **Identification of a single most preferred option**, to be implemented by transport authorities is also a common result of a MCDM application. This form of solution cannot easily be distinguished from the ranking of options, because, in most cases, the option that is ranked first is the most preferred option that will be selected for implementation.
- Another possible form of the solution provided by MCDM is the **classification of options into categories**. The type of categories may vary, depending on the specific

[1] The numerous –and as such excessive to be referenced in this Deliverable- research and case studies reviewed can be found in Deliverable 2 – Chapter 6.3 and Appendix III.

characteristics of the decision problem at hand. Categories usually found in pertinent literature are: "acceptable" or "unacceptable" options, priority categories for implementation, or identification of a short list of options for further appraisal.

- Finally, certain MCDM methods, mostly Multiple-Objective Decision Making (MODM) models result in **optimization solutions** to a decision problem, such as the recommended crew size in a mass transit system or traffic signal timing optimization.

2.4.6 Multi-Criteria Decision Making Methods Used in Transport Sector

Generally, MCDM methods that are applied in transportation problems can be classified into the following two basic categorie[65],[70],[79]:

- methods for solving problems with a discrete set of options, i.e. a **finite number of alternative solutions** (options) that are known at the beginning, and

- methods for solving problems which require selection from continuous sets of options, that encompass an **infinite or very large number of alternative solutions** that are not explicitly known in the beginning

Methods that encompass a finite number of alternative solutions (options) are appropriate for "ill-structured" problems, i.e. problems with very complex objectives, often vaguely formulated, with many uncertainties, while the nature of the observed problem gradually changes during the process of problem solving. These methods, usually called **Multiple-Attribute Decision Making** (MADM) or **Multicriteria Analysis** (MCA) models focus on solving the problem by finding the best alternative or a set of good alternatives in relation to defined attributes / criteria and their weights[65]. Examples of MADM methods include: Simple Additive Weighting (SAW), Multi Attribute Utility/Value Theory (MAUT/MAVT), ELimination and (Et) Choice Translating REality (ELECTRE), Preference Ranking Organization METHod for Enrichment Evaluations (PROMETHEE), Analytic Hierarchy Process (AHP) etc.

Methods that encompass an infinite or at least a very large number of alternative solutions are appropriate for "well-structured" problems. Well-structured problems are those in which the present state and the desired future state (objectives) are known as the way to achieve the desired state. The model encompasses an infinite or very large number of alternative solutions that are not explicitly known in the beginning, constraints are analyzed, and the best solution is reached by solving the mathematical model[65]. These methods, usually called **Multiple-Objective Decision Making** (MODM) models, in general consist of two phases, the generation of a set of efficient solutions and the exploration of this set in order to find a 'compromise solution' by means of interactive procedures[70]. Examples of Multiple-Objective Decision Making methods include: Global Criterion method, Utility Function method, Goal Programming (GP), STEp Method (STEM), Genetic Algorithms etc.

Transport sector problems usually are characterized by a finite number of alternative solutions (designs of a project, projects, policies, policy measures etc.), a complex set of objectives, criteria and indicators and many uncertainties. As such, transport sector problems are "ill-structured" problems and therefore MADM/MCA methods are usually appropriate. Examination of relevant research and case studies indicates that probably the most commonly used methods are Analytic Hierarchy Process - AHP (especially for criteria weighting), Multi Attribute Utility/Value Theory - MAUT/MAVT, Outranking methods (ELECTRE, PROMETHEE, REGIME etc.) and Simple Additive Weighting (SAW). In many occasions, a combination of methods is used (e.g. AHP for criteria weighting and MAUT or REGIME for evaluation of total performance), or certain parameters of methods are modified (e.g. introduction of fuzzy criteria, modified concordance analysis etc.), in order to better adapt the methodology to the specific decision problem. Finally, other methodologies, such as CBA scoring or GIS tools may be incorporated in the decision procedure or the presentation of the results.

The use of MODM methods in transport sector problems is less common, applied mainly in optimization problems. Relevant research examination indicated that usually some form of genetic algorithm or specialized heuristic procedures are used for that purpose.

2.4.7 Multi-criteria Decision Making (*Evaluation*) Parameters Commonly Used in Transport Sector

Although the applied MCDM methods can have significant differences, in all cases a very important part of the MCDM procedure is the definition of the hierarchy of goal, objectives, criteria and indicators of the decision problem. The goal of the decision problem is a very general statement of the desired improvement. Objectives are also statements of something that one desires to achieve, but are more specific than goals and each objective reveals an essential reason for interest in the decision situation. Criteria, or attributes, provide a measure of the degree to which an objective is met by various options/alternatives of the decision problem and indicators (quantitative or qualitative) further measure, in more specific ways, the performance of options.

Some analysts, instead of using the terms goal, objectives, criteria and indicators, prefer the structuring of the decision problem in several levels of objectives, thus the second level objectives correspond to criteria and the third level to indicators. Furthermore, it is possible that a level of the hierarchy could be missing from the analysis, e.g. indicators could be directly used for measuring the performance of options against the objectives, without explicit definition of criteria. Nevertheless, a complete typical structuring of a decision problem consists of the above evaluation parameters.

2.4.7.1 Objectives

A set of objectives in a decision problem should possess the following properties: essential, controllable, complete, measurable, operational, decomposable, non-redundant, concise and understandable[80]. Objectives specify the directions for improvement, but not the means of achieving them. In setting objectives, it is therefore important to avoid including indications of preferred solutions (e.g. "improving the environment through better public transport"), since this may cause other and possibly better policy instruments to be overlooked[72]. Setting clear and concise objectives in a decision problem has the following benefits[72]:

- helps to identify problems in the decision process
- provides guidance on the types of solutions

- can act as constraints, in clarifying what should be avoided
- provides the basis for appraisal of alternative solutions, and
- enables progress in implementation to be monitored

Since impacts from transport infrastructure projects or transport policies are wide and varied, the spectrum of common objectives in transport sector decision problems is also very broad. Objectives commonly found in transport sector decision problems[81][82][83][72][74] are the following:

- **Economic efficiency**: Economic efficiency involves minimizing implementation, operation and maintenance costs of the project or policy involved, and maximizing the financial benefits which users can gain from the transport system.
- **Transport system efficiency**: This objective refers to maximization of the efficiency of the transport system in terms of (according to each specific decision problem): reduction in travel time, reliability of travel time, minimization of congestion, integration to existing transport system, ability to effectively connect origins and destinations etc.
- **Protection of the environment**: This objective involves reducing a number of adverse impacts of the transport and land use system, such as air pollution (NO_X, CO_2, SO_2, local pollutants such as particulates etc.), their impacts on health, noise and vibration, visual intrusion, fragmentation and severance of settlements and biodiversity, urban sprawl, and loss of cultural heritage and natural habitats etc.
- **Safety**: This objective straightforwardly involves reducing the numbers of accidents for all modes, and reducing the severity of those which occur. However, since some locations, age groups and modes have higher accident rates than others, the safety objective also has equity implications.
- **Equity and social inclusion**: Under equity the principal concerns are the need for reasonably equal opportunities to travel, costs of travel and environmental and safety impacts of travel. Social inclusion mainly refers to accessibility for those without a car and accessibility for those with impaired mobility.
- **Contribution to economic growth**: Land use and transport policies should support economic growth and regional development. Transport improvements which improve access or enhance the environment can lead to increased economic activity and possibly to sustained economic growth.
- Other, less frequently used objectives are: public acceptance, privacy issues (e.g. feeling of intrusion), specific engineering objectives (staging flexibility, terrain and soil characteristics, volume of earthworks) etc.

It is important that decision-makers determine the objectives which they wish to pursue. However, it is preferable to reach agreement on them with other stakeholders and objective definition is often a key first stage in the participation of stakeholders in decision making.

Especially regarding road pricing related decision making, examination of relevant case studies in pertinent literature[84][85][70][86] reveals that four main -high level- objectives are commonly used, related to:

- economic development / growth
- transport / mobility / safety conditions
- life conditions, environment and energy conservation, and
- social cohesion, satisfaction and acceptance

2.4.7.2 Criteria and Indicators

Objectives are abstract concepts, and it is thus difficult to measure performance against them. Criteria (attributes) and indicators are ways of measuring objectives. For example, under the "protection of the environment" objective, a possible criterion would be "minimize air pollution" and a relevant indicator could be the expected CO_2 emissions.

Possible criteria related to the aforementioned objectives in transport sector decision problems could be the following[81],[82],[83],[72],[74]:

- **Economic efficiency**: Minimize construction/implementation cost, minimize maintenance cost, minimize operation cost, maximize Internal Rate of Return etc.
- **Transport system efficiency**: Minimize travel time, maximize reliability of travel time, minimize congestion, maximize comfort of service, maximize integration to existing transport system, maximize interoperability of networks, maximize ability to effectively connect origins and destinations, maximize transport network capacity, maximize passenger/freight movements, minimize construction period etc.
- **Protection of the environment**: Minimize air pollution, minimize water pollution, minimize visual intrusion, minimize land use fragmentation, minimize impacts on waterlands and natural habitats, minimize fuel consumption, minimize noise and vibration etc.
- **Safety**: minimize fatalities, minimize injuries, minimize number of accidents etc.
- **Equity and social inclusion**: Maximize accessibility for those without a car, maximize accessibility for those with impaired mobility, minimize household displacement, maximize connectivity for deprived geographical areas etc.
- **Contribution to economic growth**: Maximize regional development, maximize positive effects on tourism, maximize ease of connection between residential and employment areas, maximize positive effect on local employment etc.

In order to measure (quantitatively or qualitatively) the performance of options against criteria, indicators are constructed. There are essentially three types of indicators[71],[79]: natural, constructed and proxy. **Natural indicators** are those in general use that have a common interpretation to everyone and the impact levels reflect the effects directly (e.g. value of construction costs as an indicator for criterion "Construction Cost"). **Constructed**

indicators are developed specifically for a given decision context. In general, a constructed indicator involves the description of several distinct levels of impact that directly indicate the degree to which the associated criterion or objective is achieved (e.g. archaeological items within 50 m of the right-of-way as an indicator for criterion "Impact on Archaeological Heritage"). It is essential that the descriptions of those impact levels are unambiguous to all individuals concerned about a given decision. If no natural or constructed attribute is available, it may be necessary to utilize an indirect measure or a **proxy indicator**. When using proxy indicators, the impact levels mainly reflect the causes rather than the effects; (e.g. length of surface track as an indicator for criterion "Noise Impact").

Especially regarding road pricing related decision making, examination of relevant case studies in pertinent literature[84],[85],[70],[86],[77] reveals that several criteria are examined –in each objective category-, such as:

- **Economic development / growth**: Gross revenue generation potential, increase macroeconomic welfare, increase regional welfare, maintain / increase employment etc.
- **Transport / mobility / safety conditions**: Guarantee a minimum quality of transport, improve accessibility conditions, improve safety, improve reliability of services, decrease travel time, reduce traffic congestion etc.
- **Life conditions, environment and energy conservation**: Improve air quality, reduce energy consumption, maintenance of ecosystems' functions, reduce noise annoyance etc.
- **Social cohesion, satisfaction and acceptance**: enhance personal basic mobility, increase regional cohesion, ensure socioeconomic fairness etc.

The above criteria are further decomposed into lower level indicators, of quantitative or qualitative nature, that permit the analysts to measure the performance of each examined alternative road pricing strategy.

2.4.8 Participation of Stakeholders in Multi-Criteria Decision Making in Transport Sector

Participation of stakeholders can be a very important part of the decision making procedure in MCDM, in order to take into consideration the different aspects and opinions regarding the examined options. Participation can occur in different levels, such as information provision, consultation, deciding together, acting together or even supporting independent stakeholder groups. Each level is appropriate for different kinds of decision problems, different stages in the development of a strategy, or for strategies tackling different scales of problem. In relevant research and case studies, participation of stakeholders was found in several forms, ranging from news release, brochures and mail-outs to advisory committees and public workshops. In general, all forms of participation methods are possible in MCDM. However, different forms are more or less appropriate for different decision problems or different phases of the decision process.

2.4.9 Multi-Criteria Decision Making in Transport Policy Scenario of Consensus

Summarizing the presented context of Multi-Criteria Decision Making in the transport sector, the following conclusions can be drawn and serve as guidelines in developing the specific context of the Consensus transport policy scenario.

- Multi-Criteria Decision Making is very useful for plan-led and consensus-led approaches in decision making, or for mixed plan-led and consensus-led decision-making; **to this end such a mixed approach of decision-making it is assumed to be applied in the Consensus transport policy scenario**. More analytically, according to the vision-led approach it is assumed that the policy/ decision-makers of the Consensus transport policy scenario will have a clear view of want they want to achieve as well as of the general policy instruments needed to achieve it; that are road pricing instruments. Simultaneously, according to the concensus-led approach stakeholders' affected and/or involved in road pricing implementation will be engaged in the decision-making process focusing on the choice of options but on objectives and problems as well.

    ✓ Concerning stakeholders' identification and participation; **groups typically included in transport sector decision making and their participation methods were identified and used in the Consensus framework**.

- Based on the wide range of literature, research and case studies reviewed the evidence available on the Multi-Criteria Decision Making among policy instruments, such as road pricing, is generally very limited and/or incomplete. Typically, MCDM methods are being applied for the evaluation of transport projects (alternative solutions or different infrastructure projects) rather than transport policies or programs. This probably happens because most policy instruments, especially pricing instruments, are novel, and experience is still limited; in other cases the information gained, especially by unsuccessful implementation of measures is not made publicly available. Even where experience is available it may not be directly relevant in another context. For all of these reasons it can be difficult to transfer much experience into the Consensus concerning successful road pricing policy instruments. **To this end all possible road pricing schemes were initially considered and then through stakeholders' consultation specific road pricing schemes of interest were chosen to be examined in the Consensus framework**.

- Despite the diverse levels of decision-making approaches, the different nature/subject of decisions examined and/or the alternative desired results through a MCA application in the transport sector, in all cases the possible objectives arise from a common list and always include effects on the four basic sustainability dimensions: economy, mobility, environment and society. **To this end, these four sustainability dimensions were decided to be used as the evaluation objectives of Consensus transport policy scenario**.

- Objectives though are abstract concepts, and it is thus difficult to measure performance against them. Criteria (attributes) and indicators are ways of measuring objectives. For example, under the "protection of the environment" objective, a possible criterion would be "minimize air pollution" and a relevant indicator could be the expected reduction in specific pollutants emissions. **Based on this logic and the review of the**

numerous case studies and pertinent literature, all possible criteria related to the aforementioned objectives along with the respective indicators were initially considered; then through stakeholders' consultation specific criteria and indicators were chosen to be used in the Consensus transport policy scenario evaluation.

Finally, despite the fact that Multi-Objective Decision-Making methods usage is less common in transport sector problems -and it is applied mainly in very specific and/or narrow area problems i.e. traffic signaling optimization- the Consensus policy scenarios (including transport policy scenario) will be assessed using a multi-objective optimization tool developed specifically for this purpose.

This latter mentioned can be considered as the contribution of Consensus project to the State-of-the-Art; supporting the policy decision-maker to solve policy related problems where the set of alternative policy options encompasses a very large number of alternatives.

Especially for the transport/road pricing policy scenario this will be very useful, since the road pricing alternative options might be discrete in terms of their components but there is one component (price level) that works in a continuous way as such generating a large number of alternative options.

2.5 Visual Analytics

2.5.1 Introduction

Visual Analytics tightly couples data mining and visualization approaches to include human users in the analysis and data understanding loops, helping to make sense of data and find appropriate decisions. (Please see also the State-of-the-art report Deliverable D2.2, section 4).

In the Consensus project we deal mainly with multi-dimensional data sets which correspond to policy alternatives (input and output) and which need to be compared against each other, considering alternative weighting schemes, to arrive at assessments. To represent this kind of data, scatterplot matrices or parallel coordinate plot techniques are suitable methods. First Visual Analytics research carried out in Consensus therefore focused on developing multi-dimensional comparison techniques and testing these with first data sets obtained by partners. Specifically, first research prototypes have been implemented and deployed on the web for internal testing.

In our prototypes we make extensive use of glyph designs and the possibility to have multiple views on the data. Therefore, we here briefly introduce related research in this area to come up with a suitable glyph design. Then, we will describe functional components of our approaches in greater detail.

2.5.2 Glyph-Based Evaluation

For a detailed overview of research on data glyphs, we refer the interested reader to two summary articles[87],[88]. There exists a large amount of glyph designs and only little guidance, which design performs best for certain types of data or tasks. Domain experts in the Consensus project have to mainly perform similarity judgments to compare different scenarios. However, there is only little related work investigating the performance of glyph designs for similarity judgments.

Wilkinson[89] conducted a user study comparing star glyphs, castles, Chernoff faces and blobs. Participants had to sort 8 glyphs of each type---varied by a variety of factors---according to increasing dissimilarity. Their findings indicate that judgments on Chernoff faces were closer to the actual factor distances, followed by star glyphs, castles and blobs.

A similar sorting-based task was used by Borg and Staufenbiel[90] in their comparison of snowflakes (similar to star glyphs), suns, and factorial suns. Participants had to sort 3 times 44 shuffled cards showing data points of one type of glyph into four categories according to their similarity. Factorial suns---that make use of some preprocessing of the data---were most easily discriminated and star glyph performed the worst in this respect. Lee et al.[91] showed participants several datasets represented by one of: small-multiples Chernoff faces, star glyphs, and two plots produced with multi-dimensional scaling. For each dataset participants were given eight questions to answer, some of which included similarity judgments based on pairwise comparisons. The authors did not perform an analysis on the basis of individual similarity questions. Instead, they found that participants performed best and were most confident with one of the 2D spatial plots, in particular on global questions where the whole set of data points has to be considered.

Klippel's study[92] investigated Star Glyphs, which are well-known representatives for multi-dimensional data used in the Consensus project. They investigated the influence of shape on glyph perception based on similarity judgments. They varied shape by reordering the dimensions in a star glyph with contour. The authors studied how shape changes influenced the interpretation of data points in a similarity-based grouping task. They found that differences in shape influenced cognitive processing of the data and those perceptually salient features (such as spikes) strongly influenced how people thought about a data point.

Given the fact that only little advice exists on which glyph design should be preferred when performing similarity comparisons, we want to extend the research in this field by conducting another quantitative user study investigating the performance of star glyph variations for similarity judgments. Section 6.1 later will detail our results. Then, also later in Sections 6.2 and 6.3 we will introduce particular interaction and alignment techniques to foster the comparison of multivatiate data as per the uses cases in Consensus.

2.6 Gamification and Crowdsourcing

2.6.1 Introduction

Within a set of optimal solutions representing optimizations of multiple objectives, the decision maker needs to identify the priorities that will lead to the selection of a single policy scenario. For setting those priorities the weight of public opinion plays an important role. In order to include this information in the decision making process, Consensus aims to approach citizens through a web platform that will allow the collection of their opinion regarding the objectives in question; thus crowdsourcing the task of identifying the public opinion preferences. The challenging part of this endeavor is the incentivation of the citizens' participation and for that reason the project employs gamification techniques:

competition, challenges, visualizations, rewards and links to user reality. In what follows we provide the state of the art methods and technologies used in these techniques, classified in three major categories: gamification, crowdsourcing and serious games. These methods, even though not all used by Consensus researchers, comprise the baseline knowledge upon which the ConsensusGame implementation was inspired.

2.6.2 Gamification

Goldberg in 1989[93] proposed Pareto-based fitness which bases directly on the concept of pareto dominance. In Goldberg's method the individual s are ranked iteratively: first all non-dominated solutions are assigned rank 1 and then the next non-dominated solutions are assigned rank 2 and so forth.

Fonseca and Flemming[94] stated that an individual's rank corresponds to the number of solutions in the population by which it is dominated.

Srinivas and Deb[95] created Non-dominating Sorting Genetic Algorithm (NSGA) based on Goldberg's suggestions, analogous to Goldberg the fitness assignment is carried out in several steps in each step the non-dominated solutions constituting a non-dominating front are assigned the same dummy fitness value these solutions are shared with their dummy fitness values and ignored in the further classification process. The dummy fitness is set to a value less than the smallest shared fitness value in the current non dominated front and the next front is extracted. This procedure is repeated until all individuals are classified. In the original study this fitness assignment method was combined with a stochastic remainder selection. The complexity of the algorithm is $O(mN^3)$ where m is the number of objectives and N is the population size.

Deb, Pratap, Agarwal and Meyarivan in 2002[96] created NSGA-II in which for each solution two entities are calculated: domination count n_p, the number of solutions which dominate the solution p, and S_p, a set of solutions that the solution p dominates. This requires $O(mN^2)$ comparisons. In the algorithm all solutions p in the beginning are marked with $n_p = 0$. For each solution with $n_p = 0$ each member (q) of its set is visited and its domination count is reduced by one. In doing so, if for any member the domination count becomes zero, we put it in a separate list Q. These members form the second non-dominated front. The above procedure is continued with each member of Q and the third front is identified. This process continues until all fronts are identified.

Zitzler and Thiele[97] created an elitist multi-criterion EA with the concept of non-domination in their strength Pareto EA (SPEA). In their algorithm an external population was maintained at every generation storing all non-dominated solutions discovered so far beginning from the initial population. At each generation the external and current population are combined, all non-dominated solutions in the combined population are assigned a fitness based on the number of solutions they dominate and dominated solutions are assigned fitness worse than the worst fitness of any non-dominated solution. This assignment of fitness makes sure that the search is directed towards the non-dominated solutions. To ensure diversity among non-dominated solutions a deterministic clustering technique is used. The implementation suggested is $O(mN^3)$.

Knowles and Corne ([98],[99],[100]) implemented a simple MOEA using an evolution strategy (ES). In their Pareto-archived ES (PAES) with one parent and one child, the child is compared to the parent. If the child dominates the parent, the child is accepted as the next parent and the iteration continues. If on the other hand the parent dominates the child, the child is discarded and a new child is found. If the child and the parent do not dominate each other, the choice between the child and the parent considers the second objective of keeping diversity among obtained solutions. In order to keep diversity an archive of non-dominated solutions is maintained. The child is compared with the archive to check for dominance. If the child dominates any other member in the list it is accepted as the new parent and the dominated solution is eliminated from the archive, if not then both parent and child are checked for their nearness with the solutions of the archive. If the child resides in a least crowded region in the parameter space among the members of the archive, it is accepted as a parent and a copy of added to the archive. The overall complexity of the algorithm is$O(mN^2)$. Knowles and Corne in their other implementation PESA, based it on the degree of crowding in different regions of the archive. Replacing the selections in the archive file is also based on a crowding measure. PESA uses binary tournament selection and for selective fitness the squeeze factor (the chromosome with the lowest squeeze factor is chosen).

Greenwood, Hu, and D'Ambrosio[101] suggested a solution using no preference information (in the case of Pareto rankings) and aggregation methods like weighted sum. They extended the concept of Pareto dominance by elements of imprecisely specified multi-attribute value theory in order to incorporate preference in the search process. By systematically varying the numerical scalar weights in an aggregate objective function (AOF), each set of weights results in a corresponding Pareto solution.

Generally in the process of maximizing the objectives and acquiring the pareto-optimum solutions we have three distinct categories that are formed by the non-dominated values:

- When we witness 1% of the total population of solutions then most of the solutions are dominated
- When we witness 10% of the total population then there is a complete and tight distribution
- When we witness more than 20% of the total population then the algorithm prematurely converged

Conventional GA wisdom states that strongly elitist strategies result in premature convergence[102].

2.6.2.1 Game Theory Models

2.6.2.1.1 Repeated Games

Repeated games are a series of games that get repeated. In infinitely repeated games the average reward given an infinite sequence of payoffs r_1, r_2,… for player i is: $\lim_{k\to\infty} \sum_{j=1}^{k} \frac{r_j}{k}$

Given an infinite sequence of payoffs r_1, r_2r1,r2,… for player i and discount factor β with 0<β<1 its future discounted reward is $\sum_{j=1}^{\infty} \beta^j r_j$.

There are two types of learning in repeated games: fictitious play and no-regret learning. Fictitious play was originally proposed as a method for computing Nash equilibrium. In that scenario each player maintains explicit belief about the other players. They start by initializing their beliefs about the opponent's strategies and by each turn they play a best response to the assessed strategy of the opponent, later they observe the opponent's actual play and update their beliefs accordingly. Formally the player maintains counts of opponent's actions. For every a ϵ A let w(a) be the number of times the opponent has player action a which can be initialized to non-zero starting values. Assess opponent's strategy using these counts:

$$\sigma(a) = \frac{w(a)}{\sum_{a' \in A} w(a')}$$

(pure strategy) best respond to this assessed strategy.

The regret an agent experiences at time t for not having played s is:

$R^t(s) = \max\left(a^t(s) - a^t, 0\right)$. The agent will try to exhibit no regret from the strategy he follows. At each time step each action is chosen with probability proportional to its regret. That is $\sigma_i^{t+1}(s) = \frac{R^t(s)}{\sum_{s' \in S_i} R^t(s')}$ where $\sigma_i^{t+1}(s)$ is the probability that agent i plays pure strategies at time t + 1. No-regret learning (Regret matching) converges to a correlated equilibrium for finite games.[103][104]

2.6.2.1.2 Stochastic Games

A stochastic game is a generalization of repeated games where agents repeatedly play games from a set of normal-form games and the game played at any iteration depends on the previous game played and on the actions taken by all agents in that game. A stochastic game is a tuple (Q, N, A, P, R), where Q is a finite set of states, N is a finite set of n players, A = (A_1,..., A_n), where Ai is a finite set of actions available to player i, P:QxAxQ→ [0, 1] is the transition probability function. P(q, a,q̂) is the probability of transitioning from state q to state $\hat{q}$ after joint action a, and R = r1,...,rn, where ri: Q x A→R is a real-valued payoff function for player i[105][104][103].

2.6.2.1.3 Bayesian Games

Bayesian game is a set of games that differ only in their payoffs, a common prior defined over them, and a partition structure over the games for each agent. A Bayesian game is a tuple (N,G,P,I) where N is a set of games, G is a set of games with N agents each such that if g, g' ∈ G then for each agent i ∈ N the strategy space in g is identical to the strategy space in g'. P∈Π(G) is a common prior over games where Π(G) is the set of all probability distributions over G, and I=(I1,...,IN) is a set of partitions of G one for each agent.

Another definition for Bayesian games states a tuple (N,A,Θ,p,u) where N is a set of agents, A=(A1,...,An) where Ai is a set of actions available to player i , Θ=(Θ1,...,Θn) where Θi is the type space of player i , p:Θ→[0,1] is the common prior over types and u=(u_1,..,u_n) where u_i:AxΘ→Ris the utility function for player i .

The expected utility has three standard notions of expected utility: ex-ante where the agent knows nothing about anyone's actual type, interim where the agent knows her own type but not the types of the other agents and ex-post where the agent knows all agent types.

It is assumed that a player who has only partial knowledge about the state of nature has some beliefs, a prior distribution, about the parameters which he does not know or he is uncertain about. In a multiplayer game the decisions of others players are relevant, so are their beliefs, since they affect their decisions. Thus a player must have beliefs about other player's beliefs in order to form a strategy.

In Bayesian games we have the Bayesian (Nash) Equilibrium according to which players choose strategies to maximize their payoffs in response to others accounting for strategic uncertainty about how others will play and payoff uncertainty about the value to their actions. [106][103][104]

2.6.2.2 Gamification Elements

2.6.2.2.1 Game with a Purpose (GWAP)

Games With A Purpose (GWAP)[107] , propose that using computer games can gather human players and solve open problems as a side effect of playing. GWAP approach is widely used for image tagging [108], [109] collecting common-sense facts[110], music annotation[111],economic games design[112],transportation solutions[113]. Most GWAP implementations valuate results according to three game-structure templates : output-agreement games, inversion-problem games and input-agreement games.

In Output-agreement games[110] a three-step procedure is followed[114]:

***Initial setup**.* The game chooses two players randomly among all players.

***Rules**.* Players are provided with the same input and indulged to produce the same output as their partners. Players cannot see another's output or communicate with each other.

***Winning condition**.* Both players get rewarded for producing, at some point, the same output. Due to the fact both players cannot contact each other they result in the same output related to the only thing they have in common, the input. The output is verified because the same result occurred from two independent sources.

In Inversion-problem games[110][109] a three-step procedure is followed [114]:

Initial setup .The game chooses two players randomly among all players.

Rules. In each round one player is the "describer" and the other is the "guesser". The describer is given the input and has to produce outputs in order for the guesser to find the original input.

Winning condition. The guesser produces the original input given to the describer.

In input-agreement games[111] a three-step procedure is followed [114]:

Initial setup. Two The game chooses two players randomly among all players.

Rules. In each round both players are given the same or different inputs (known by the game but not the players). Players are prompted to produce outputs describing their input.

Winning condition. Players decide whether the input is the same for both players given the outputs the other player provides.

Agreement in GWAP games can be used to verify results only in a global scale. In the task of finding public preference on a policy implementation we will use output agreement to verify all users result in the same general perspective of what should be implemented. To make this clearer, provided we collect a specific amount of user implementations in a specific scenario we can check if users' preference creates patterns that indicate public preference in a specific section and we will check if new users' implementations agree with this selection. If there is indeed agreement that means users agree with public preference. In the scenario of "output agreement" among the choices of the same player in each game session a solid preference will be verified from the last policy implementation made.

2.6.2.2.2 Reward Model

There are four things players enjoy while playing games. Achievement within the game context, exploration of the game, socializing with others and imposition upon others. Therefore creating four basic player categories as Bartle suggested in 1996 achievers, killers, socializers and explorers[115].

All forms of rewards apply to those basic categories of players. There are eight forms of rewards[116]:

1. **Score systems** (use numbers to mark player performance).Scores which generally serve as tools for self-assessment and comparison sometimes affect game play indirectly.
2. **Experience point reward systems** (Avatars earn experience points during game play, and "level up" when specified goals are achieved) These systems differ from score systems in at least three ways, Rather than single game plays or specific players they are bound to specific avatars, they reflect time and effort rather than player skill which results to rarely being used for purposes of player ranking, they directly affect game play by making certain tasks easier to accomplish, as well as by expanding the number of ways that a game can be played.
3. **Item granting system rewards** (that consist of virtual items that can be used by players or much more commonly avatars) Item granting mechanisms encourage players to explore game worlds.
4. **Resources** (valuables that can be collected and used in a manner that affects game play) Resources differ from items in at least one important aspect, resources are mostly for practical game use or sharing, whereas items have collecting and social comparison value. Experience points in leveling system mark the growth of avatars and create a feeling of progress, while resources create feelings mainly about timely support.
5. **Achievement systems** (consist of titles that are bound to avatars or player accounts; users collect them by fulfilling clearly stated conditions). Achievement systems make players complete specific tasks, play in challenging ways, or explore game worlds. Achievements are the type of reward systems classified as glory." Collectable titles

serve as meta-goals, and thus provide "multiple level goals" for various challenges" [117],[118].

6. **Feedback messages** (mostly used to provide instant rewards instant positive feedback that players receive in response to successful actions). Feedback messages create positive emotions, pictures, sound effects, and video clips are also commonly used as feedback mechanisms. They are neither collectable nor available for player comparisons, and do not directly affect game play.
7. **Plot animations and pictures** (used as rewards following important events such as the defeat of a major enemy, clearing a new level, or ending a game) They motivate players to advance game stories. They create fun in at least two ways they are visually attractive and serve as milestones marking player achievement.
8. **Unlocking mechanisms** (they give players access to game content (e.g., new levels, access to special virtual environments, and mini-games) once certain requirements are met). This kind of reward is best classified as access[119] .As Malone suggests that one of the most important features of intrinsically motivating environments is providing incomplete information about a subject. These mechanisms don't reveal all possibilities and choices at the beginning of games, instead they reward players as games progress by gradually exposing hidden parts of game worlds.

2.6.3 Croudsourcing

In Crowdsourcing needed services, ideas, or content are obtained by soliciting contributions from a large group of people, and especially from an online community, rather than from traditional employees or suppliers. Crowdsourcing combines the efforts of numerous self-identified volunteers or part-time workers, where each contributor of their own initiative adds a small portion to the greater result.

In implicit crowdsourcing , crowdsourcing is less obvious because users do not necessarily know they are contributing, yet can still be very effective in completing certain tasks. Users are not actively participating in solving a problem or providing information, but instead do another task entirely where a third party gains information for another topic based on the user's actions. In our case users play the game with other users and try to excel in levels and we on the back end collect information about user preference on specific policies according to their selections and comments during the game.

Other crowdsourcing applications include Verbosity a game that collects common sense facts [110], Tagatune a game that annotates music and sounds[111],Peekaboom a game that locates objects in images [109], ESP game, a game that labels images [108] and reCAPTCHA which asks people to solve CAPTCHAs to prove they are human, and then provides CAPTCHAs from old books that cannot be deciphered by computers, to digitize them for the web [120].

2.6.4 Serious Games

Serious games are simulations of real-world events or processes designed for the purpose of solving a problem. Although serious games can be entertaining, their main purpose is to train or educate users. In consensus one of the main goals is to educate citizens about policy making relative to Biofuels and transportation and also inform them of the tradeoffs and consequences theirs decisions suggest.

Other serious games applications relative to Biofuel and transportation policies include CO2GO[121] a mobile application that claims to calculate carbon footprint in real-time while on the move, IBM City One Game a city-building simulation game introducing the effects of various policies[122], I-Gear uses gamification as a way to optimize mobility patterns within a heavily congested European City[123], SimCityEDU: Pollution Challenge is a game-based learning and assessment tool for middle school students covering the Common Core and Next Generation Science Standards[124] and intelenBIG claims to enable an organization to reduce its overall energy consumption through behavioral change at the same time, it is able to raise environmental awareness among its premises' occupants in an efficient and entertaining way[125].

3 The GLOBIOM Optimization Model

GLOBIOM is a global recursive dynamic partial equilibrium bottom-up model integrating the agricultural, bioenergy and forestry sectors. In this section we will focus solely on the optimization approach in the model. For a more complete model description we refer to "D.3.2.1 Models and Simulators Report", "D.2.1.1 User requirements" and "D.2.4.1 System Architecture".

GLOBIOM is an economic linear optimization model wherein the global forestry and agriculture market equilibrium is determined by choosing economic activities to maximize social welfare (consumer and producer surplus) subject to resource, technological, demand and policy constraints following McCarl and Spreen [126]. GLOBIOM is a linear mathematical programming model. This type of model is derived from aggregation of more simplified linear programming models of production used in microeconomics [127] which have been long used in economics for many sectoral problems, in particular in agricultural economics. Development of recent computation capacities allowed application of this framework to large scale problems with a high level of details.

The optimization problem in GLOBIOM is a linear programming (LP) problem which can be described in the following simplified form:

$$Max \sum_j c_j X_j$$

$$s.t. \sum_j a_{ij} X_j \leq b_i \; for \; all \; i$$

$$X_j \geq 0$$

In the LP problem, decision variables x_j (i.e. production activities) are chosen so that a linear objective function value c_jX_j (in GLOBIOM the consumer and producer surplus) of the decision variables is optimized given a simultaneous set of linear constraints involving the decision variables. The a_{ij}, b_i, and c_j are the exogenous parameters of the LP model where a_{ij} are the resource requirements, b_i the resource endowments and c_j the benefit coefficients. Different resources are represented by *i* and different production activities by *j* [128].

As GLOBIOM is a linear model, non-linear relationships (i.e. non-linear downward sloped demand function) need to be linearized. In this type of approach, the supply side can be very detailed, in particular benefiting from the possibility of linearizing the non-linear elements of the objective function, the model can be solved as a LP model, allowing a large quantity of data to be used for production characteristics. The GLOBIOM model for instance can optimize the production for each sector on a large number of geographic units. Additionally, many technologies and transformation pathways can be defined for the different sectors. This detailed representation on the production side however induces a trade-off on the demand side. Because of the linear optimization structure, demand is represented through separated demand functions, without a representation of total households budget and the associated substitution effects McCarl and Spreen [126].

GLOBIOM is a price endogenous model compared to the standard LP model, where input and output prices or quantities are assumed fixed and exogenous. In price endogenous models as GLOBIOM, the level of output influences equilibrium prices. The objective function maximizes the integral of the area underneath the demand curve minus the integral underneath the supply curve, subject to different constraints such as a supply-demand balance. The resultant objective function value is commonly called consumer plus producer surplus. Producer surplus is determined by the difference between equilibrium prices and the cost of the different production factors (labor, land, capital) and purchased inputs. On the consumer side, surplus is determined by the level of consumption on each market: the lower the equilibrium price is, the higher the consumption level can be as well as the consumer surplus. The objective function in GLOBIOM includes the following cost term: production cost for the crop- and livestock sector, costs for irrigation water, land use change costs, processing costs, trade costs and a potential tax on greenhouse gas emissions.

GLOBIOM covers the whole world aggregated to 57 market regions. It is based on the spatial equilibrium approach developed by Takayama and Judge [129] which enables optimization across different regions. Production and consumption usually occurs in spatially separated regions, each having supply and demand relations. In a solution, if the regional prices differ by more than the interregional cost of transporting goods, then trade will occur and the price difference will be driven down to the transport cost[128].

Objective function

$$\mathrm{Max}\ WELF_t = \begin{aligned}[t] & + \sum_{r,y}\left[\int \varphi^{demd}_{r,t,y}\left(D_{r,t,y}\right)\mathrm{d}(\cdot)\right] - \sum_{r}\left[\int \varphi^{splw}_{r,t}\left(W_{r,t}\right)\mathrm{d}(\cdot)\right] \\ & - \sum_{r,l,\tilde{l}}\left[\int \varphi^{lucc}_{r,l,\tilde{l},t}\left(\sum_{c,o,p,q} Q_{r,t,c,o,p,q,l,\tilde{l}}\right)\mathrm{d}(\cdot)\right] \\ & - \sum_{r,c,o,p,q,l,s,m}\left(\tau^{land}_{c,o,p,q,l,s,m} \cdot A_{r,t,c,o,p,q,l,s,m}\right) \\ & - \sum_{r}\left(\tau^{live}_{r} \cdot B_{r,t}\right) - \sum_{r,m}\left(\tau^{proc}_{r,m} \cdot P_{r,t,m}\right) \\ & - \sum_{r,\tilde{r},y}\left[\int \varphi^{trad}_{r,\tilde{r},t,y}\left(T_{r,\tilde{r},t,y}\right)\mathrm{d}(\cdot)\right] \\ & - \sum_{r,e}\left(\tau^{emit}_{t,e} \cdot E_{r,t,e}\right) \end{aligned} \qquad (1)$$

The supply – demand balance ensures that for each region, product and time period the endogenous demand is met by supply of the different crop-, livestock, bioenergy and forest product plus imports from other regions minus exports to other regions.

Supply - demand balance

$$D_{r,t,y} \leq \sum_{c,o,p,q,l,s,m}\left(\alpha^{land}_{t,c,o,p,q,l,s,m,y} \cdot A_{r,t,c,o,p,q,l,s,m}\right) + \alpha^{live}_{r,t,y} \cdot B_{r,t} + \sum_{m}\left(\alpha^{proc}_{r,m,y} \cdot P_{r,t,m}\right) + \sum_{\tilde{r}} T_{\tilde{r},r,t,y} - \sum_{\tilde{r}} T_{r,\tilde{r},t,y}$$

(2)

Equation 3 limits available land for the production activities in the different sectors (crop-, livestock- and forest sector) to total land available in that land cover category i.e. the area of crops planted cannot exceed the area of cropland available. In the land use change equation (4), land available in each land cover class is defined as the initial land endowments at the beginning of a period, plus land converted to that class minus land being converted to another class. After each period, initial land endowments in each land cover class get updated for the next period. In equation 5, maximum land conversion is limited to the available land suitable for conversion i.e. inside Europe conversion of forests and grassland is restricted.

Land use balance

$$\sum_{s,m} A_{r,t,c,o,p,q,l,s,m} \leq L_{r,t,c,o,p,q,l} \tag{3}$$

$$L_{r,t,c,o,p,q,l} \leq L^{init}_{r,t,c,o,p,q,l} + \sum_{\tilde{l}} Q_{r,t,c,o,p,q,\tilde{l},l} - \sum_{\tilde{l}} Q_{r,t,c,o,p,q,l,\tilde{l}} \tag{4}$$

$$Q_{r,t,c,o,p,q,l,\tilde{l}} \leq L^{suit}_{r,t,c,o,p,q,l,\tilde{l}} \tag{5}$$

Variables

D demand quantity [tonnes, m3, kcal]

W irrigation water consumption [m3]

Q land use/cover change [ha]

A land in different activities [ha]

B livestock production [kcal]

P processed quantity of primary input [tonnes, m3]

T inter-regionally traded quantity [tonnes, m3, kcal]

E greenhouse gas emissions [t CO_2eq]

L available land [ha]

Functions

φ^{demd} demand function (constant elasticity function)

φ^{splw} water supply function (constant elasticity function)

φ^{lucc} land use/cover change cost function (linear function)

φ^{trad} trade cost function (constant elasticity function)

Parameters

τ^{land} land management cost except for water [$ / ha]

τ^{live} livestock production cost [$ / kcal]

τ^{proc} processing cost [$ / unit (t or m3) of primary input]

τ^{emit} potential tax on greenhouse gas emissions [$ / t CO_2eq]

d^{targ} exogenously given target demand (e.g. biofuel targets) [EJ, m3, kcal,…]

α^{land} crop and tree yields [tonnes / ha, or m3 / ha]

α^{live} livestock technical coefficients (1 for livestock calories, negative number for feed requirements [t/kcal])

α^{proc} conversion coefficients (-1 for primary products, positive number for final products [e.g. GJ/m3])

L^{init} initial endowment of land of given land use / cover class [ha]

L^{suit} total area of land suitable for particular land uses / covers [ha]

ω irrigation water requirements [m3/ha]

ε^{land}, ε^{live}, ε^{proc}, ε^{lucc} emission coefficients [t CO_2eq/unit of activity]

Indexes

r economic region (57 aggregated regions and individual countries)

t time period (10 years steps)

c country (203)

o altitude class (0 – 300, 300 – 600, 600 – 1100, 1100 – 2500, > 2500, in meter above see level)

p slope class (0 – 3, 3 – 6, 6 – 10, 10 – 15, 15 – 30, 30 – 50, > 50, in degree)

q soil class (sandy, loamy, clay, stony, peat)

l land cover/use type (cropland, grassland, managed forest, fast growing tree plantations, pristine forest, other natural vegetation)

s species (18 crops, managed forests, fast growing tree plantations)

m technologies: land use management (low input, high input, irrigated, subsistence, “current”), primary forest products transformation (sawnwood and woodpulp production), bioenergy conversion (first generation ethanol and biodiesel, energy production from forest biomass – fermentation, gasification, and CHP) *y* outputs (primary: 18 crops, sawlogs, pulplogs, other industrial logs, fuel wood, plantations biomass, processed products: forest products (sawnwood and woodpulp), first generation biofuels (ethanol and biodiesel), second generation biofuels (ethanol and methanol), other bioenergy (power, heat and gas)

e greenhouse gas accounts: CO_2 from land use change, CH4 from enteric fermentation, rice production, and manure management, and N_2O from synthetic fertilizers and from manure management, CO_2 savings/emissions from biofuels substituting fossil fuels

To solve the optimization problem described above, GLOBIOM uses the GAMS/Cplex solver. This solver allows combining the high level modeling capabilities of GAMS (General Algebraic Modeling System) software with the power of Cplex optimizers. Cplex optimizers are designed to solve large, difficult problems quickly and with minimal user intervention applying the simplex method. Cplex provides solution algorithms for linear, quadratically constrained and mixed integer programming problems.

4 Multi-Objective Optimization and Visualization Tool (MOOViz)

4.1 Introduction to the MOOViz Tool

Decision makers are often required to account for multiple conflicting objectives when selecting a policy for a problem, overall resulting in a potentially large number of candidate policies to consider. The MOOViz tool is aimed at assisting decision makers in the process of selecting a preferred policy amongst a set of candidate policies.

Within a given dataset, an ideal policy is one that achieves better objective results than all other policies. The problem is that usually no such policy exists due to tradeoffs among different criteria. Often, when one objective is improved, other objectives decrease. The task of the decision maker is to find a policy that makes a good compromise of the objective values. Finding a good policy is particularly difficult when the number of options is large and many objectives must be simultaneously considered.

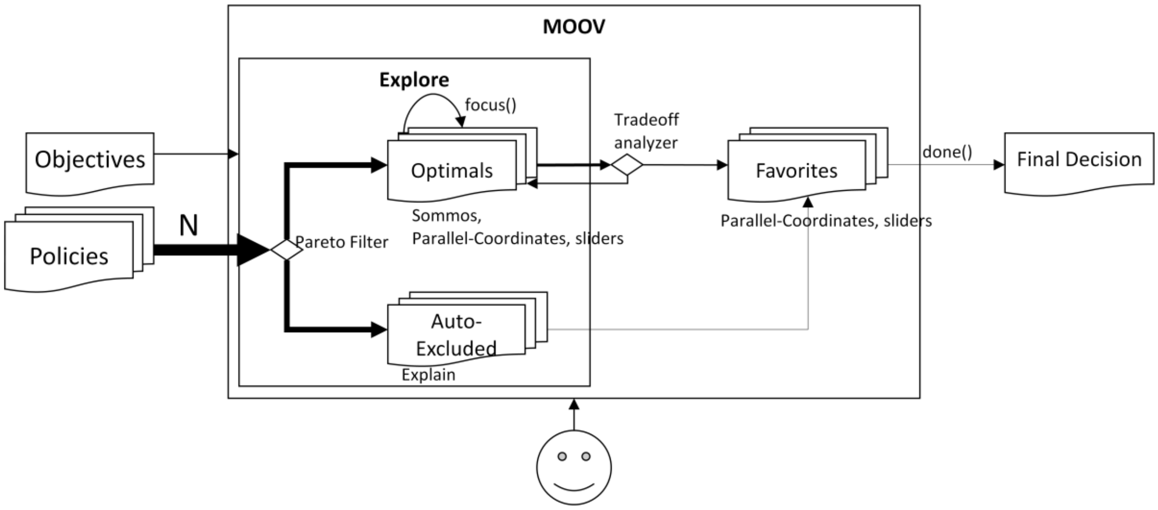


Figure 3: High-level view of MOOViz workflow

The MOOViz tool uses analytics, rich visualizations, and interactions to guide the decision making process until a decision is made. Figure 3 shows a high-level view of the MOOViz workflow. MOOViz accepts two inputs: a set of *objectives* to optimize (maximize or minimize) and a set of alternate *policies*. Each policy represents a possible action and carries numeric measures for each objective. The output is the best policy according to the user preferences. For example Table 4, presents a problem of selecting one of four candidate policies considering three objectives.

Table 4: MOOViz inputs – a domain-definition containing three objectives and a corresponding scenario containing four policies

O1	Maximize
O2	Maximize
O3	Maximize

	O1	**O2**	**O3**
A	100	100	100
B	80	90	70
C	110	90	100
D	100	140	70

One analytics that MOOViz uses is *Pareto filtering*. The Pareto filter removes policies that are *dominated* by other better policies in all objectives. For example, considering the policy in Table 4, policy *B* is dominated by policy *A* as it is worse in all objectives. On the other hand, there is no domination between policies *A* and *C* as each policy has its benefits and drawbacks. Applying the Pareto filter on this dataset will result with policies *A, C* and *D*.

The result dataset after applying a Pareto-Filter is called the *Pareto Frontier* or the *Optimal set*. A decision-maker should consider only the policies on the optimal set. Indeed, MOOViz initially presents the optimal policies. MOOViz also provides the ability to look at the *Auto-Excluded* policies and provides explanation why a particular policy was excluded.

For the optimal policies, MOOViz provides two visualizations techniques (Sommos[2] and parallel-coordinates[3]) for exploring and analyzing the data. When the user clicks on a particular policy a popup is showing details for the policy.

Sliders can be used for filtering policies by their objectives values. Finally, the user can *focus* on the filtered policies showing a 'zoomed' view of the filtered policies.

As the user observes the data, she can add policies to the list of *favorites*. The 'favorites' is a narrow subset of finalist policies – making the decision among them easier. The user compares the favorite policies using a parallel-coordinate chart. Again, the user can filter-out policies using sliders and details are provided on demand.

When the user reaches the decision that a particular policy is the right approach, the user marks the policy as *final* and click the *done* button. The chosen policy is returned back to the hosting application.

4.2 Introduction to Multi Objective Optimization Problems

A multi objective optimization problem is defined as an optimization problem in which there are multiple objectives that need to be optimized in simultaneously. In most cases, there is no single solution that optimizes all objectives, because the objective functions are usually

[2] In the next sections it is referred to as Map or Polygon view

[3] In the next sections it is referred to as Lines view

conflicting. In other words, optimizing one objective will worsen others. A solution is called Pareto optimal or non-dominated if all other solutions are worse in at least one objective value. In other words, a solution is Pareto optimal if none of the objective functions can be improved without damaging other objective function(s). Clearly, if a solution is not Pareto optimal, than there exists a solution which is better than it on all objectives. Thus, it is natural to focus in such Pareto solutions when this is computationally feasible. This set of solutions is called the Pareto front of the optimization problem.

Solving[4] multi objective problems is a difficult task. There are several approaches for that. The most intuitive one is to convert the multi objective optimization problem into a single objective optimization problem (for examples, by using a weighted sum of the multi objectives), and applying single objective optimization methods. Other approaches include the no preference method, a priori methods, a posteriori methods, and more.

Multi optimization problems are encountered in many applications in economics, engineering and science. In the context of decision making, each solution refers to a certain policy. As stated, policies that reside on the Pareto front are considered equally good, and the final policy (solution) chosen depends on the user and involves subjective biases.

4.2.1 Mathematical Background

Let X be a set and $f_1(x),,...,f_N(x)$ functions from X to R^1. A multi objective optimization problem is defined as follows:

$$\min\left(f_1(x), f_2(x),..., f_N(x)\right)$$

Subject to $x \in X$

The set X represents the space of feasible solutions. Note that if an objective function needs to be maximized, the representation still holds when replacing $f_n(x)$ with $(-f_n(x))$.

In order to define a Pareto optimal solution, let us first define *dominate* solution.

Let $x_i, x_j \in X$ be two solutions to the multi objective optimization problem.

x_i dominates x_j if the following conditions hold:

- $f_n(x_i) \leq f_n(x_j),\ n = 1,2,...,N$ namely for each objective functions, the value of x_i does not exceed the value of x_j
- $\exists k, 0 < k \leq N$ such that $f_k(x_i) < f_k(x_j)$ namely for at least one objective function for which the value of x_i is smaller than the value of x_j

A solution is Pareto optimal if no other solution dominates it.

---

[4] Solving in this context refers to finding the set of solutions that reside on the Pareto front

4.3 Research Overview

Decision processes that involve Multi-Objective Optimization problems raise many challenges. The first challenge is solving the optimization problem, namely finding the Pareto optimal solutions, or at least filtering the dominated solutions out of a given set of solutions. The second challenge is visualize the Pareto optimal solutions. This challenge can be divided to two different problems: how to visualize the Pareto optimal solutions in 2D when typically the number of objectives is above 3, and how to visualize the Pareto front in a way that would assist the decision maker to better understand the tradeoffs between the various objective functions.

The research conducted in IBM focused on these topics, and in addition on validating the suggested approaches on various problems. [130] is focused on the challenge of visualizing the Pareto front of the Multi-Objective Optimization problem. The suggested solution (implemented in MOOViz tool) is using Self-Organizing Map. This approach was demonstrated on two real world problems, and was found to provide consistent orientation of the 2D mapping and an appropriate visual representation of the Pareto optimal solutions.

A question that emerges from the visualization challenge involves the ability to evaluate the various visualizations. There exist several methods for visual representation of the Pareto front, but not all of them are equally good. In order to compare between them, a framework is required that would be able to provide evaluation of the various options.[131] suggests a suitable method that focuses on the ability of the visualization to facilitate a better understanding of inter-objective trade-offs to assist in the decision making process. The method was used to evaluate two visualization aids: Parallel Coordinates and an adaption of Self Organizing Maps. The visualizations were compared with tabular data presentation. The results show that the first visualization is more effective than tabular visualization.

The offered visualization using Self Organizing maps was further tested on another application[132]: simulation performance which is evaluated according to multiple quality measures, some of them conflicting. The various performance criteria serve as multiple objective functions, and vector optimization is performed. The approach was applied to a specific Artificial Neural Network simulation with several quality measures. The used visualization as implemented in MOOViz tool assisted in the process of understanding the tradeoffs and choosing the optimal configuration for the simulation process.

Another challenge in the domain of multi objective optimization in the context of decision making, is how to efficiently find a Pareto optimal solution, starting from an initial sub-optimal solution given by the decision maker.[133] suggests a mechanism to handle this challenge using two different methods, which are analyzed and tested.

4.4 MOOViz technical Model Specification

4.4.1 Domain Definition for MOOViz Tool

As a generic technology, the MOOViz tool requires the definition of the domain of interest. A domain consists of a set of policy objectives, constraints (optional), and a set of decision variables. Using MOOViz the decision maker aims at evaluating different candidate alternatives to the decision problem. Each policy alternative consists of a specific assignment to the decision variables and its corresponding objectives. Typically, the policy domain definition is set once when setting the tool for handling a new policy domain. A definition of the domain would rarely change during the decision making process. However, it may be that in future interactions with the decision maker, the domain specification would dynamically change to accommodate to the cognitive model of decision maker.

4.4.1.1 Attributes

A '*DomainDefinition*' JSON[5] object specifies a multi-objective decision problem. The '*objectives*' section lists the objectives that have to be simultaneously minimized or maximized. The '*designParams*' section lists the definition of decision variables comprise a policy alternative.

- `key` [mandatory, string] – identifies this domain
- `objectives` – [mandatory, list]. Each objective is specified using the following attributes:
  - `key` – [mandatory, string] technical identification of an objective
  - `fullName` – [optional, string] human readable name of the objective. This name will appear in all UI interactions. If this attribute is not specified the 'name' attribute is used instead
- `description` – [optional, String] human readable description of the objective
- `format` – [optional, String] a number formatting pattern used to stringify numbers. The pattern string is according to http://www.unicode.org/reports/tr35/tr35-numbers.html#Number_Format_Patterns
- `enumVals` – [optional, list of strings] zero based enumeration labels
- `isMin` – [mandatory, Boolean] specifies whether this objective should be minimized (true) or maximized
- `range` – [optional, object] – specifies the lower and upper bounds of the objective values. When the range is not specified in a domain then the concrete scenario automatically computes the range to the minimum and maximum values of this objective in the scenario solutions
  - `low` – [optional, number] specifies the objective scale lower bound. If not specified, the lower bound is compensated by a percentage denoted by the configuration file. (A document specifying an application configuration would be provided separately)
  - `high` – [optional, number] specifies the objective scale high bound. If not specified, the lower bound is compensated by a percentage denoted by the configuration file
- `designParams` [optional, list]. – Similar to 'objectives', but a design parameter has no isMin attribute because it cannot be optimized

Note, that within a domainDefinition the key attribute of the objectives and designParams must be unique.

4.4.1.2 Domain Definition Sample for the Biofuel Use Case

Following, a sample JSON file for describing the objectives data in the MOOViz tool for the biofuel policy scenario is provided. This domain definition is expected to evolve when additional metrics from the GLOBIOM model will be included in the MOOViz tool.

[5] See http://en.wikipedia.org/wiki/JSON

```
{
  "objectives": [
    {
      "key":"biodiv",
      "fullName":"Bio Diversity",
      "description":"Bio Diversity Change (%)",
      "isMin":false,
      "format":"+#.#%;-#.#%"
    }, {
      "key":"co2",
      "fullName":"CO2 Emission",
      "description":"CO2 Emission Change (%)",
      "isMin":true,
      "format":"+#.#%;-#.#%"
    },{
      "key":"costfood",
      "fullName":"Cost of Food",
      "description":"Cost of Food Change (%)",
      "isMin":true,
      "format":"+#.#%;-#.#%"
    },{
      "key":"forestland",
      "fullName":"Forest Land",
      "description":"Forest Land Change (%)",
      "isMin":false,
      "format":"+#.#%;-#.#%"
    }
  ],

  "key":"Land_Use"
}
```


Figure 4: Biofuel Scenario Domain Definition

4.4.1.3 Domain Definition Sample for the Road Pricing Use Case

Following, a sample JSON file for describing the objectives data in the MOOViz tool, for the road pricing policy scenario. This Domain definition is expected to evolve after the integration of the transportation models with the MOOViz tool.

```
{
  "objectives":[
    {
      "key":"cost",
      "fullName":"Gross Investment Cost",
      "isMin":false,
      "enumVals":[
        "Very Expensive",
        "High Cost",
        "Medium Cost",
        "Low cost",
        "Inexpensive"
      ]
    },{
      "key":"revenues",
      "fullName":"Gross Revenue",
      "description":"Gross Revenue (Million Euro)",
      "isMin":false,
      "format":"#.#M €"
    },{
      "key":"traffic",
      "fullName":"Traffic Volume",
      "description":"Traffic Volume Change (%)",
      "isMin":true,
      "format":"+#.#%;-#.#%"
    },{
      "key":"emission",
      "fullName":"Emission Level",
      "description":"Emission Level Change (%)",
      "isMin":true,
      "format":"+#.#%;-#.#%"
    },{
      "key":"convenience",
      "fullName":"User Convenience",
      "isMin":false,
      "enumVals":[
        "Inconvenient",
        "Low",
        "Medium",
        "High",
        "Very high"
      ]
    }
  ],

  "key":"Transportation"
}
```

Figure 5: Transportation Scenario Domain Definition

4.4.2 Policy Alternatives Data

An '*Alternative Policy*' object represents a possible assignment to the multi-objective problem. A typical decision problem has multiple candidate policies that together form a Pareto-Frontier. The policy Alternatives are JSON objects that automatically generated out of the simulation results.

4.4.2.1 Attributes

Each policy alternative is specified using the following attributes:

- id – [Mandatory, number] technical identification of a solution
- name – [optional, string] human readable name
- objectives – [mandatory, map<string, number>] Each entry in this map specifies the value of a particular objective (referenced using its 'key' attribute)
- designParams – [optional, map<string, number>] Each entry in this map specifies the value of a particular Design-Parameter (referenced using its 'key' attribute)
- descriptionHtml – [optional, String] An Html snippet describing this solution. This html can be used in a web client solution tooltip for example
- appData – [optional, map<String, String>] a placeholder to carry domain-specific applicative data
- status – [generated, enum: "FRONT", "EXCLUDED", "INCOMPLETE"] classifies this solution as being on the Pareto frontier, being Pareto-dominated, or having incomplete data
- statusCause – [generated, object] carry error information
    - errorCode – [String. one of: "MISSING_OBJECTIVE_VALUE", "RANGE_MISMATCH", "MISSING_DESIGN_VALUE"]
    - tokens – [array of strings] carry the error information
    - message – a human readable message (English)

4.4.3 Decision Scenario Data

A `Scenario` is a unit of information that couples a DomainDefinition with a set of policy alternatives:

- `key` – [mandatory, string] unique identifier of the scenario
- `embeddedDomainDefinition` – [optional, DomainDefinition] an inline\embedded DomainDefinition
- `domainDefinitionRef` – [optional, String] the key of a referenced domain definition
- `policies` – [mandatory, list of policy alternatives] – the alternatives for solving the optimization problem. The solutions' objective values do not need to be on the

Pareto frontier. The solution 'id' attribute must be unique within all solutions in a scenario

Note: Exactly one of the attributes '*domainDefinition*' or '*domainDefinitionRef*' must be provided. This is for relating the scenario to one Domain Definition.

4.5 Using MOOViz Tool for Bio-Fuel Scenario

4.5.1 Introduction

The MOOViz tool for Bio-Fuel scenario is aimed at assisting policy makers to explore policy alternatives, to better understand the trade-offs between objectives, and coming into educated decision that is taking into account the entire aspects.

Note that a link to the prototype is provided within Deliverable D4.1.1 Optimization and Visual Analytics Prototypes (due to confidentiality considerations).

Following the link will open the application main page:

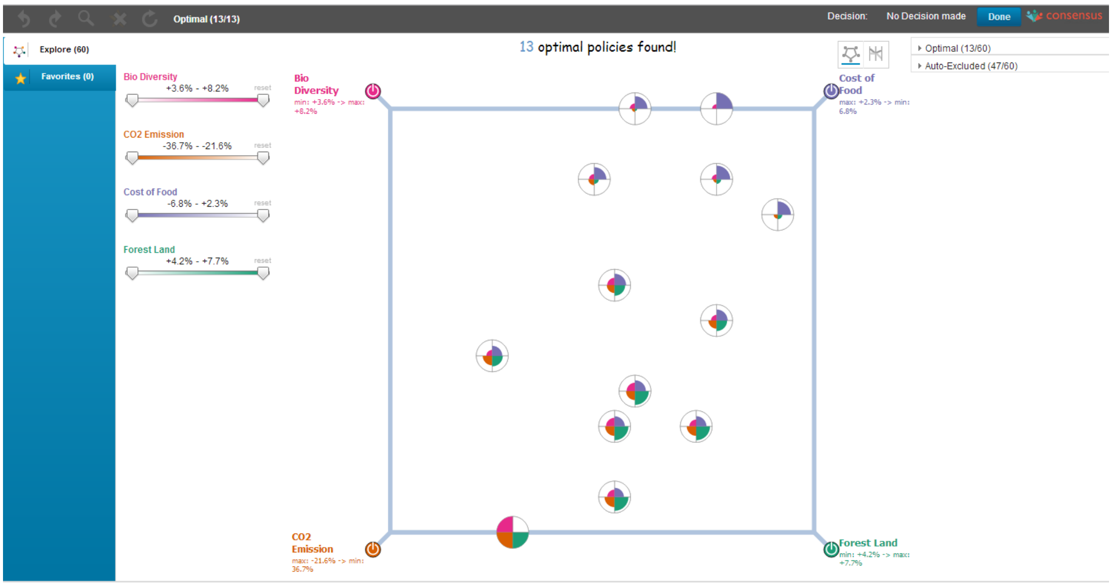


Figure 6: MOOViz tool in bio-fuel scenario – Snapshot of application

The page is composed of several viewports. Figure 7 contains the names of the various components that are described in the following sections.

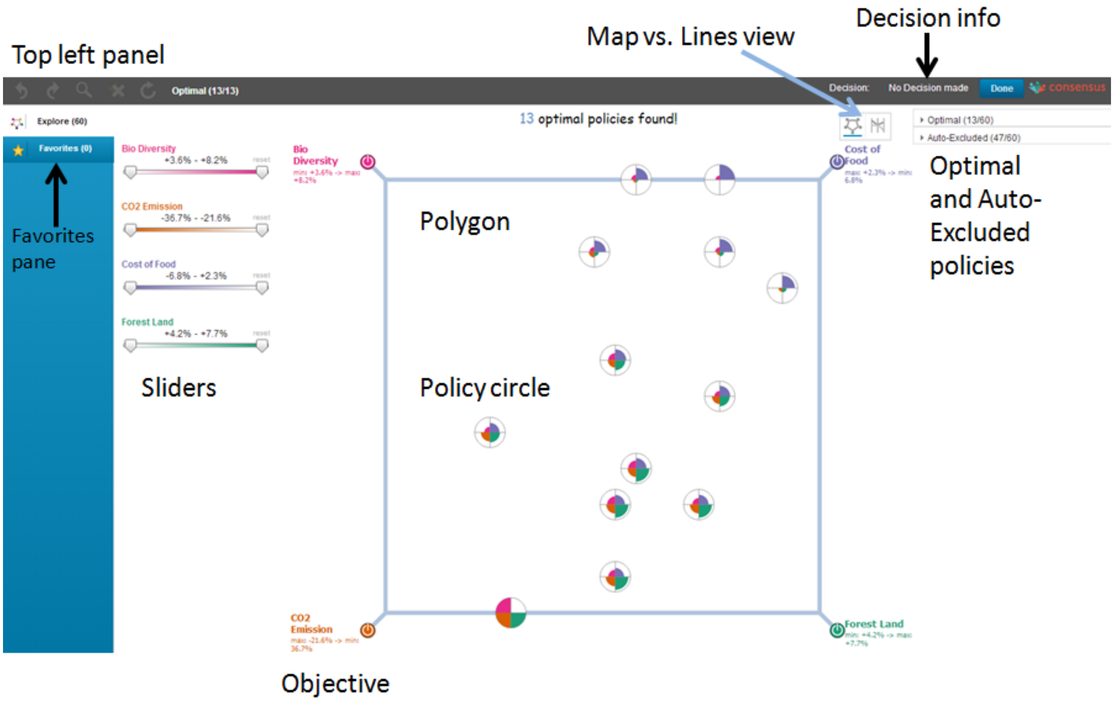


Figure 7: : MOOViz tool in bio-fuel scenario – main components

4.5.2 The Polygon

The main viewport of the application contains a polygon, with a vertex for each objective function. Each objective function has a different color. The corners form a symmetric polygon, which is used to visualize, in two dimensions, the natural high dimensional space in which each dimension represents a different objective function. In each corner, the name of the objective function and its range of values are presented:

Figure 8: MOOViz tool in bio-fuel scenario – objective visualization

The order of the values indicates whether we maximize or minimize the objective. If the first value is “min”, the objective is being maximized. If the first value is “max”, the objective is being minimized:

Figure 9: MOOViz tool in bio-fuel scenario – optimization direction

By clicking the colored circle, the corresponding objective function is disabled:

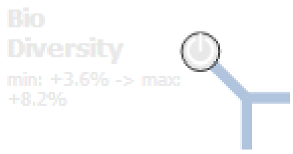


Figure 10: MOOViz tool in bio-fuel scenario –disabled objective

4.5.3 Policy Glyphs

A glyph on the polygon represents a solution of the multi-objective optimization problem, that resides on the Pareto fron. Each policy on the Pareto front is visualized by a circle inside the polygon. This circle is termed glyph.

Figure 11: MOOViz tool in bio-fuel scenario – policy glyph

Each glyph is divided into equal slices where each slice faces the value of its corresponding objective. The slice color is indicating to which objective it refers:

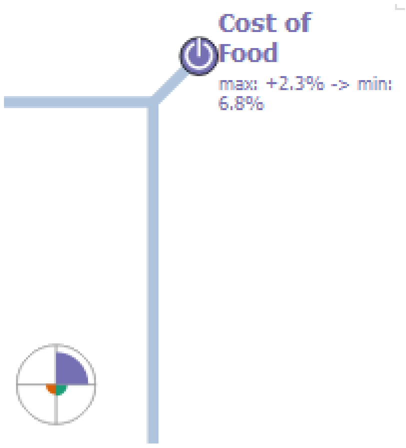


Figure 12: MOOViz tool in bio-fuel scenario – referring slice to objective

The size of the colored slice indicates how large the value is. The location of the policy glyph inside the polygon aimed at reflecting the “distance” from the optimal objective value. These locations are determined using a complex optimization process, and the final locations are a local optimum of this optimization process.
Note that the glyph location optimization problem is not the policy multi objective optimization problem and is aimed at optimizing visualization parameters such as the orientation of the points in space, the distance between them and their distance from the polygon corners (anchors).

When the mouse is placed over a certain policy glyph, the glyph’s border becomes bold, and its details presented:

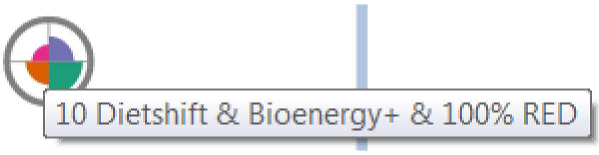


Figure 13: MOOViz tool in bio-fuel scenario – details of policy

And when the user clicks on a policy glyph, a tooltip window is shown. The tooltip window presents the values of the various objectives for that policy.

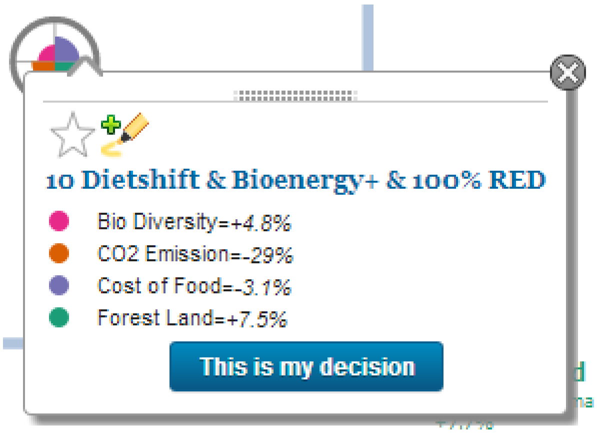


Figure 14: MOOViz tool in bio-fuel scenario – tooltip window details

In addition it allows the following actions:

- Adding the policy to the list of favorite policies: Usually the decided policy is chosen from the list of favorites. The list allows the user to concentrate only on a subset of the optimal policies that reflect her preferences. Adding a policy to the favorites set is done by clicking the star on the top left part of the window (). After clicking on the star, it changes its color to yellow and a small yellow star appears next to the policy glyph:

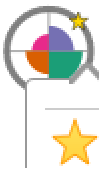


Figure 15: MOOViz tool in bio-fuel scenario – policy added to favorites

It should be noted that after this action, the window closes and in order to perform additional actions, the policy glyph needs to be clicked again.

- Highlight the policy. Highlight the policy can assist in reducing the set of policies from which favorites are chosen. In addition, when viewing policies on the lines view (see corresponding section), highlighting policies can assist in analysis of the information. This is done by clicking the highlight button at the left of the star on the top right part of the window (). The policy glyph turns to yellow and the window closes:

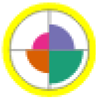


Figure 16: MOOViz tool in bio-fuel scenario – policy highlighted

- Mark a policy as the decided policy. This should be done after considering the favorite policies, and upon taking a final decision. is the choice is selected by clicking the “This is my decision” button at the lower part of the window of the chosen policy: 

  After clicking this button the window closes, and the policy glyph turns to bold blue:

Figure 17: MOOViz tool in bio-fuel scenario – policy marked as decision

Comments on policies:

- In certain cases, in the tooltip window opened by clicking the policy glyph, the application suggests other policies that may be more appealing. In this case, clicking on the blue text opens a larger window that allows performing tradeoff analysis. This will be explained in the continuation.

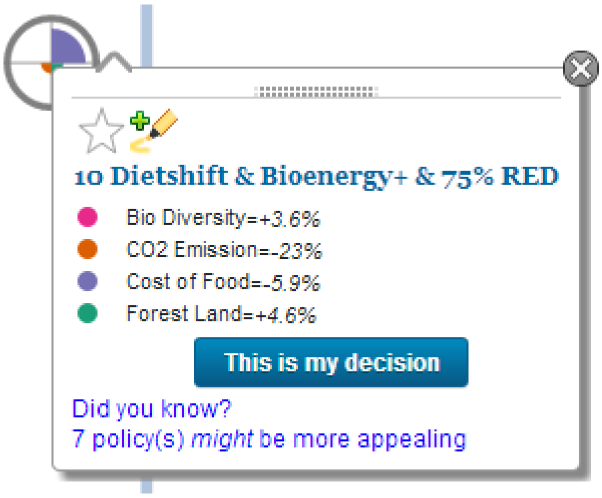


Figure 18: MOOViz tool in bio-fuel scenario – example of more appealing policies in window tooltip

- After the chosen policy is marked, its name appears on the top right part of the application, and the user can press the done button which indicates the session is done:

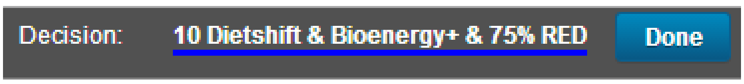


Figure 19: MOOViz tool in bio-fuel scenario – decision panel

When the mouse is placed on the name of the chosen policy, a red X appears to the left of the name. Clicking it will remove the chosen policy so that another policy can be decided instead:

Figure 20: MOOViz tool in bio-fuel scenario – delete decision example

- When an objective is disabled, some policies may be removed from the Pareto front. In this case, the slice related to the disabled objective will also be disabled:

Figure 21: MOOViz tool in bio-fuel scenario – policy glyph with a disabled objective

4.5.4 Range sliders

Range sliders are located to the left of the polygon. The range sliders allow the user to change the range of the various objectives to reflect her preferences:

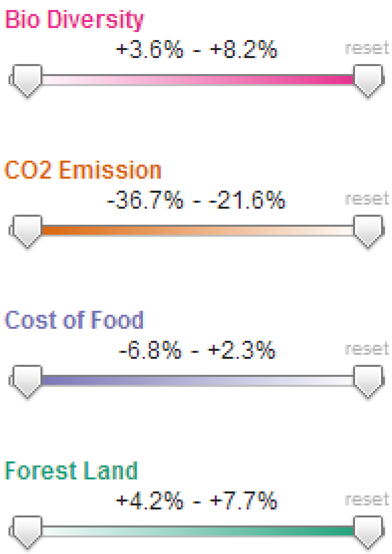


Figure 22: MOOViz tool in bio-fuel scenario – range sliders

Each range slider is associated with a certain objective (the colors and names correspond to the vertices of the polygon). Each slider has two buttons, one on the right hand side of the slider, and on the left hand side of the slider. By clicking the buttons and moving them along the slider, the range of legitimate values of the objective change. This allows filtering out policies that do not meet the values induced by the new range. The color change reflects the direction of the values in the objectives: the darker the color of the objective, the better it is for that objective. For example in the Figure 22 above, the aim is to maximize bio diversity (dark color is on the right) and to minimize CO2 Emission (dark color is on the left).

4.5.5 Optimal and Auto-excluded policies

Located in the top right part of the screen below the upper panel, these tabs are used to list all the available policies (divided to optimal and auto-excluded):

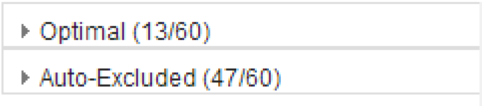


Figure 23: MOOViz tool in bio-fuel scenario – optimal and auto-excluded tabs

The Optimal tab contains the list of policies (solutions) that belong to the Pareto front of the multi objective optimization problem. The Auto-Excluded tab contains the policies (solutions) that were filtered out because they do not reside on the Pareto front (in other words – there are dominated by other policies).

Clicking on each of the tabs opens a detailed list of policies:

Figure 24: MOOViz tool in bio-fuel scenario – optimal policies tab

- When the mouse is over a certain policy, the corresponding policy glyph inside the polygon becomes bold gray:

Figure 25: MOOViz tool in bio-fuel scenario – example of gray policy glyph when mouse is placed over an optimal policy name

- Clicking the name of the policy that appears as a link, opens a window identical to the one opened when clicking a policy glyph inside the polygon. The same actions are available (see corresponding section).

- A policy can be added to the favorites list by clicking the star to the left of the policy name. In the list, the star becomes yellow, and a small star appears next to the policy glyph inside the polygon.

Figure 26: MOOViz tool in bio-fuel scenario – adding policy to favorites from optimal policies tab

- When a policy is added to the favorites list, sometimes more appealing policies exist. For example, there may be other Pareto optimal policies in which the values of one objective is slightly less appealing, but in other objectives it can be much more appealing[6]. In a similar fashion to the option in the window of the policy glyph, a message may appear on the bottom right part of the screen. This message notifies on options that can be more appealing. Clicking on the blue “Consider…” link opens a tradeoff analysis window. Its functionality will be described later.

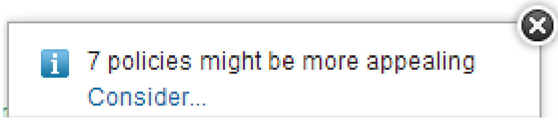


Figure 27: MOOViz tool in bio-fuel scenario – example of notification on existing appealing policies

- A highlighted policy appears highlighted in the list:

Figure 28: MOOViz tool in bio-fuel scenario – highlighted policy in the optimal policies tab

- When policies are excluded as a result of changing the slider values, the corresponding policy in the list is disabled:

Figure 29: MOOViz tool in bio-fuel scenario – excluded policy in the optimal policies tab

---

[6] For more information see the section related to tradeoff analysis

- When a policy is chosen ("This is my decision" in the window), it is marked with a blue underline:

Figure 30: MOOViz tool in bio-fuel scenario – decided policy in the optimal policies tab

- A policy can be a combination of subsets or all the above mentioned characteristics:

Figure 31: MOOViz tool in bio-fuel scenario – combination of excluded, highlighted, decided and favorite policy

The same operations can be done on the Auto-Excluded policies. Note that an auto excluded policy can still be added to the favorites list or chosen as the decided policy although it is not optimal. In addition, since these policies are not reflected visually in the polygon, the effects will only be seen in the list.

4.5.6 Top left panel

This panel consists of several action buttons. When the tool is uploaded, before any changes are done, all are disabled (top part).

After operations are done, subsets or all of them can be enabled (bottom part)

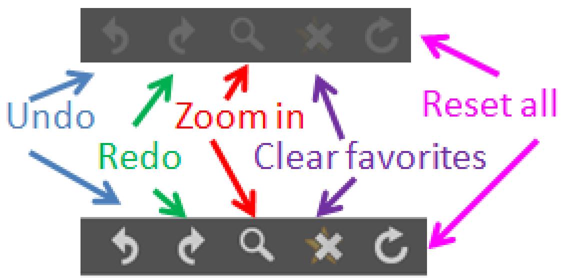


Figure 32: MOOViz tool in bio-fuel scenario – top left panel

The functionality of the various buttons:

- Undo – undo last operation. Multiple consecutive undo operations are allowed (history of operations is saved).
- Redo – redo last operation. Multiple consecutive redo operations are allowed (history of operations is saved).
- Zoom in – see corresponding section on zooming.
- Clear favorites – clear the list of favorite policies. Star tagging is removed from favorite policies.

- Reset all – get back to the initial state. In particular, favorite policies are cleared, zoomed views are removed, sliders are reset to initial ranges and filtered out measures are enabled.

4.5.7 Lines view

In addition to the polygon view (which is entitled map), there is an option to view the policies in lines. This view assists in visualization of the policies in a different fashion compared to the polygon representation. Changing the view is done by clicking on the view icons that are located on the top right part of the screen to the left of the optimal and auto-excluded lists:

Figure 33: MOOViz tool in bio-fuel scenario – moving from polygon to lines view

The left button is for map view and the right is for lines view. The current view is marked with a blue line below.

Below is a snapshot of the lines view:

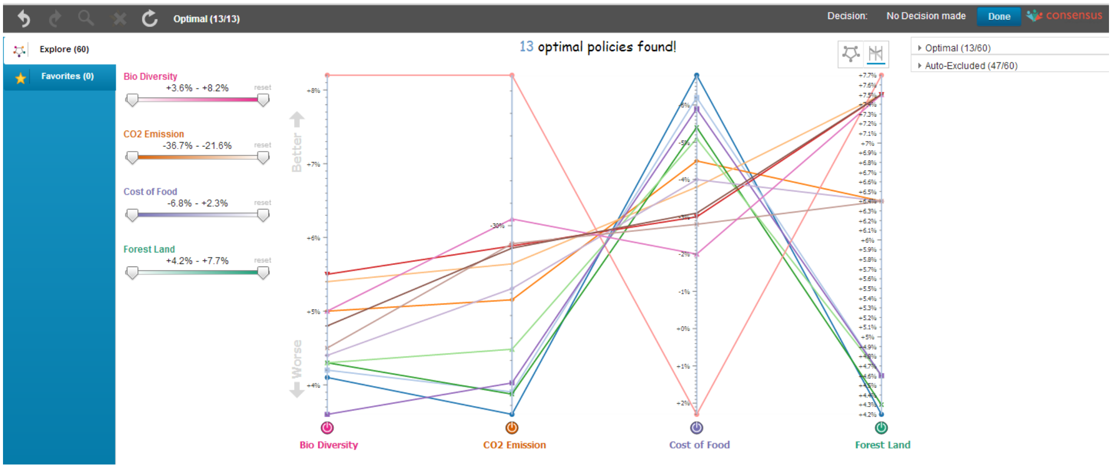


Figure 34: MOOViz tool in bio-fuel scenario – lines view

And a zoom on the lines:

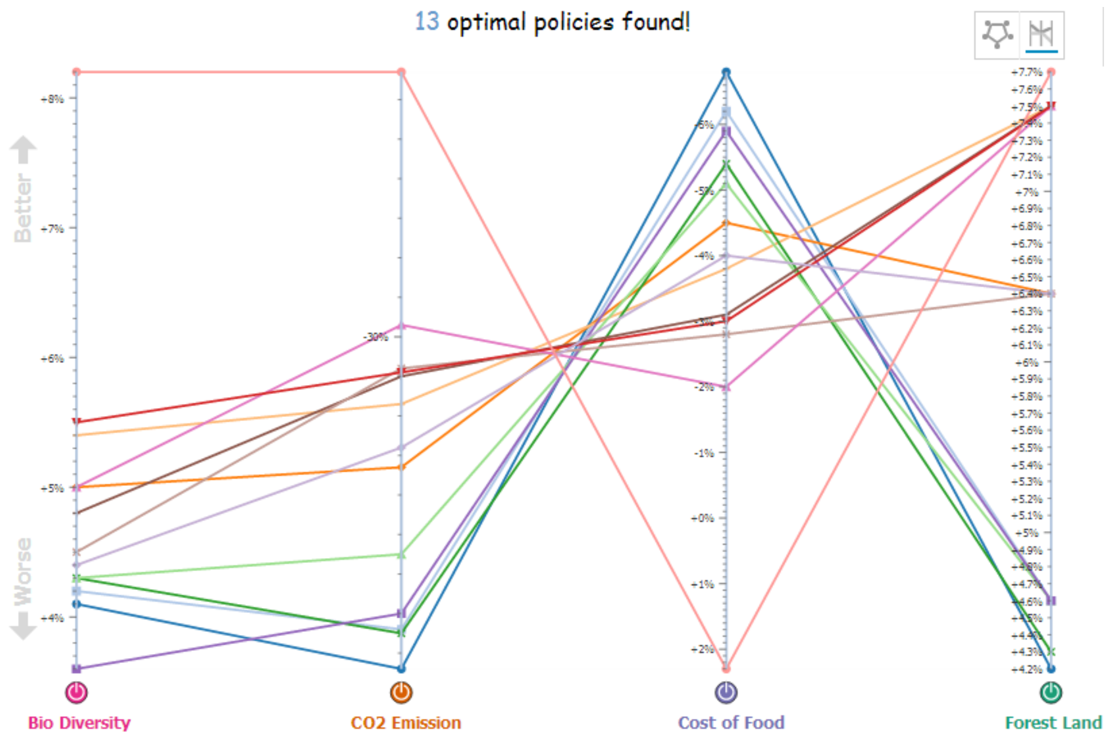


Figure 35: MOOViz tool in bio-fuel scenario – lines view, zoom on lines area

In the lines view, each vertical line corresponds to a certain objective. The names and colors correspond to the map (polygon) view and are listed below the line. Each policy is represented by a line of a different color. The points in which a policy meets the various objective lines corresponds to the value of the objective for this policy. The upper parts of the lines correspond to better values in the corresponding objective. When the mouse is moved on a certain policy, its line becomes thicker and its name appears:

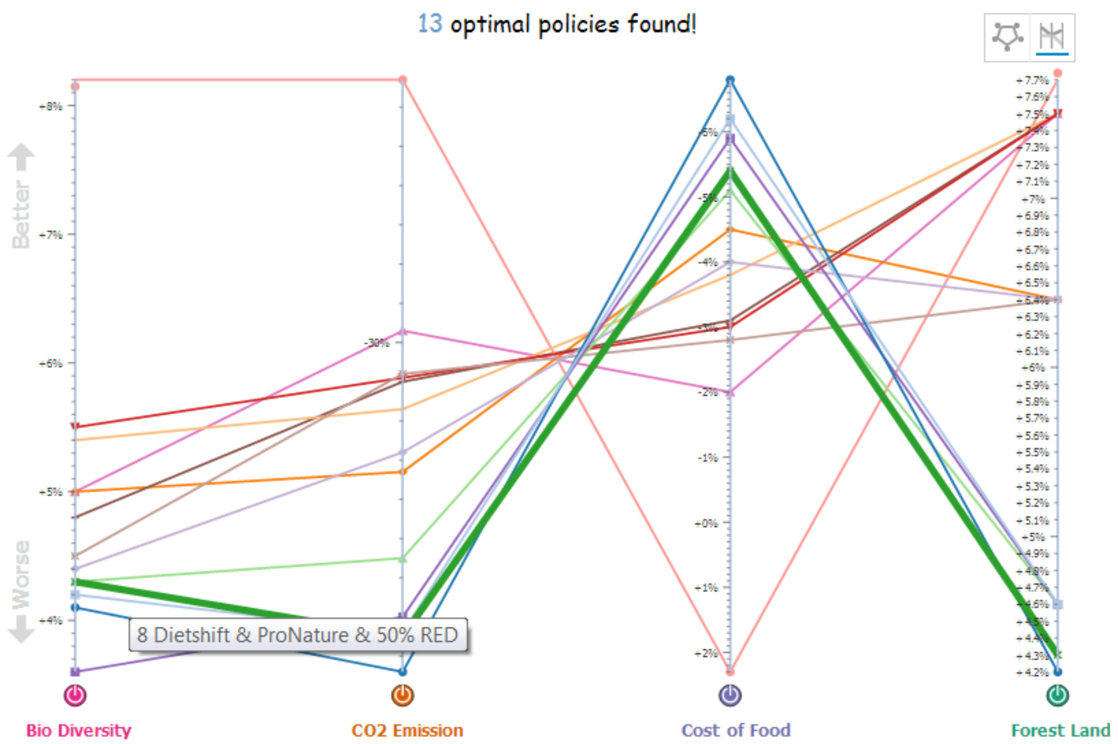


Figure 36: MOOViz tool in bio-fuel scenario – information on policy when mouse placed over corresponding lines

When clicking on the policy line, a window appears. It has the same options as the window that opens when clicking on a policy glyph:

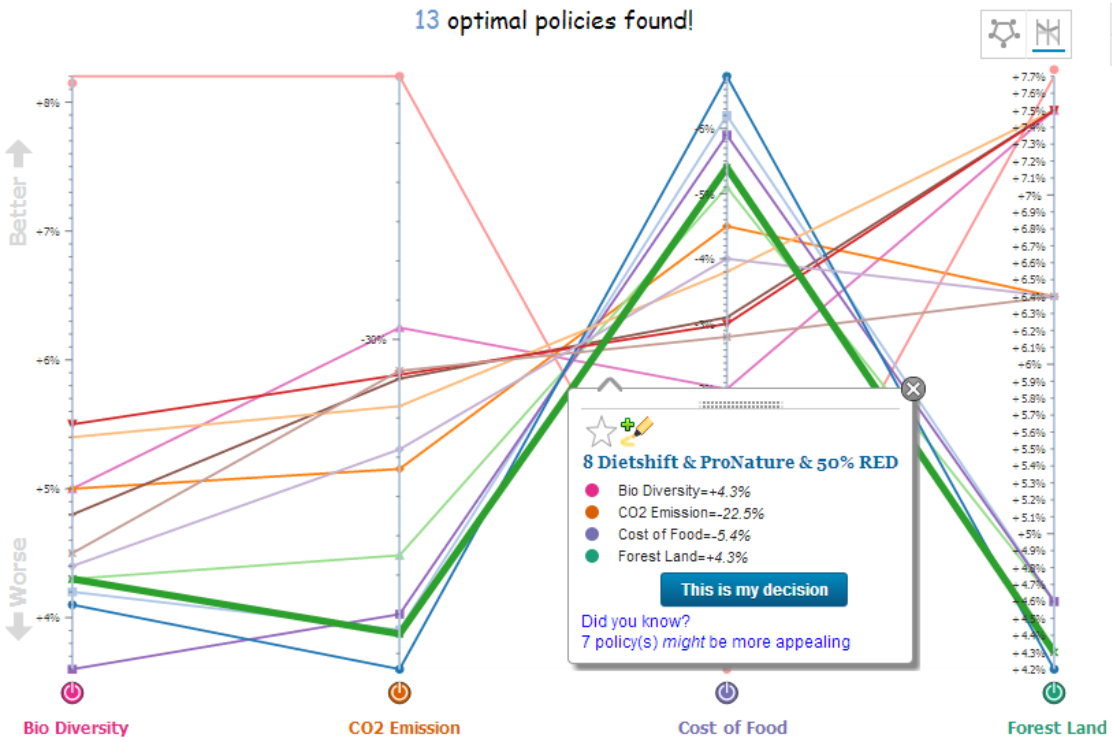


Figure 37: MOOViz tool in bio-fuel scenario – tooltip window in lines view

The functionality is identical to the one in the map (polygon) view. Below is a snapshot with various options used on a few policies:

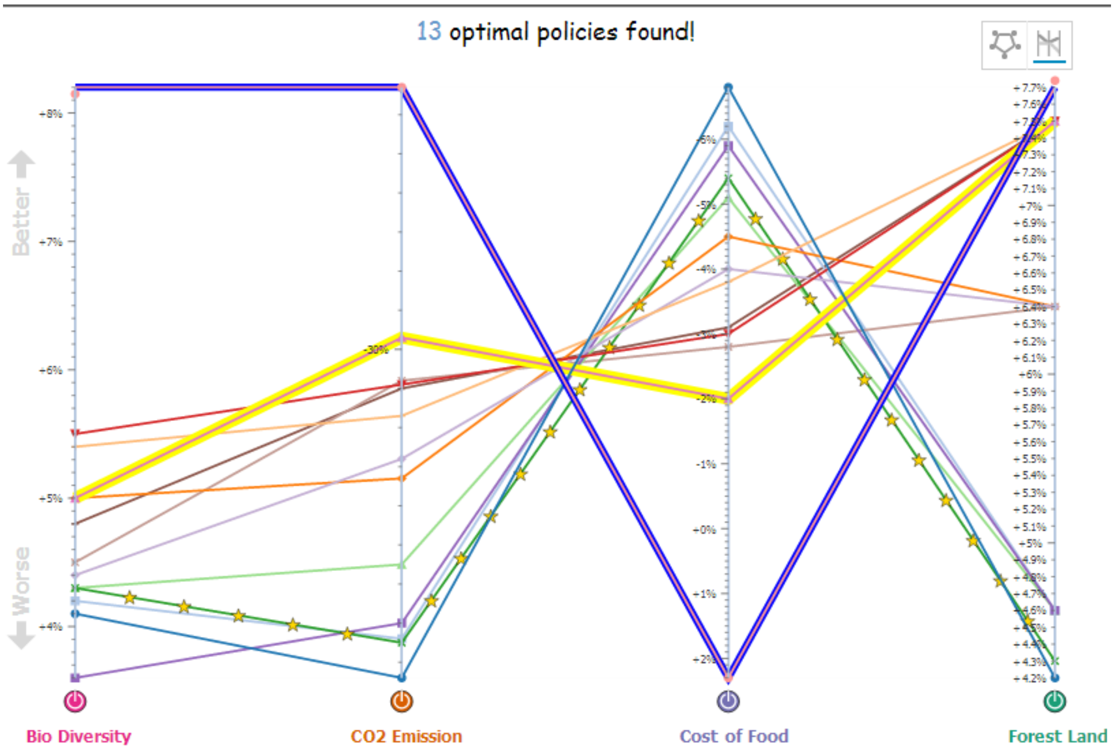
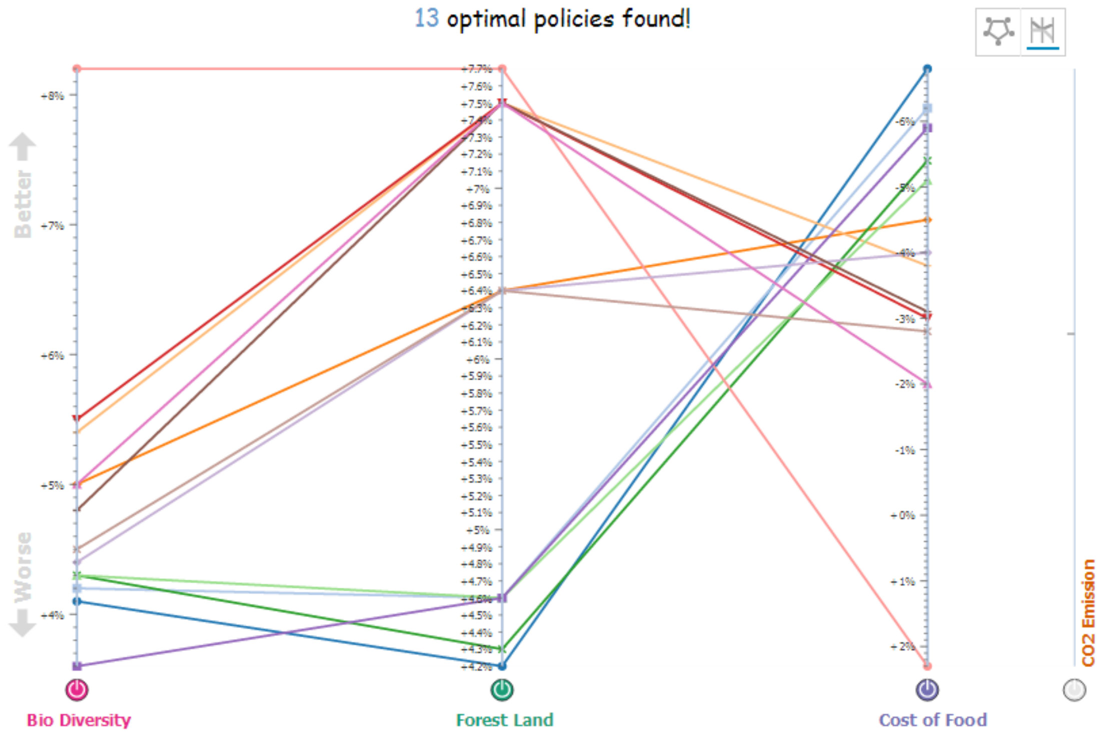


Figure 38: MOOViz tool in bio-fuel scenario – examples of highlighted, favorite and decided policy in lines view

The dark green policy is a favorite policy (yellow stars). The purple policy is highlighter (yellow marker). The pink policy is the decision (highlighter with blue).

All the options available in the map (polygon) view are valid in the lines view (e.g., optimal and auto-excluded policies, reset/undo/redu etc). Slider functionality remains, and moving sliders can disable policies.

When an objective is disabled (clicking on the name of the objective), the lines are reordered to that the disabled objective is moved to the right and the enabled objectives are in on the left part of the screen:



Finally, objectives can be switched to desired location. In other words suppose we want the cost of food objective to be on the left side of CO2 emission objective. We can drag the vertical line of cost of food between the lines of CO2 emission and bio diversity, and the objectives are reordered according to the new order:

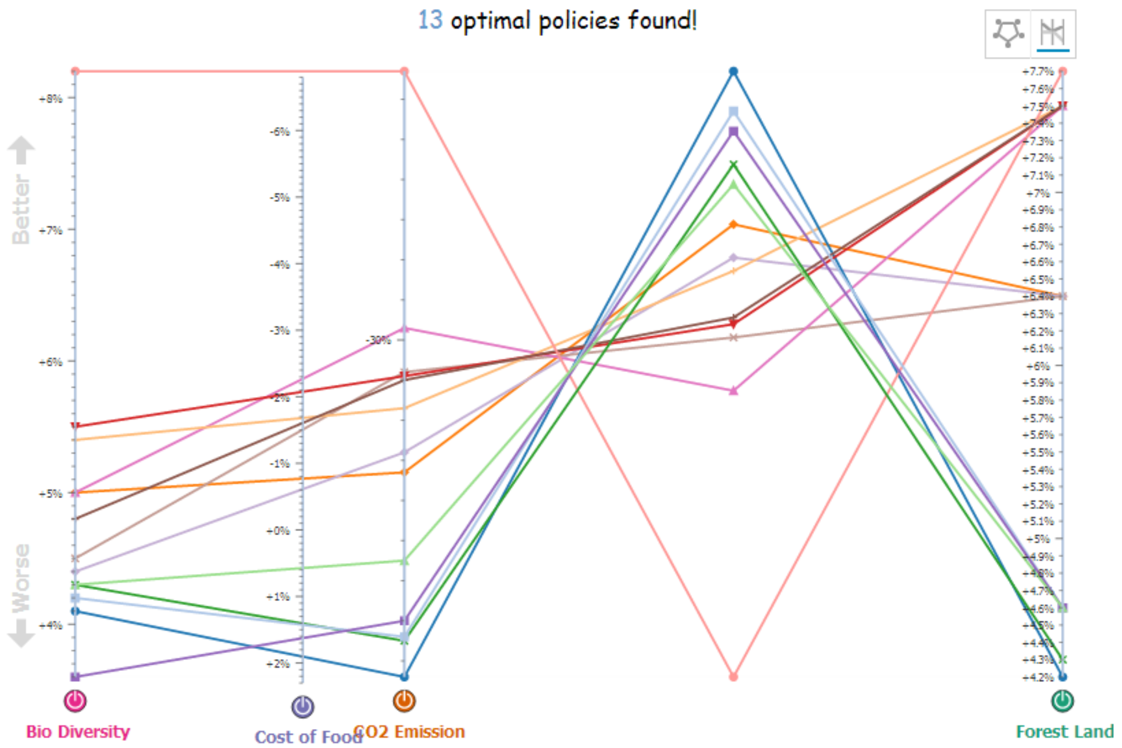


Figure 39: MOOViz tool in bio-fuel scenario – rearranging order of objectives – in process

And after automatic rearrange:

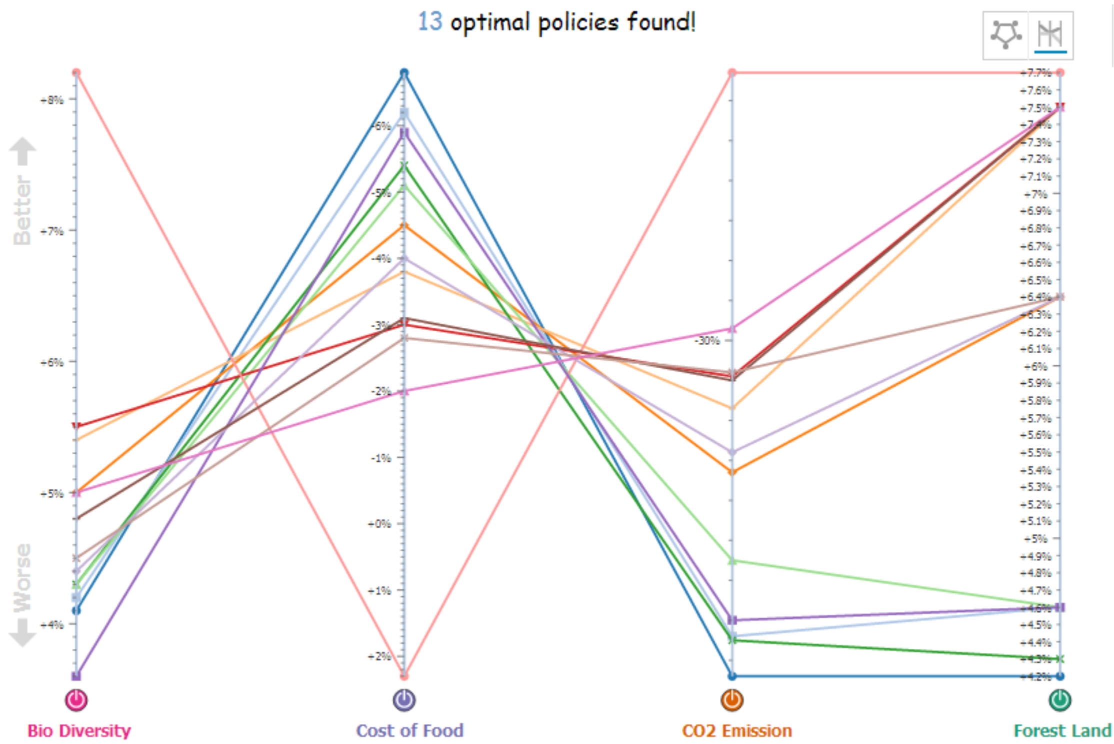


Figure 40: MOOViz tool in bio-fuel scenario – rearranging order of objectives – after reorder

4.5.8 Favorites pane

When one or more policies are marked as favorites, there is an option to move to Favorites pane. The favorites view allows analysis of the favorites polices as a step towards decision on the chosen policy. By clicking on the Favorites pane on the top left side of the screen, below the top left panel:

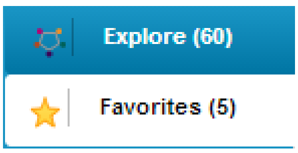


Figure 41: MOOViz tool in bio-fuel scenario – favorites pane

After pressing, the tab turns white (active). The favorites pane is always shown in lines view:

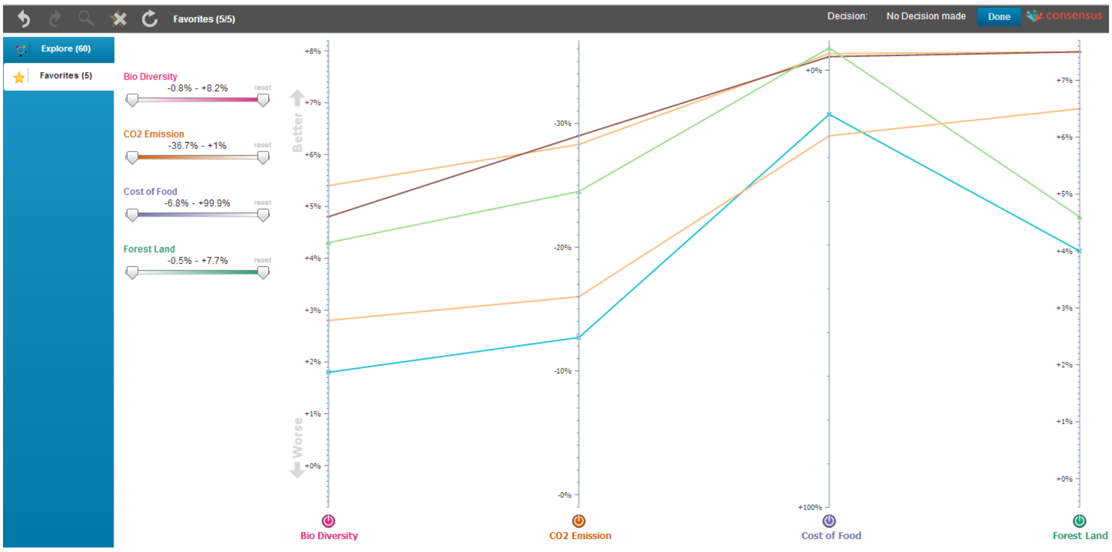


Figure 42: MOOViz tool in bio-fuel scenario – favorites view

Note that here, there are 3 optimal policies (3 “top” ones) and 2 auto-excluded policies (2 “bottom” ones). Here is visible why the auto-excluded policies do not reside on the Pareto front – the optimal policies dominate them (there is no objective in which they equal or are better than the optimal)

All the functionality discussed in the lines view holds here (name when mouse is over policy, window when clicking, sliders, highlight, decision, redu/undo/reset etc). Note that clicking on the favorite button of a policy removes it from the list of favorite policies and it disappears from the view.

4.5.9 Tradeoff analyzer

The tradeoff analyzer allows the user to check tradeoffs between various policies and supports the process of decision on which policy to choose (or add to favorites/highlight). As explained in the policies section, sometimes when clicking on a policy glyph, a window may appear, containing a recommendation on other policies that may be more appealing:

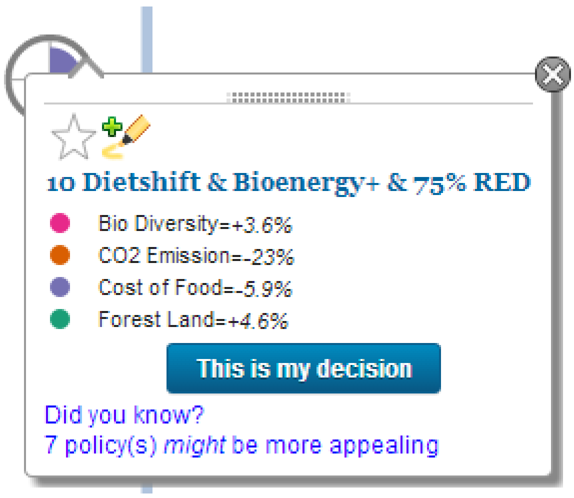


Figure 43: MOOViz tool in bio-fuel scenario – examples of more appealing policies

Clicking on the blue link at the bottom of the window, a large application window opens:

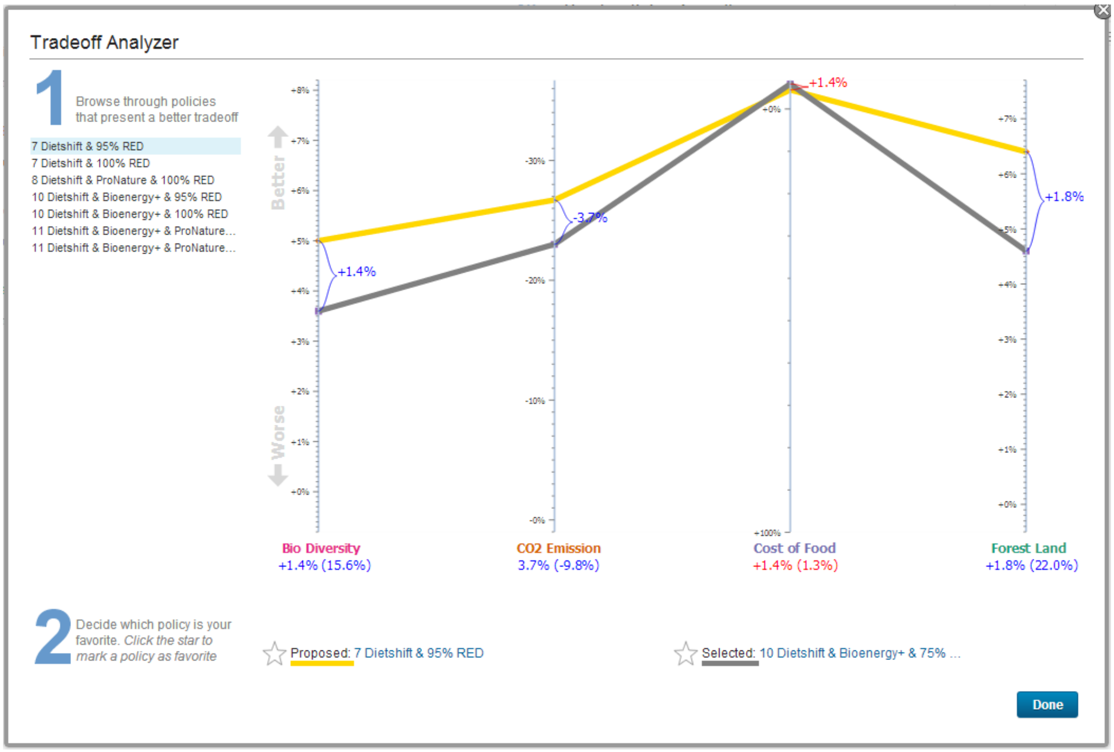


Figure 44: MOOViz tool in bio-fuel scenario – tradeoff analyzer

The right hand side of the window contains a view similar to the lines view. In this view two policies are presented: the policy clicked by the user, and one of the suggested policies that may be more appealing. For each objective, the difference in percentage of the objective value is plotted. Browsing between the suggested appealing policies is done by clicking on the policy names on the left hand side of the screen. This is considered as the first part of the tradeoff analysis. At the bottom of the window, the proposed policy and selected policy (from the appealing set) are listed with their names. It is possible to add them to the favorites list (using the star to the left of the name). Adding the policy to the favorites list would turn the star to yellow and mark the policy with stars:

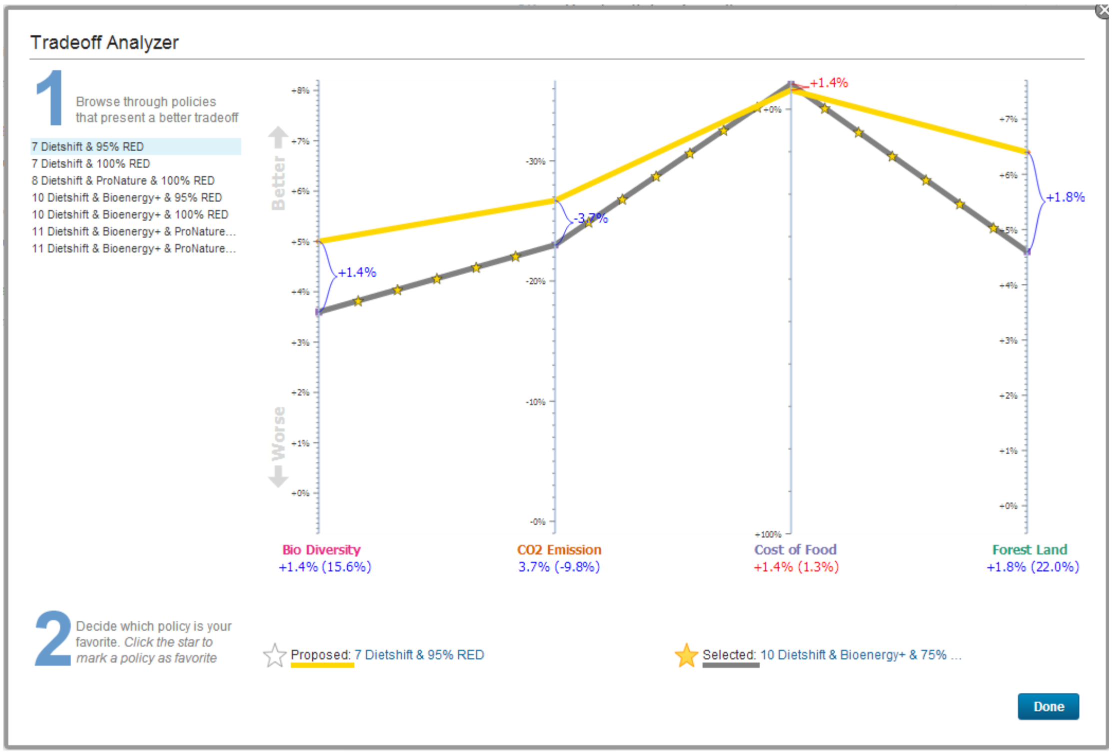


Figure 45: MOOViz tool in bio-fuel scenario – tradeoff analyzer – adding policy to favorites

When the tradeoff analysis is done (after going over the chosen appealing policies), the user clicks the done button at the bottom right side of the window and returns to the main application page.

Access to the tradeoff analysis can be done in a similar fashion from the lines view. In this case, clicking the done button closes the window and returns to the lines view.

4.5.10 Zoom in on views

When the sliders are moved and some solutions are disabled, it is possible to zoom in. Zoom in allows analysis of chosen subset of policies. This subset is induced by the changed in the range of various objectives (see section on range sliders for additional information). Zooming is done by clicking the button on the top left panel:

Figure 46: MOOViz tool in bio-fuel scenario – zoom in button

The button is only enabled when there are disabled policies as a result of changing the sliders. Pressing the zoom will open a zoomed view, in which disabled policies do not appear in the polygon (as opposed to transparent glyphs), and the slider's range is automatically adjusted to the new range of values.

This is a snapshot before zoom:

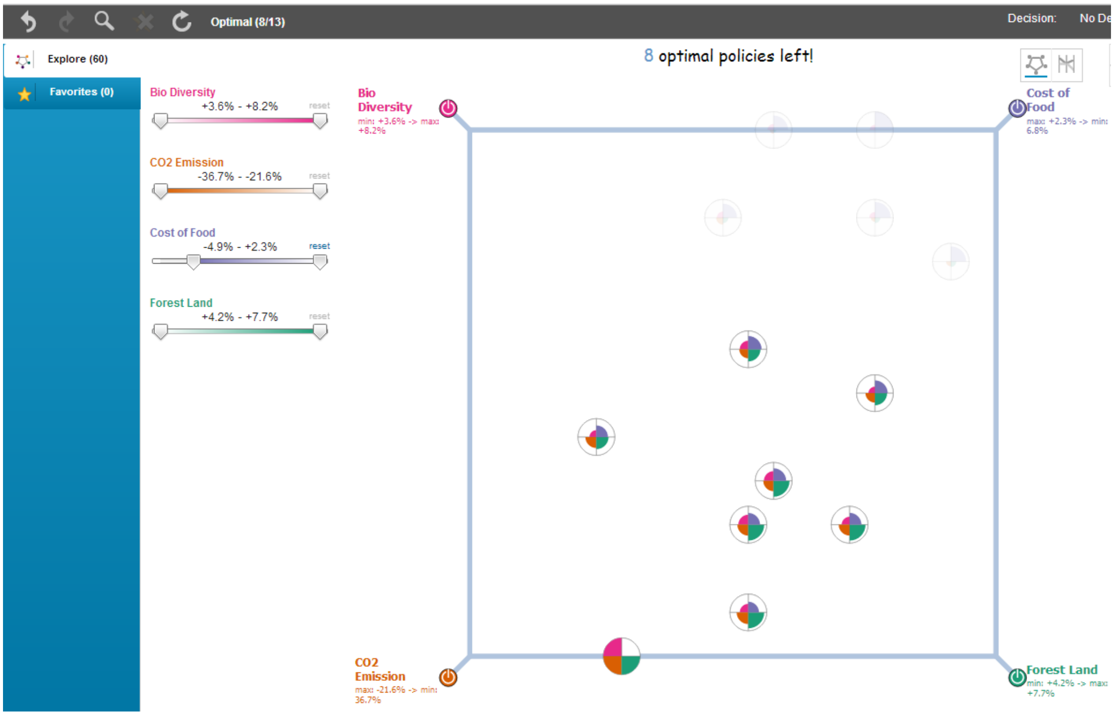


Figure 47: MOOViz tool in bio-fuel scenario – snapshot before zoom in

In particular, note the disabled policies in the top right part of the polygon. The zoom button is now enabled. Clicking it performs the zoom in and the obtained snapshot is as follows:

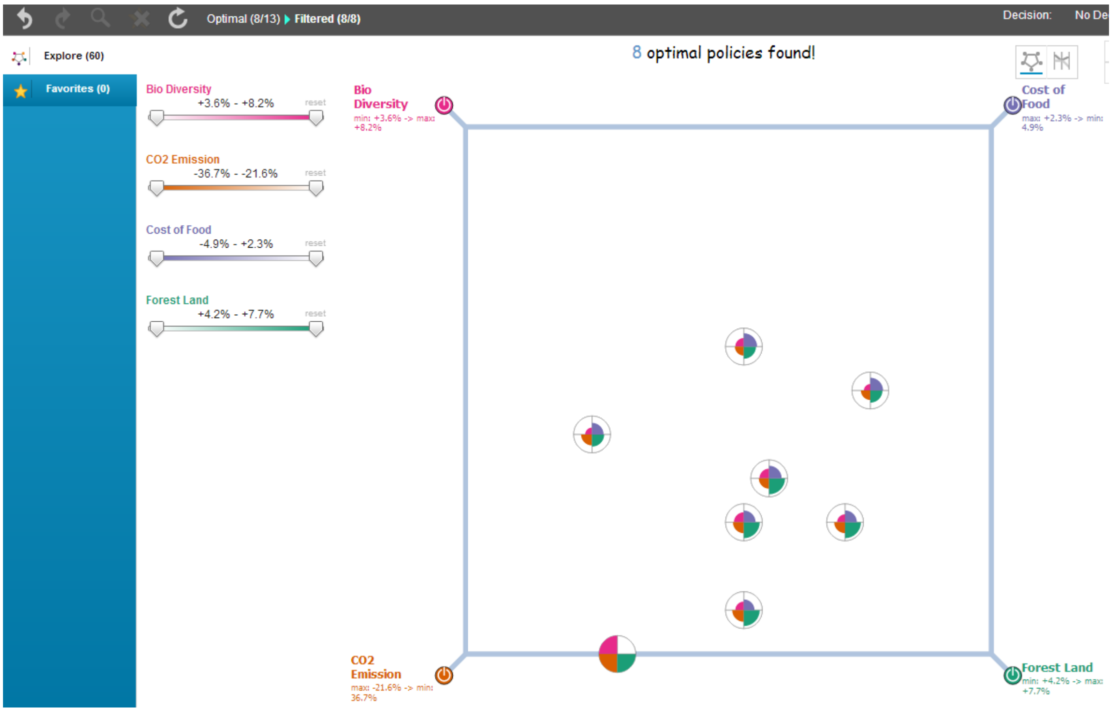


Figure 48: MOOViz tool in bio-fuel scenario – snapshot after zoom in

There are three changes compared to the previous view. First, the disabled policies disappeared. Second, the range of the sliders is changed to reflect the change in value (see the slider corresponding to cost of food in the previous snapshot). Finally, on the top panel next to the Optimal link, a link to the current view was added:

Figure 49: MOOViz tool in bio-fuel scenario – link to zoom in view

In this view, the shown policies are the subset obtained by changing the range of values of the objectives. Clicking the Optimal link moves us back to the previous view. In addition, we can change the sliders in the filtered (zoomed) view and zoom in again. In this case the link chain would look as follows:

Figure 50: MOOViz tool in bio-fuel scenario – link to nested zoom in

We can continue the zoom in the same fashion.

4.6 Using MOOViz Tool for Transportation Scenario

4.6.1 Introduction

The MOOViz tool for the Transportation scenario is aimed at assisting policy makers to explore policy alternatives and better understand the trade-offs between objectives, and coming into educated decision that is taking into account the entire aspects.

Note that a link to the prototype is provided within Deliverable D4.1.1 Optimization and Visual Analytics Prototypes (due to confidentiality considerations).

Following the link will open the application main page:

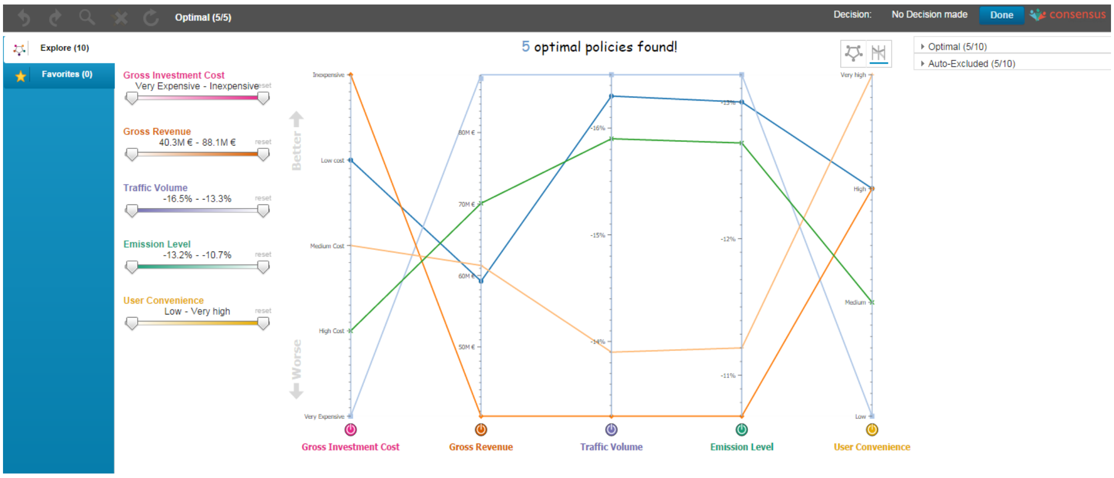


Figure 51: MOOViz tool in transportation scenario – snapshot of application

The page is composed of several viewports. Figure 52 contains the names of the various components that are described in the following sections.

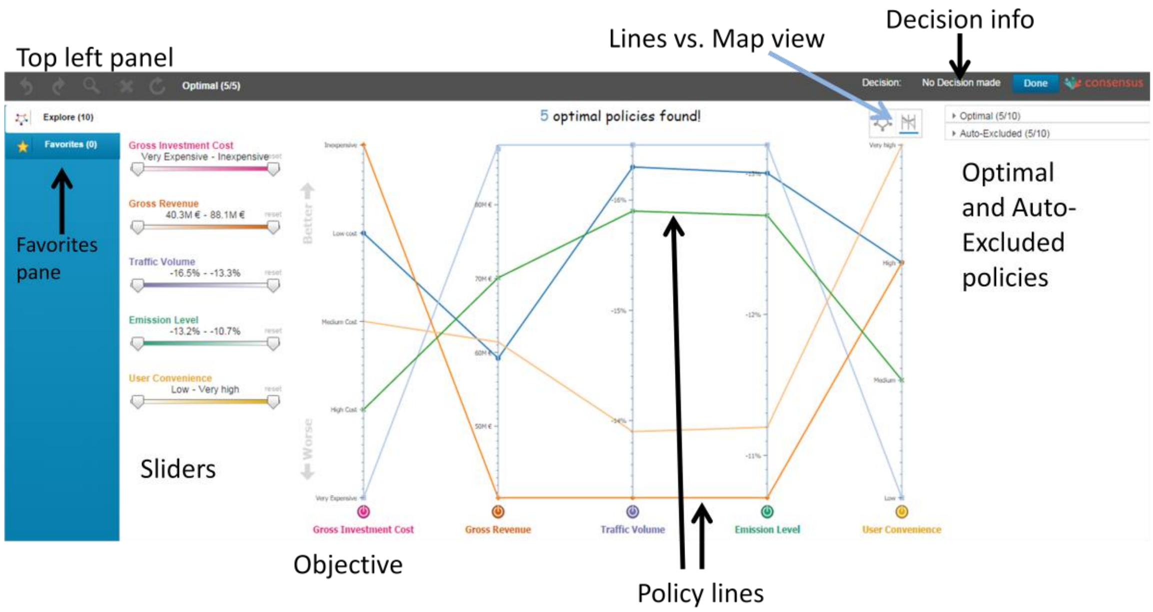


Figure 52: MOOViz tool in transportation scenario – main components

4.6.2 Lines view

The main view of the application is the lines view. This view assists in visualization of the various policies. Each vertical line corresponds to a certain objective function, and the various pseudo horizontal lines represent different policies. Each Objective function has a different color, and the name of the policy is listed below the corresponding line. The values of an objective function are derived from the values of the objective function of all Pareto optimal policies. This range is listed on the vertical lines, where values at the top correspond to better values.

Below is a snapshot of the lines view:

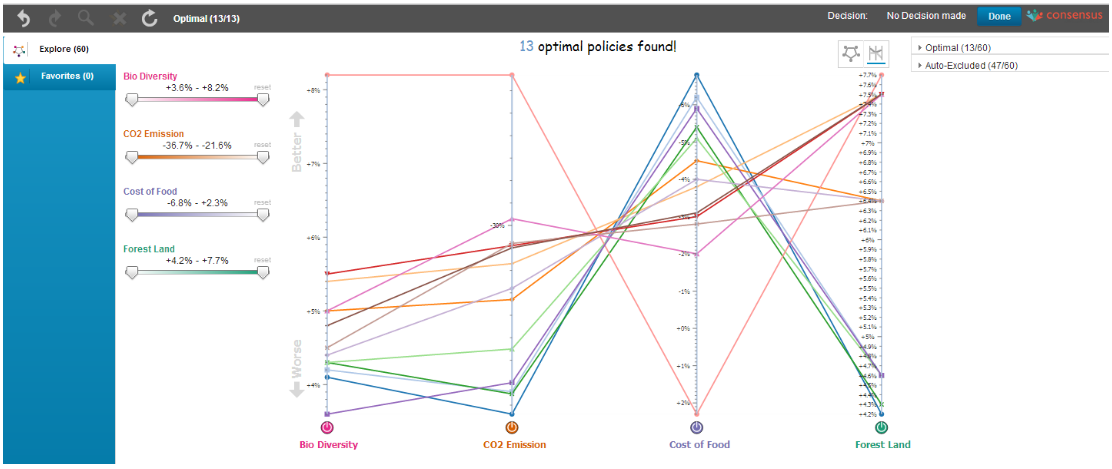


Figure 53: MOOViz tool in transportation scenario – lines view

And a zoom on the lines:

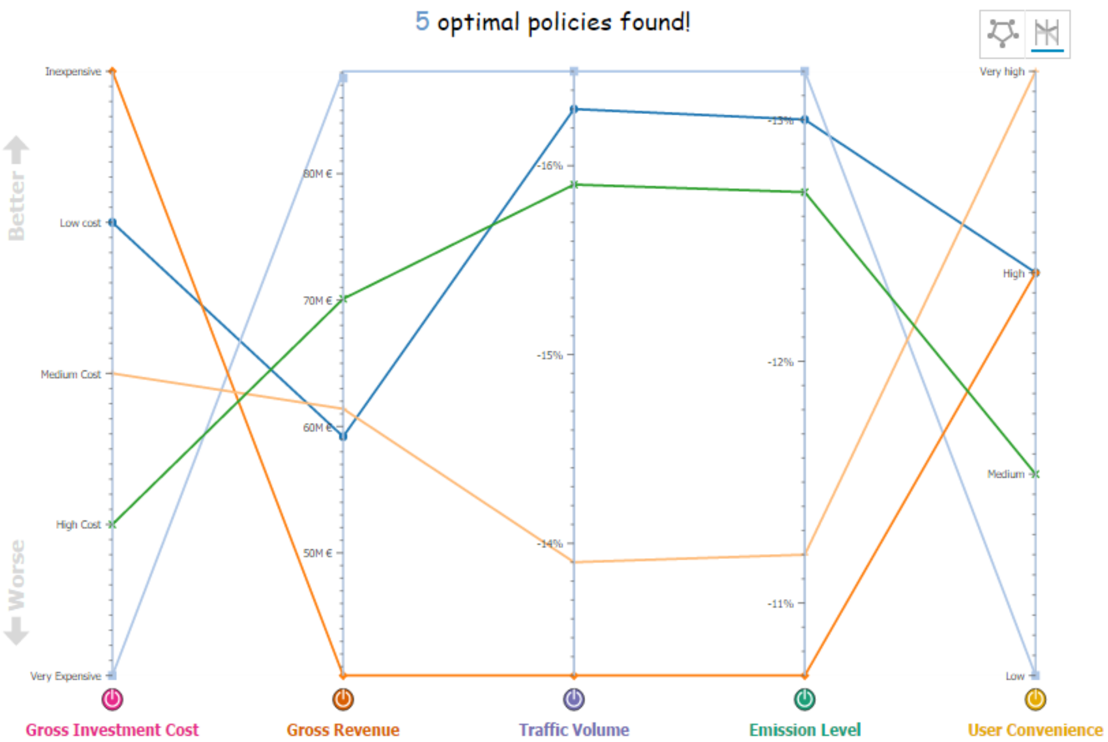


Figure 54: MOOViz tool in transportation scenario – lines view, zoom on lines area

Objective can be disabled by clicking on the name of the objective. In this case the vertical objective lines are reordered to that the disabled objective is moved to the right and the enabled objectives are in on the left part of the screen:

When an objective is disabled (clicking on the name of the objective), the vertical objective lines are reordered to that the disabled objective is moved to the right and the enabled objectives are in on the left part of the screen:

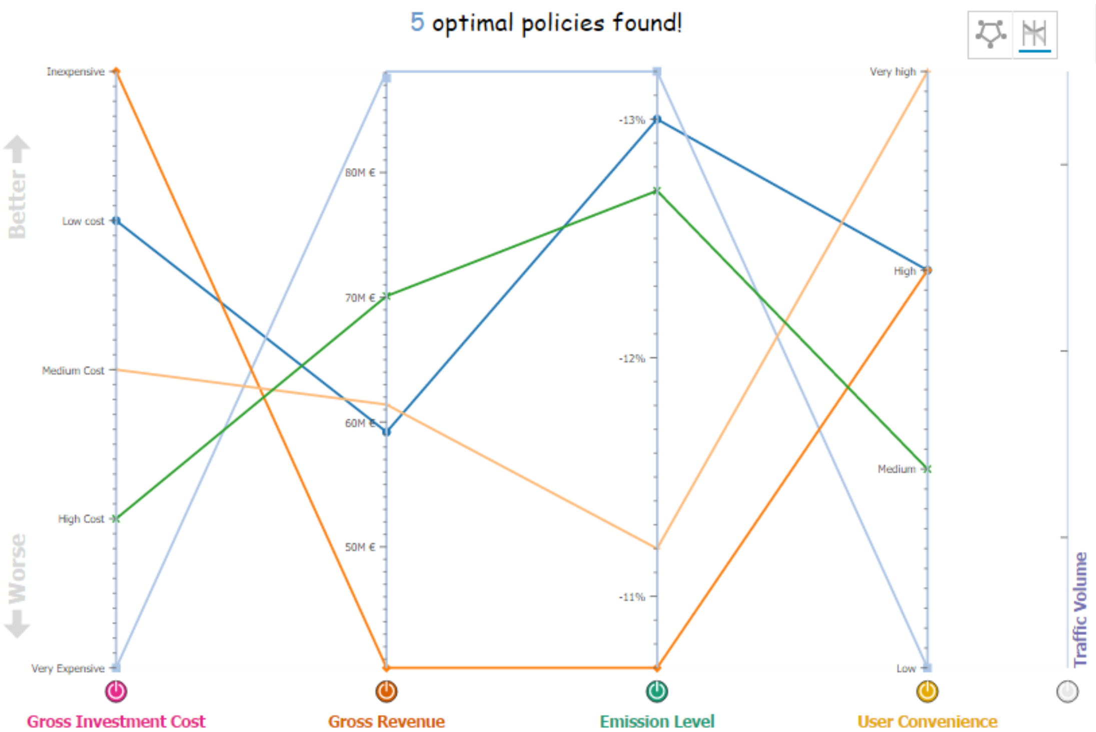


Figure 55: MOOViz tool in transportation scenario – disabled objective

Finally, objectives can be switched to desired location. In other words suppose we want the emission level objective to be on the left side of gross revenue objective. We can drag the vertical line of emission level between the lines of gross revenue and gross investment cost, and the objectives are reordered according to the new order:

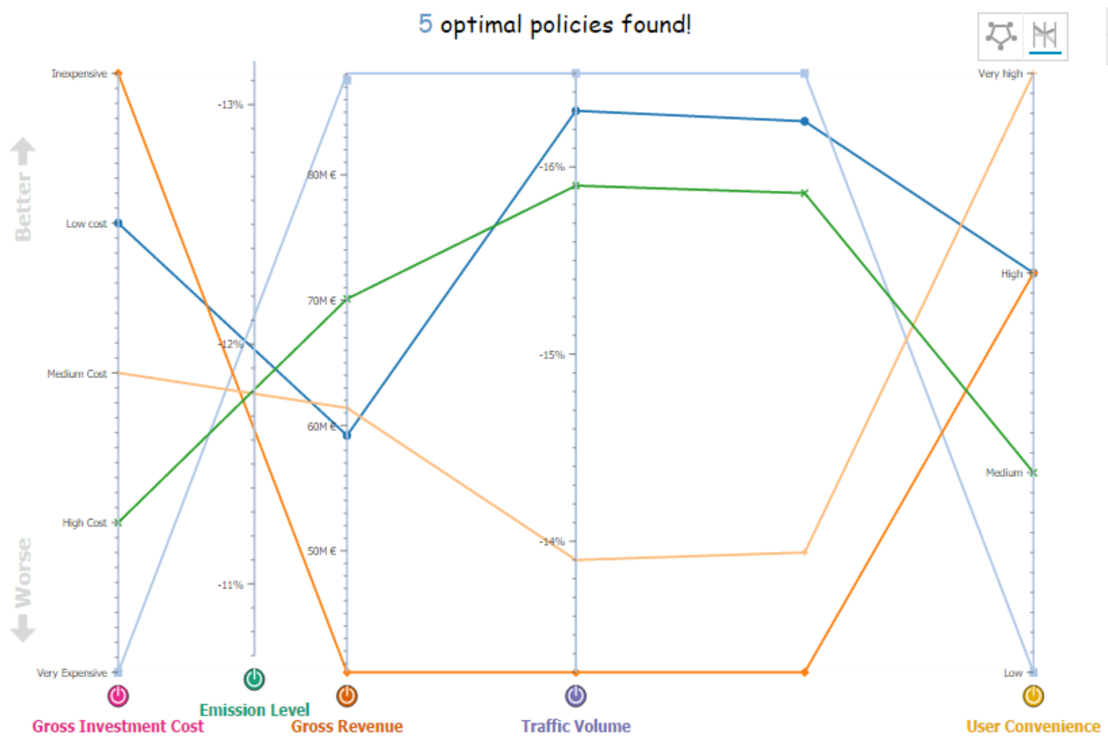


Figure 56: MOOViz tool in transportation scenario – rearranging order of objectives – in process

And after automatic rearrange:

Figure 57: MOOViz tool in transportation scenario – rearranging order of objectives – after reorder

4.6.3 The policy line

Each policy is represented by a line of a different color. The points in which a policy line meets the various objective lines corresponds to the value of the objective for this policy. When the mouse is moved on a certain policy, its line becomes thicker and its name appears:

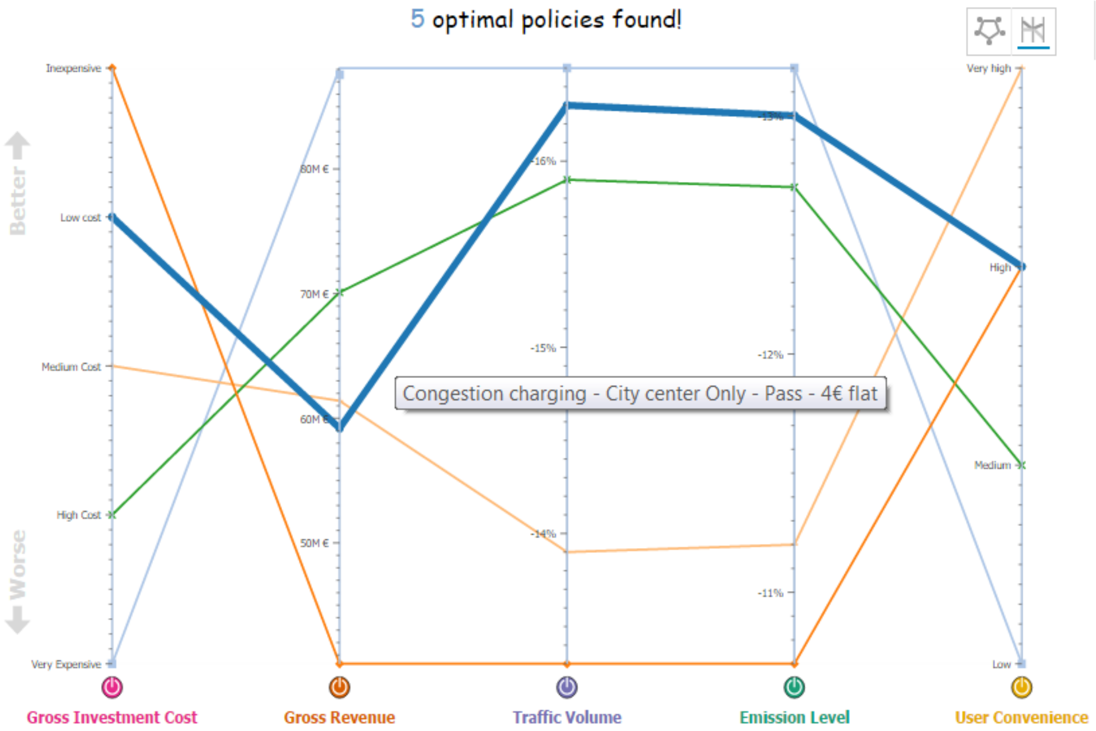


Figure 58: MOOViz tool in transportation scenario – details of policy

When clicking on the policy line, a tooltip window is shown. The tooltip window presents the values of the various objectives for that policy.

Figure 59: MOOViz tool in transportation scenario – tooltip window details

In addition it allows the following actions:

- Adding the policy to the list of favorite policies. Usually the decided policy is chosen from the list of favorites. The list allows the user to concentrate only on a subset of the optimal policies that reflect her preferences. Adding a policy to the favorites set is done by clicking the star on the top left part of the window ( ). After clicking on the star, it changes its color to yellow and the policy line has small yellow stars over it:

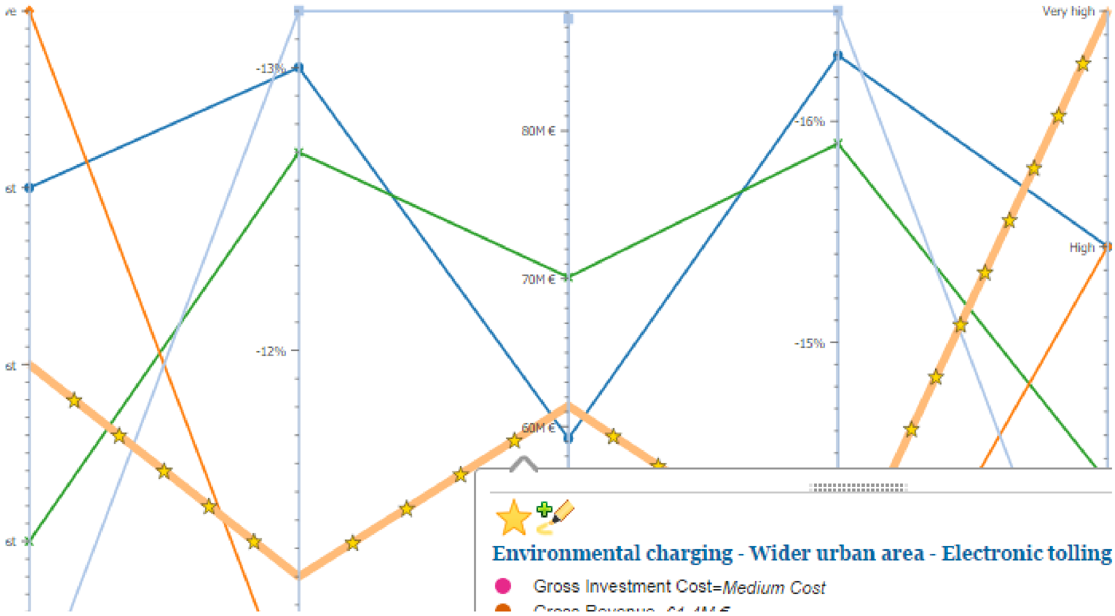


Figure 60: MOOViz tool in transportation scenario – policy added to favorites

It should be noted that after this action, the window closes and in order to perform additional actions, the policy line needs to be clicked again.

- Highlight the policy. Highlight the policy can assist in reducing the set of policies from which favorites are chosen. In addition, when viewing policies on the lines view, highlighting policies can assist in analysis of the information. This is done by clicking the highlight button at the left of the star on the top right part of the window ( ). The policy line turns to yellow and the window closes:

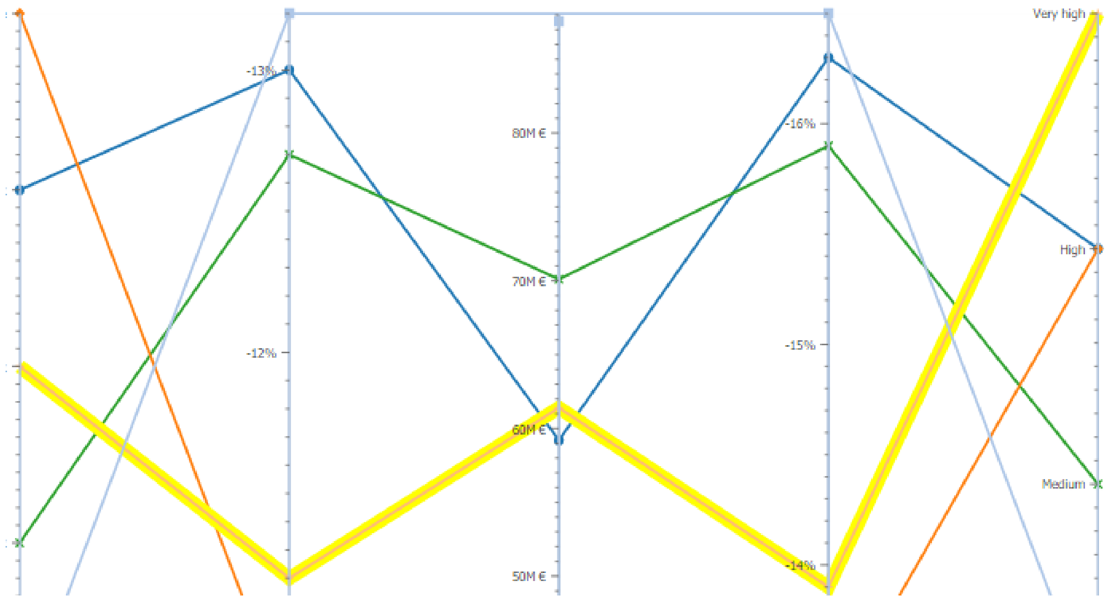


Figure 61: MOOViz tool in transportation scenario – policy highlighted

- Mark a policy as the decided policy. This should be done after considering the favorite policies, and upon taking a final decision. is the choice is selected by clicking the "This is my decision" button at the lower part of the window of the chosen policy: 

  After clicking this button the window closes, and the policy line turns to bold blue:

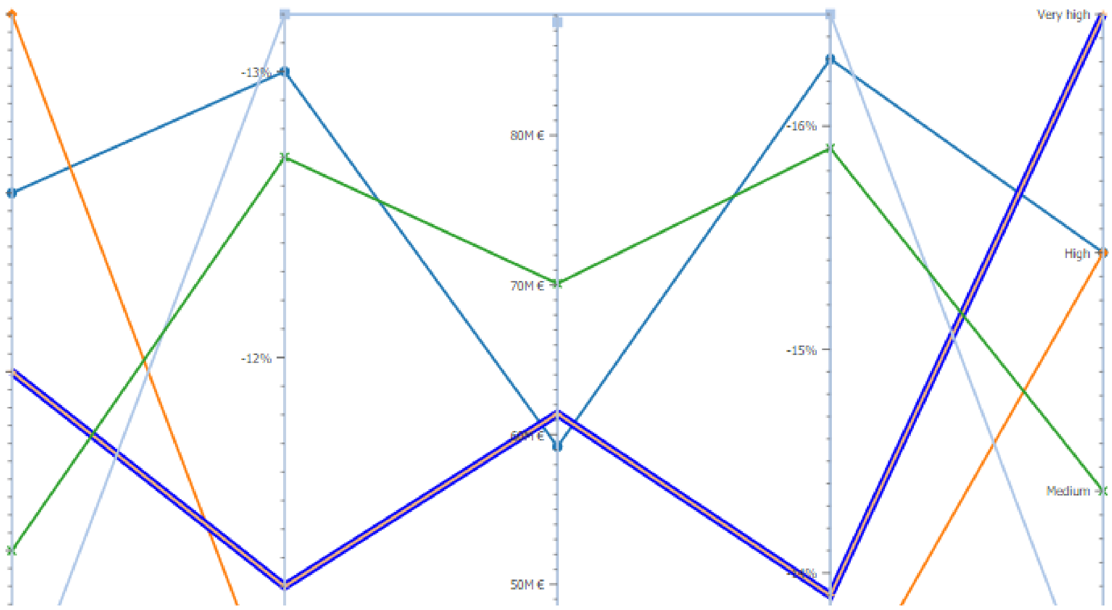


Figure 62: MOOViz tool in transportation scenario – policy marked as decision

Comments on policies:

- In certain cases, in the tooltip window opened by clicking the policy line, the application suggests other policies that may be more appealing. In this case, clicking on the blue text opens a larger window that allows performing tradeoff analysis.

This will be explained in the continuation. Note that it in the current transportation scenario this option is not demonstrated, but may happen in the future. To see an example please see comments in the section related to policy glyphs in the bio fuel scenario.

- After the chosen policy is marked, its name appears on the top right part of the application, and the user can press the done button which indicates the session is done:

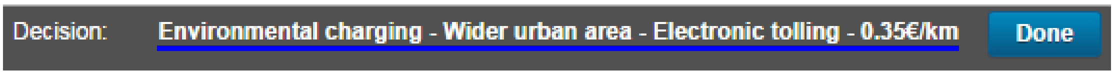


Figure 63: MOOViz tool in transportation scenario – decision panel

When the mouse is placed on the name of the chosen policy, a red X appears to the left of the name. Clicking it will remove the chosen policy so that another policy can be decided instead:

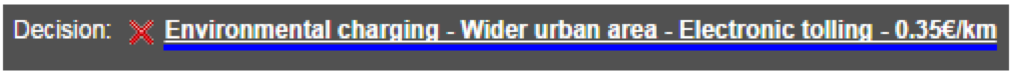


Figure 64: MOOViz tool in transportation scenario – delete decision example

4.6.4 Range sliders

The range slides are located to the left of the lines view. The range sliders allow the user to change the range of the various objectives to reflect her preferences:

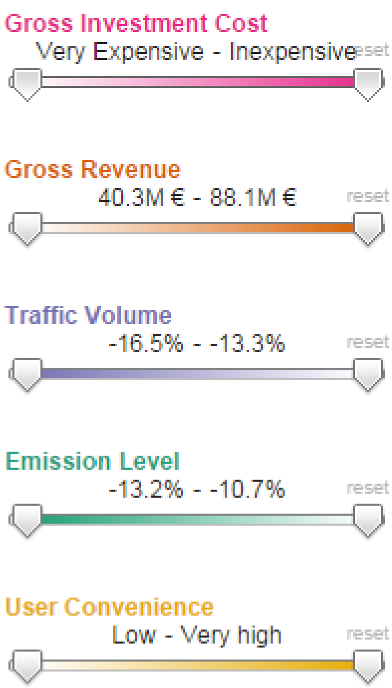


Figure 65: MOOViz tool in transportation scenario – range sliders

Each range slider is associated with a certain objective (the colors and names correspond to the vertical lines of the lines view). Each slider has two buttons, one on the right hand side of the slider, and on the left hand side of the slider. By clicking the buttons and moving them along the slider, the range of legitimate values of the objective change. This allows filtering out policies that do not meet the values induced by the new range. In other words, changing the range of an objective can exclude a policy from the optimal set, and in such cases the line corresponding to that policy is disabled.

The color change of the sliders reflects the direction of the values in the objectives: the darker the color of the objective, the better it is for that objective. For example in the Figure 65 above, the aim is to maximize gross revenue (dark color is on the right) and to minimize traffic volume (dark color is on the left).

4.6.5 Optimal and Auto-excluded policies

Located in the top right part of the screen below the upper panel, these tabs are used to list all the available policies (divided to optimal and auto-excluded):

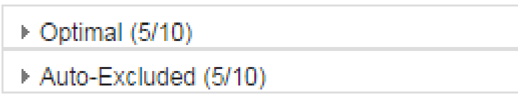


Figure 66: MOOViz tool in transportation scenario – optimal and auto-excluded tabs

The Optimal tab contains the list of policies (solutions) that belong to the Pareto front of the multi objective optimization problem. The Auto-Excluded tab contains the policies (solutions) that were filtered out because they do not reside on the Pareto front (in other words – there are dominated by other policies).

Clicking on each of the tabs opens a detailed list of policies:

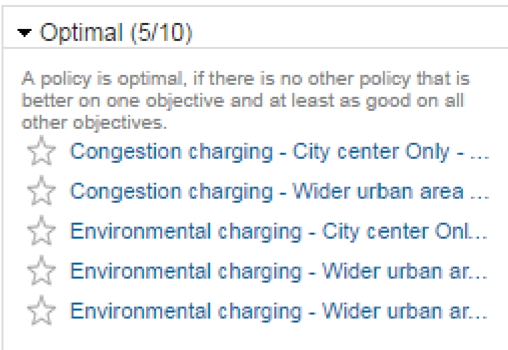


Figure 67: MOOViz tool in transportation scenario – optimal policies tab

- When the mouse is over a certain policy, the corresponding policy line becomes bold:

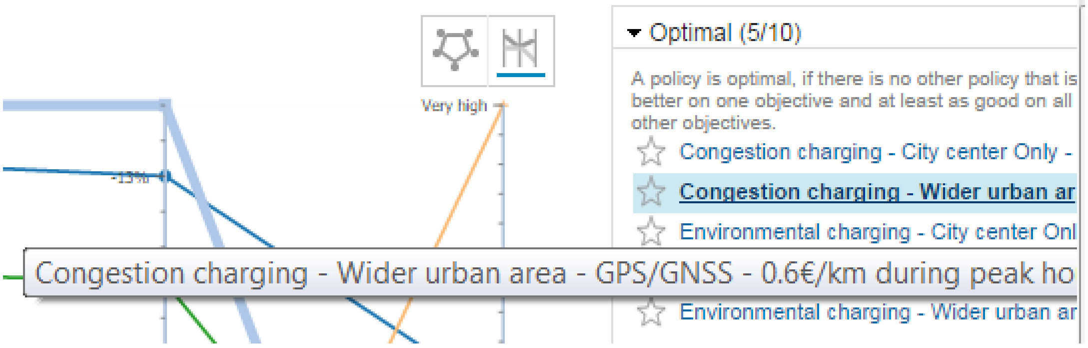


Figure 68: MOOViz tool in transportation scenario – example of bold policy line when mouse is placed over an optimal policy name

- Clicking the name of the policy that appears as a link, opens a window identical to the one opened when clicking a policy line inside the lines view. The same actions are available (see corresponding section).

- A policy can be added to the favorites list by clicking the star to the left of the policy name. In the list, the star becomes yellow, and small stars appear along the policy line.

Figure 69: MOOViz tool in transportation scenario – adding policy to favorites from optimal policies tab

- When a policy is added to the favorites list, sometimes more appealing policies exist. For example, there may be other Pareto optimal policies in which the values of one objective is slightly less appealing, but in other objectives it can be much more appealing[7]. In a similar fashion to the option in the window of the policy line, a message may appear on the bottom right part of the screen. This message notifies on options that can be more appealing. Clicking on the blue "Consider…" link opens a tradeoff analysis window. Its functionality will be described later. Note that as stated earlier, the option for "more appealing policies exploration" is not described in the current transportation scenario. It only happens when clicking an auto excluded policy:

[7] For more information see the section related to tradeoff analysis

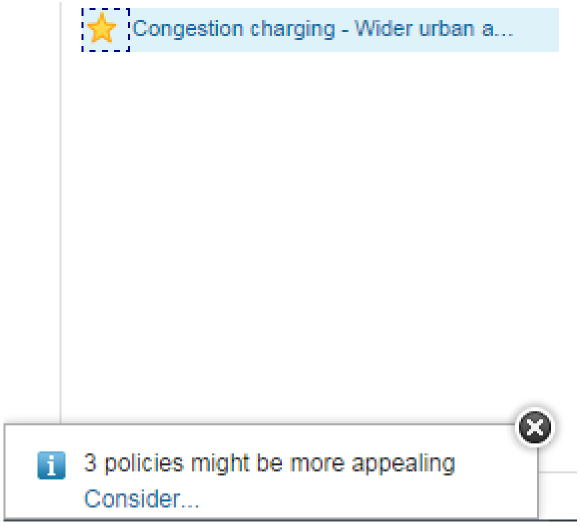


Figure 70: MOOViz tool in transportation scenario – example of notification on existing appealing policies

- A highlighted policy appears highlighted in the list:

Figure 71: MOOViz tool in transportation scenario – highlighted policy in the optimal policies tab

- When policies are excluded as a result of changing the slider values, the corresponding policy in the list is disabled:

Figure 72: MOOViz tool in transportation scenario – excluded policy in the optimal policies tab

- When a policy is chosen ("This is my decision" in the window), it is marked with a blue underline:

Figure 73: MOOViz tool in transportation scenario – decided policy in the optimal policies tab

- A policy can be a combination of subsets or all the above mentioned characteristics:

Figure 74: MOOViz tool in transportation scenario – combination of excluded, highlighted, decided and favorite policy

The same operations can be done on the Auto-Excluded policies. Note that an auto excluded policy can still be added to the favorites list or chosen as the decided policy although it is not optimal. In addition, since these policies are not reflected visually in the lines view, the effects will only be seen in the list.

4.6.6 Top left panel

This panel consists of several action buttons. When the tool is uploaded, before any changes are done, all are disabled (top part).

After operations are done, subsets or all of them can be enabled (bottom part)

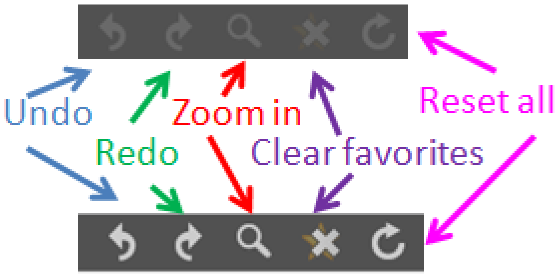


Figure 75 MOOViz tool in transportation scenario – top left panel

The functionality of the various buttons:

- Undo – undo last operation. Multiple consecutive undo operations are allowed (history of operations is saved).
- Redo – redo last operation. Multiple consecutive redo operations are allowed (history of operations is saved).
- Zoom in – see corresponding section on zooming.
- Clear favorites – clear the list of favorite policies. Star tagging is removed from favorite policies.

Reset all – get back to the initial state. In particular, favorite policies are cleared, zoomed views are removed, sliders are reset to initial ranges and filtered out measures are enabled.

4.6.7 The Polygon

In addition to the lines view, there is an option to view the policies in a polygon (also termed map). This view assists in visualization of the policies in a different fashion compared to the lines representation. Changing the view is done by clicking on the view icons that are located on the top right part of the screen to the left of the optimal and auto-excluded lists:

Figure 76: MOOViz tool in transportation scenario – moving from lines view to polygon

The left button is for map view and the right is for lines view. The current view is marked with a blue line below.

The main viewport of this view contains a polygon, with a vertex for each objective function. The corners form a symmetric polygon, which is used to visualize, in two dimensions, the natural high dimensional space in which each dimension represents a different objective function. In each corner, the name of the objective function and its range of values are presented:

Figure 77: MOOViz tool in transportation scenario – objective visualization

The order of the values indicates whether we maximize or minimize the objective. If the first value is “min”, the objective is being maximized. If the first value is “max”, the objective is being minimized:

Figure 78: MOOViz tool in transportation scenario – optimization direction

By clicking the colored circle, the corresponding objective function is disabled:

Figure 79: MOOViz tool in transportation scenario – disabled objective

4.6.8 Policy Glyphs

A glyph on the polygon represents a solution of the multi-objective optimization problem, that resides on the Pareto front of the optimization problem. Each policy on the Pareto front is visualized by a circle inside the polygon. This circle is termed glyph. Each policy glyph corresponds to a line policy in the lines view. It is a different representation of a Pareto optimal policy.

Figure 80: MOOViz tool in transportation scenario – policy glyph

Each glyph is divided into equal slices where each slice faces the value of its corresponding objective. The size of the colored slice indicates how large is the value. The slice color is indicating to which objective it refers:

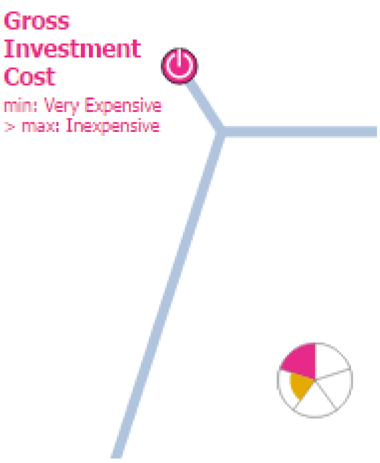


Figure 81: MOOViz tool in transportation scenario – referring slice to objective

The location of the policy glyph inside the polygon aimed at reflecting the “distance” from the optimal objective value. These locations are determined using a complex optimization process, and the final locations are a local optimum of this optimization process.
Note that the glyph location optimization problem is not the policy multi objective optimization problem and is aimed at optimizing visualization parameters such as the orientation of the points in space, the distance between them and their distance from the polygon corners (anchors).

Clicking a policy glyph provides the same functionality of clicking a policy line. When the mouse is placed over a certain policy glyph, the glyph’s border becomes bold, and its details presented:

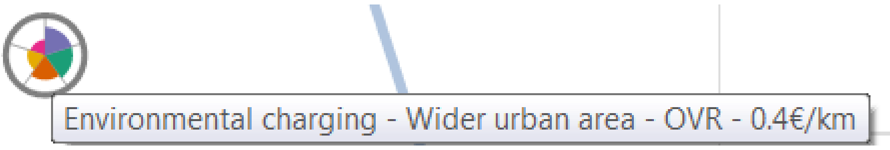


Figure 82: MOOViz tool in transportation scenario – details on policy

When the user clicks on a policy glyph, a tooltip window is shown. The tooltip window presents the values of the various objectives for that policy.

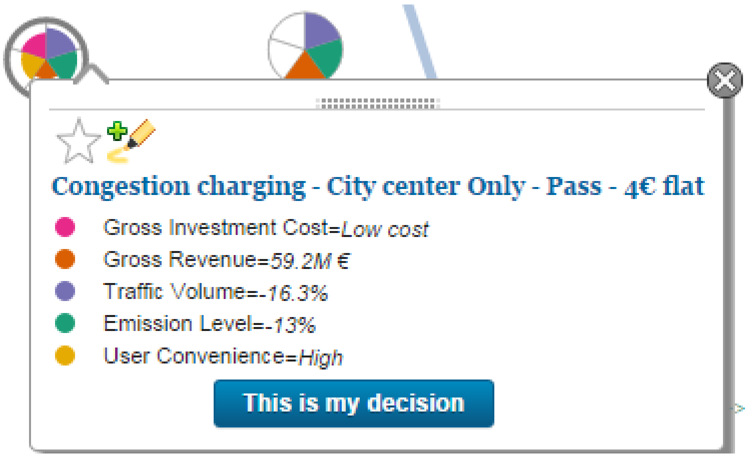


Figure 83: MOOViz tool in transportation scenario – tooltip window details

In addition it allows the same actions as in the lines view (adding to favorites, highlight, decision).

4.6.9 Favorites pane

When one or more policies are marked as favorites, there is an option to move to Favorites pane. The favorites view allows analysis of the favorites polices as a step towards decision on the chosen policy. By clicking on the Favorites pane on the top left side of the screen, below the top left panel:

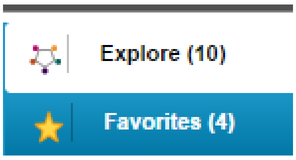


Figure 84: MOOViz tool in transportation scenario – favorites pane

After pressing, the tab turns white (active). The favorites pane is always shown in lines view:

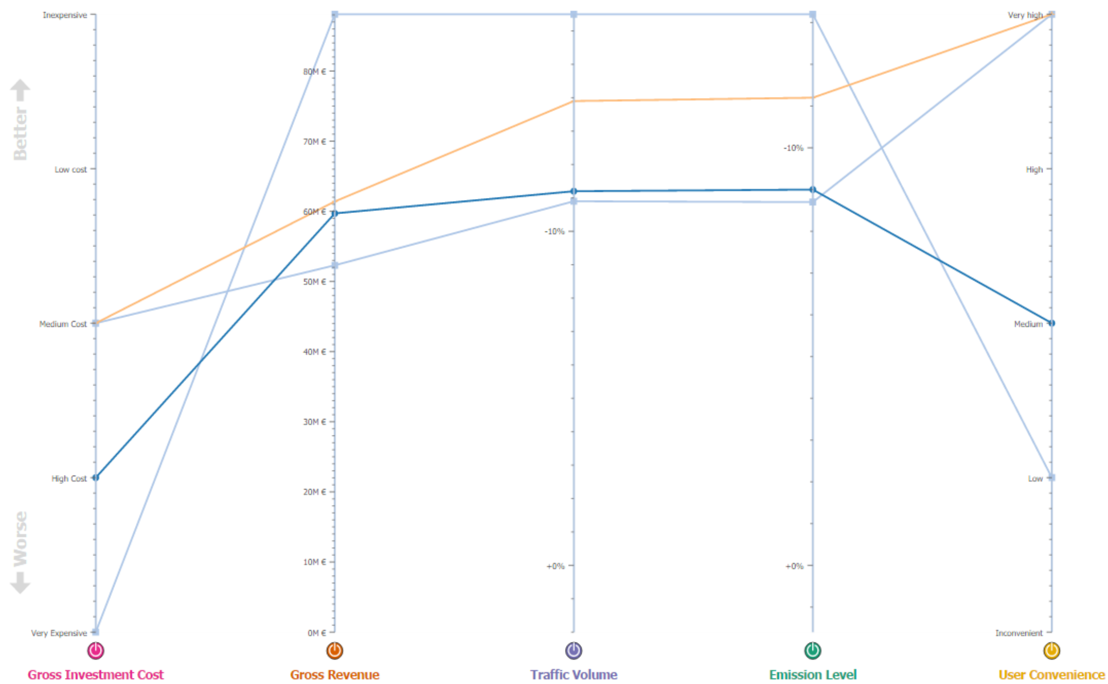


Figure 85: MOOViz tool in transportation scenario – favorites view

Note that here, there are 2 optimal policies (2 “top” ones) and 2 auto-excluded policies (2 “bottom” ones). Here is visible why the auto-excluded policies do not reside on the Pareto front – the optimal policies dominate them (there is no objective in which they equal or are better than the optimal)

All the functionality discussed in the lines view holds here (name when mouse is over policy, window when clicking, sliders, highlight, decision, redu/undo/reset etc). Note that clicking on the favorite button of a policy removes it from the list of favorite policies and it disappears from the view.

4.6.10 Tradeoff analyzer

The tradeoff analyzer allows the user to check tradeoffs between various policies and supports the process of decision on which policy to choose (or add to favorites/highlight). As explained in the policies section, sometimes when clicking on a policy line, or an auto excluded policy, a window may appear, containing a recommendation on other policies that may be more appealing:

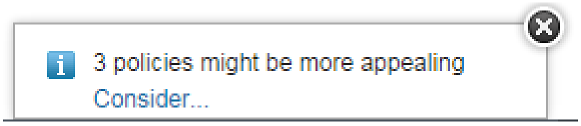


Figure 86: MOOViz tool in transportation scenario – example for more appealing policies

Clicking on the blue link at the bottom of the window, a large application window opens:

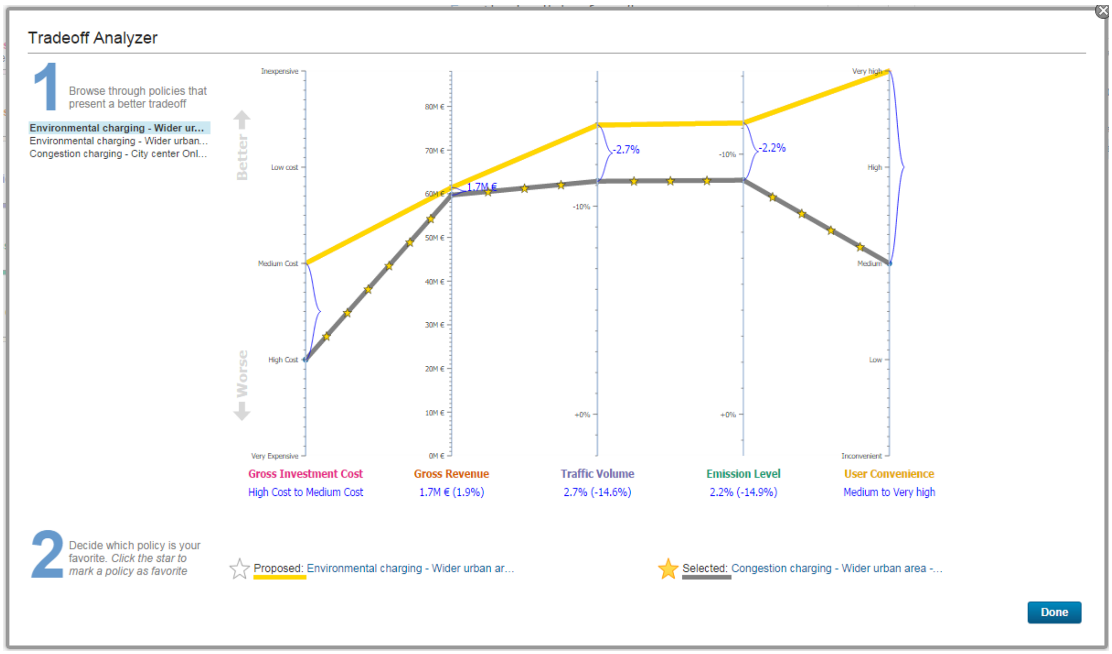


Figure 87: MOOViz tool in transportation scenario – tradeoff analyzer

The right hand side of the window contains a view similar to the lines view. In this view two policies are presented: the policy clicked by the user, and one of the suggested policies that may be more appealing. For each objective, the difference in percentage of the objective value is plotted. Browsing between the suggested appealing policies is done by clicking on the policy names on the left hand side of the screen. This is considered as the first part of the tradeoff analysis. At the bottom of the window, the proposed policy and selected policy (from the appealing set) are listed with their names. It is possible to add them to the favorites list (using the star to the left of the name). Adding the policy to the favorites list would turn the star to yellow and mark the policy with stars.

When the tradeoff analysis is done (after going over the chosen appealing policies), the user clicks the done button at the bottom right side of the window and returns to the main application page.

Access to the tradeoff analysis can be done in a similar fashion from the polygon view. In this case, clicking the done button closes the window and returns to the lines view.

4.6.11 Zoom in on views

When the sliders are moved and some solutions are disabled, it is possible to zoom in. Zoom in allows analysis of chosen subset of policies. This subset is induced by changed in the range of various objectives (see section on range sliders for additional information). Zooming is done by clicking the button on the top left panel:

Figure 88: MOOViz tool in transportation scenario – zoom in button

The button is only enabled when there are disabled policies as a result of changing the sliders. Pressing the zoom will open a zoomed view, in which the slider's range is automatically adjusted to the new range of values.

This is a snapshot before zoom:

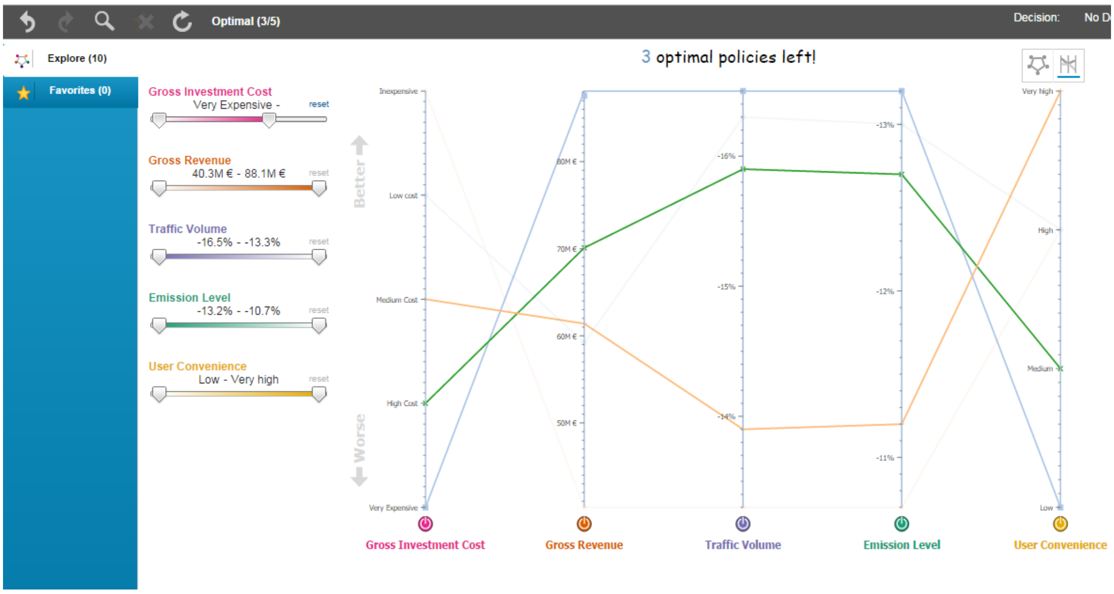


Figure 89: MOOViz tool in transportation scenario – snapshot before zoom in

The zoom button is now enabled. Clicking it performs the zoom in and the obtained snapshot is as follows:

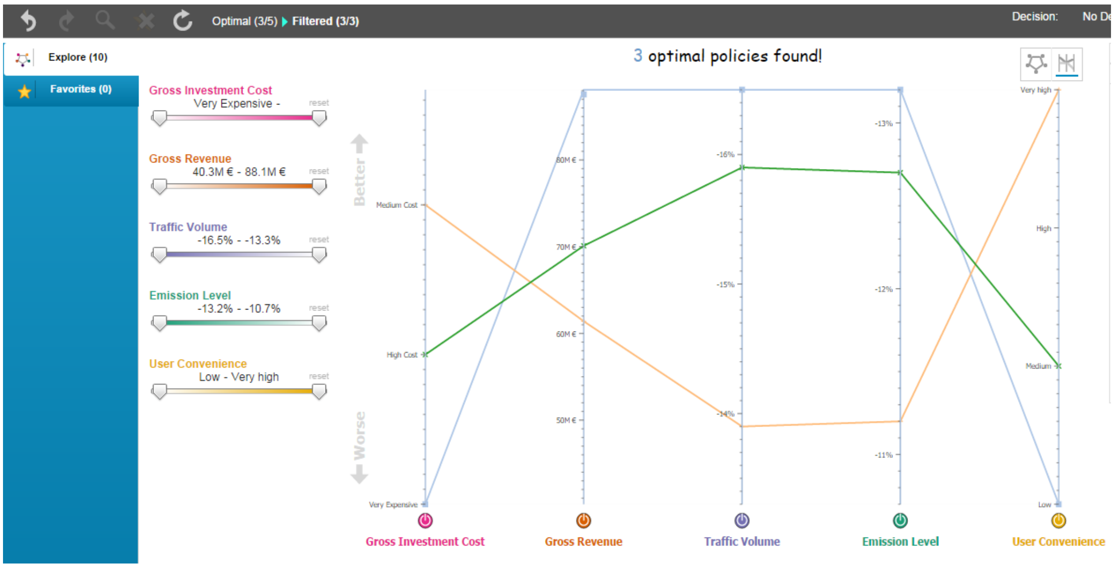


Figure 90: MOOViz tool in transportation scenario – snapshot after zoom in

There are two changes compared to the previous view. First, the range of the sliders is changed to reflect the change in value (see the slider corresponding to gross investment cost in the previous snapshot). Second, on the top panel next to the Optimal link, a link to the current view was added:

Figure 91: MOOViz tool in transportation scenario – link to zoom in view

In this view, the shown policies are the subset obtained by changing the range of values of the objectives. Clicking the Optimal link moves us back to the previous view. In addition, we can change the sliders in the filtered (zoomed) view and zoom in again. In this case the link chain would look as follows:

Figure 92: MOOViz tool in transportation scenario – link to nested zoom in

We can continue the zoom in the same fashion.

5 Consensus Game

5.1 Introduction

Policy making is a task affecting many aspects of a community, from the natural environment and the artificial constructs all the way to the people and the animals that live and take part in the frame of this community. That is the reason that a policy maker must always consider the effects that a policy will have on the community, on some occasions a negative effect, caused by a poorly chosen policy can have consequences that harm the community for a large time frame. In order to predict that factor a crowdsourcing game is implemented as to educate citizens in policy making and collect citizens' preference.

Policy implementations rarely manage to tackle all objectives simultaneously and often a prioritization needs to take place before the policy maker decides to implement a measure. This prioritization must be made clear to the public because citizens' involvement in decision making and the transparency of the process would also support better acceptance of any decision. One of the reasons that policy implementations often fail is because their respective objectives and their tradeoffs, are not adequately explained to the public. In order to educate citizens about policies and their consequences it is imperative to present to them all the available options of policy implementations and link them to the potential outcome.

To tackle this issue, we propose a web tool and underlying model that enables citizens to collaboratively examine the objective space and link it to consequences. For incentivizing participation and in order for this tool to maintain an educational character, gamification techniques are used, rewarding decisions that are near-optimal. At the same time, it is expected that this process will enforce a transparent and more straightforward approach from the policy makers who will then have to justify their decisions. This Section (Section 5) is comprised by two major units: the web game, in which the game tactics and the tools used to implement it are described and the reward model. The latter, even though an integral part of the game (developed by ATC, co-designed by NTUA and ATC), it is actually developed in isolation (by NTUA) and the integration took place through loosely coupled interfaces (web services).

5.2 Research Challenges

The main foreseen challenge will be the incentivization of citizen participation, i.e. making the system attractive for the citizens to use it so as to achieve educating and harvesting of user preferences including achieving a critical mass of user preferences that will make them a useful and reliable source of information. This challenge will be addressed by infusing gamification concepts in the main platform, thus leveraging on the people's natural desires for competition, achievement, status, self-expression, altruism, and closure.

Another challenge is creating a mathematical model for the award system in order to assign points for citizen's decisions during the game session. In order to tackle this challenge we divide policy making in two main categories; objective and subjective policy making. Policies to assist in objective decision making are usually operational in nature and can be objectively tested .On the contrary policies to assist in subjective decision making would usually consist of a number of factors that citizens consider subjectively and are unknown.

The objective criteria of the reward model are generated by the deduction of policy-making into a multi-objective optimization problem (MOP). In MOPs there is no single optimal solution but rather a set of alternative solutions. These solutions are optimal in the wider sense that no other solutions in the search space are superior to (dominate) them when all objectives are simultaneously considered. These solutions are called pareto-optimal solutions. Pareto optimality provides flexibility for the human decision maker. Ultimately, a designer or decision maker needs to make tradeoffs between disparate and conflicting design objectives (i.e. design metrics). So far, the prevailing approach for addressing this decision-making task is to solve an optimization problem which yields candidate solutions. To acquire a single optimum solution, a statement of user preference is required. In the mathematical model for the award model, policy implementations are assigned score points according to their optimality in the solution pool.

Another challenge is to create citizens' hierarchy of objectives in order to generate a more realistic system with subjective decision making. To tackle this challenge policy implementations selected in the game will be marked with a number indicating how many times they were selected, this numbering will provide a more precise image of user preference. Policy implementations will be evaluated by parameter and areas of most selected values in each parameter will indicate users' preference for this specific parameter provided their appearance count surpasses a certain threshold. Policy implementations will also be assigned score points that represent consensus with the public opinion.

As final challenge stands analyzing the results of the game, assigning correctly most preferred policy on each area with policies implemented in game sessions. In order to avert this challenge statistical data should be saved from all final policies implemented in game sessions by players. According to the majority of specific policies combined with players' experience level choosing these policies, a more precise image of citizens preferred implemented policies will be generated.

In what follows we provide a description of the web game (Section 5.3) and in sequel the corresponding description of the Reward model (Section 5.4). The combination of these two is expected to resolve the major challenges mentioned above.

5.3 Consensus Game

The game is designed and implemented in such a way that hides the complexity of the decision making process and the underlying policies and instead challenges the users to find good solutions based on their experience and review solutions of others. The environment is entirely user friendly and metaphors and assistants are created in order to enhance user experience and relay basic messages regarding the concepts of the policies in question. Users are rewarded based on their selections, something which aims to engage them in a competitive environment in which they will want to return. The game scenario mentioned below gives a hint of the process that the user has to traverse while playing the game, learning and providing indications about preferences.

5.3.1 Game Scenario

The following list describes the actions taken by a user in order to play a full game.

1. **Login/Register**. Users can create an account in the game or login using an existing account of other social media like Google, Facebook or Twitter.

2. **Select the field of interest.** Users can select one of the two fields (Transportation & Biofuels) that is of their interest to play.

3. **Click a game** to open the game page.

4. **Start the game**. When the player clicks start, the timer of the game starts the countdown.

5. **Order the objectives.** The players sees the problem description and sets the order of the objectives from highest to lowest priority.

6. **Submit the objectives' order.** The players submits the selected order of the objectives.

7. **Select a proposed solution.** After submitting the order of the objectives, the game proposes three different solutions and the player decides which of the proposed solutions represents his opinion.

8. **Finish the game & View the results.** After submitting his choice, the player can finish the game by getting the game's response which indicates if his chosen solution was an optimal solution or not. Also the user gets a score according to the reward model described later in Section 5.4.

5.3.2 Implementation Details

The Consensus Game is implemented with ASP.NET[8] and SignalR[9] technologies. The architecture of the technical solution is depicted in the following Figure 93:

[8]http://www.asp.net/

[9]http://signalr.net/

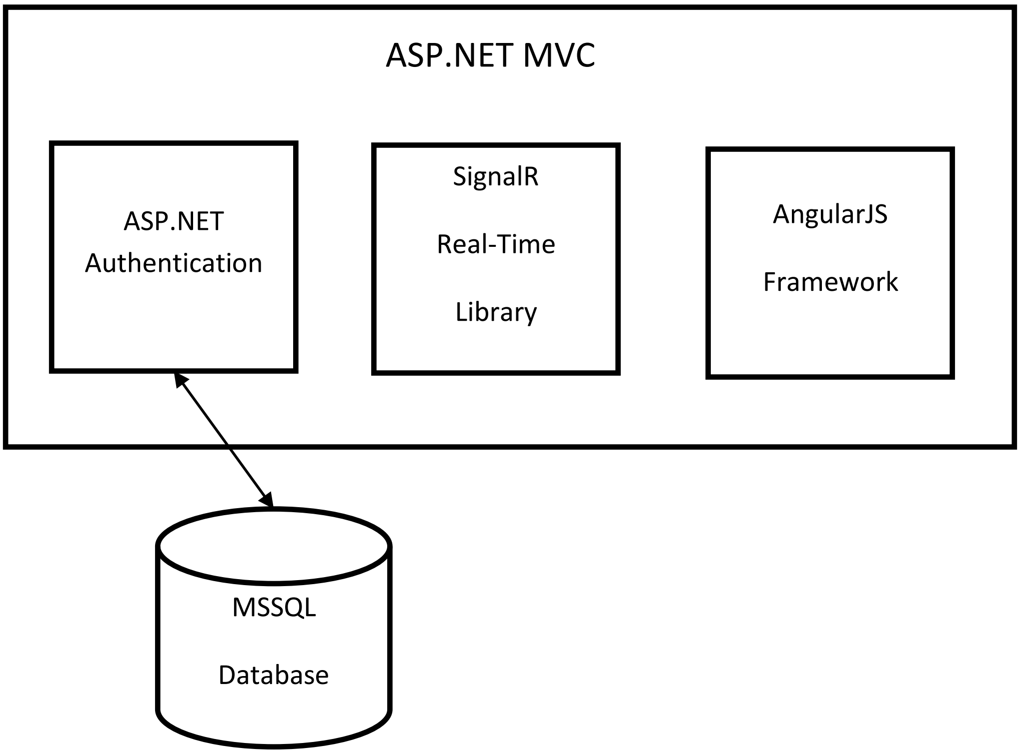


Figure 93: Consensus Game Architecture

The Consensus Game is an ASP.NET MVC web site developed with Visual Studio 2013. The most recent version (ASP.NET MVC 5, Visual Studio 2013) of the technologies and tools are used. ASP.NET MVC makes it very easy to add login from external services (Google, Facebook, Twitter, Microsoft Account). This can be achieved with a single line of code and only requires to add the authentication identifiers of the external service. This is highly desirable for the Consensus Game, as the use of external services for authentication will greatly facilitate the adoption of the game.

The persistence layer of the game is created with a MSSQL database. This database is mainly used for supporting the ASP.NET authentication scheme and storing user information. Among others, the achievements of a player are stored in this database. The database is currently deployed in a local Windows Server machine. However, it can also be deployed in Microsoft Azure cloud environment, without code changes.

Although ASP.NET MVC provides great features for implementing modern and robust UIs, these are only used at a bare minimum. Instead, the UI is created with the aid of AngularJS[10] framework. The game is basically a Single Page Application. This makes the user experience more fluid and also the integration of real-time characteristics easier.

The core of the game is the ASP.NET SignalR real time library. SignalR allows bi-directional communication between server and client. Servers can push content to connected clients instantly, as it becomes available. SignalR supports Web Sockets, and falls back to other compatible techniques for older browsers. SignalR includes APIs for connection

[10] https://angularjs.org/

management (for instance, connect and disconnect events), grouping connections, and authorization. These functionalities allow the players to see their co-players decisions almost instantly, without the need of long polling. The players are also notified instantly, when another player joins or leaves the game.

5.3.2.1 Installation Guide

The software required to setup the first prototype of the game is presented below:

- ASP.NET MVC framework
- ASP.NET SignalR library
- MSSQL Server
- Angular JS framework
- Operating System: Windows 7
- Web Server: IIS 7.5

The latest versions of these frameworks and libraries were used during the development of the game and they will be updated to the new ones, if available, throughout the development period. The standard configuration options have been used for setting up the various frameworks and libraries. The final deployed game is a website which can be accessed by all major modern browsers with no other software or hardware requirements.

5.3.3 Prototype Access Details

The first prototype of the Consensus game was developed based on the technologies described in the previous paragraph and is currently deployed and publicly available at the following url: http://atc-dnn.atc.gr/consensus/ by using the below credentials:

Username: demo
Password: 123456

The prototype includes the basic game features and logic and will be continuously developed according to the needs and feedback received from the users' testing and evaluation.

Minimum client requirements:

The deployed game is a web application which can be accessed by all major modern browsers with no other software or hardware requirements.

5.4 Reward Model

As Zichermann explained gamification is able to stimulate levels of user engagement because "in the regular world, the dopamine rush doesn't happen that often. But in the gamification world, we engineer challenge and achievement to happen continuously to drive engagement forward". Rewards make certain a user understands when targets and achievements get fulfilled. In order to acquire data indicating user preference in specific policy implementations the need to engage citizens in the crowdsourcing game is obvious.

Citizen engagement will be achieved in two steps; generating a reward model to indicate target fulfillment and adding gamification elements to create enjoyable experiences.

In the concept of creating a game we implement a Bayesian game in the sense that it has a set of players, in our case it is a multiplayer game with four players. It has a set of actions *A* available to each player *i*, where players choose from a set of available policy implementations. *T* is the type space of the player; each player favors one objective more than the rest therefore her actions are in consideration to maximizing that objective, indicating a specified spectrum of choices he can make in the total population of options. The common prior among types $p: T \rightarrow [0,1]$is generated by the probability distribution across each player's type space in the total population of options. And finally u is the utility function we generated by marking all solutions of the population with pareto-optimality criteria and assigning profits to them.

The reward model (utility function) will be structured by a mathematical model indicating objectively and subjectively optimal solutions as follows.

5.4.1 Mathematical Model

5.4.1.1 Notation of Solutions with Utility Functions

In the mathematical model all solutions are sorted by Euclidean distance and all n-vectors descending from the highest distance to the lowest are compared with each other in order to establish each solutions domination count. Then we ranked the pareto–optimal solutions depending on their domination count and created k-means clusters from the pool of solutions (with k the number of parameters each vector had). Non-dominated solutions were assigned profit of 3.0 provided they didn't exceed 19% of the total population (in order to prevent the algorithm from prematurely converging) marking solutions from each cluster in order to diversify the pool of optimal solutions. If they exceed the 19% of the total population of solutions they were assigned with score points from the next score point category. Later we define the next rank of optimal solutions as the next 5% of the total population. Again solutions are ordered by the number of solutions they are dominated by. The rest of the data are divided in categories of 5% and assigned with profits descending by 0.5 points. The complexity of the algorithm is $O(mN^2)$.

5.4.1.2 Preference Induction

In order to help citizens learn how public opinion is crucial in their selections we collect data of user's selections in each game session. For each objective the policy-making problem suggests (i^{th} –vector parameter) we find the margin of values that indicate user's preference with a threshold above 45% of the total selections made. If the threshold is below 45% then no suggestion margin is set for this objective therefore the mathematical model does not verify consensus with the users' preference for this parameter.

For each game session users participate in, they will be rewarded with score points according to their game behavior and choices according to the utility function:

$$U = Pareto(f) * 100 + Consensus + Agreement$$

where:

$Pareto(f)$is the function used to establish pareto-optimal solutions and assign profits to them (from 3.0-0.0)

$$Agreement = \begin{cases} 0\ ,0\ user\ select\ users\ implementation\ policy \\ 25\ ,1\ user\ selects\ users\ implementation\ policy \\ 50\ ,2\ users\ select\ users\ implementation\ policy \end{cases}$$

In this manner players in order to get the extra bonus will be prompted to debate about their options and analyze them therefore better understand them and on the other hand we will get more valid preference indicators. The agreement parameter indicates that in a game where the best implementation policy is voted from many users it is also the one to be implemented.

$$Consensus = 25 * percentage\ of\ consensus$$

Where the percentage of consensus will be estimated by each vector parameter e.g. if in parameter i of the *N*-vector the area of preference is between$[value_1, value_2]$ by 47% then if players implementation in parameter *i* is between $[value_1, value_2]$it will receive

$$\sum_{i=0}^{m} = \begin{cases} 0, & if\ not\ in\ the\ space \\ 1/m, & if\ in\ the\ space \end{cases} * 100\%$$

If all parameters match the assigned margins the user will increase the percentage of consensus with the public opinion.

5.4.2 Other Forms of Reward Models

Gamification techniques strive to leverage people's natural desires for competition, achievement, status, self-expression, altruism, and closure while simultaneously specify targets and create enjoyable experiences.

In order to accomplish playfulness feedback messages will be implemented mostly to provide positive feedback in response to successful actions during the game session. Plot animations and pictures will be added to help better understand the problem the policy implementations are solving and also as a means of creating a sense of fun since they are visually attractive. Unlocking mechanisms will be used in specific game sessions in order for specific games to be played by users with better understanding of the policy making process (which will assist in acquiring more efficient data of the user preference when people are aware of the policy making process and tradeoffs). Achievements in the form of badges will be used to mark the completion of goals and the steady progress of play within the system. Users enjoy the sudden rush of surprise or pleasure when an unexpected badge shows up in a gamified system. At the same time a well-designed, visually valuable badge can also be compelling for purely aesthetic reasons.[116]

5.5 Game Evaluation

The crowdsourcing game will be evaluated according to two parameters; the learning curve and the participation. In order to evaluate the performance of the game after being deployed, there are specific metrics we can use.

5.5.1 Learning Curve

The Learning Curve (LC) forecasts how fast future costs will drop as more of an item is produced. The LC contributes a simple mathematical relationship between some metric (performance measure) such as cost, quality, or cycle time of producing an item. In our case Learning Curve forecasts how fast future decisions will be optimal as more games are played. Using the power law of learning we have:

$$M(q) = M(1)q^{-p}$$

Where M(1) is the amount of solutions tested in order to acquire an optimal solution. And M(q) is the amount of solutions tested so far in order to acquire the q^{th} optimal solution, p stands for the parameter of learning for each candidate.

Using the exponential law of learning we have:

$$M(q) = M(0)e^{-pq}$$

Where M(0) is the amount of solutions tested before we acquire the first optimal solution.[118]

In simpler terms we need to collect data from user selections mid game sessions and see if the later choices they make are more optimal than the previous ones. Also in later game sessions if they also make choices that are more optimal than the previous made in other game sessions and other game levels. In this manner we test whether users understand the tradeoffs each policy implementation suggests and which solution from the pool of policy implementations provided is optimal for each scenario.

5.5.2 Participation

In order to measure participation, as in all games with a purpose (GWAP) we have specific metrics. We count Throughput (the average number of problems instances solved per human per hour), ALP (average(across all people who play the game) overall amount of time the game will be played by an individual player) and Expected contribution (throughput*ALP).[114]

5.5.3 Implementation details

The reward model is invoked as a SOAP web service. There are various interfaces in order to accomplish access and modification of the database in order to accomplish elements of the game sessions such as choosing a policy implementation, presenting policies according to preference order etc. Also interfaces for generating statistical data of users selections and making further study of the data collected in the game, are invoked. The list of interfaces includes:

1. decreaseChosen (decreases the number a policy implementation has been selected –usefull when users play the game session again and choose a different policy)
2. decreaseLiked(decreases the number a policy implementation is liked by a player)
3. findAll(returns all policy implementation data)
4. findByBiodiv(finds policies according to the biodiversity percentage completeness)

5. findByChosen(finds policies according to the number of times they have been selected)
6. findByCo2(finds policies according to the Co2 emission reduction percentage completeness)
7. findByCostfood(finds policies according to the Cost of food reduction percentage completeness)
8. findByDistance(finds policies according to the Euclidean distance the policy implementation vector has)
9. findByForestland(finds policies according to the forest land percentage completeness)
10. findById(finds policies according to their assigned ID)
11. findByLiked(finds policies according to how many times they are liked by users)
12. findByMyorder(finds policies according to the preference order)
13. findByPolicy(finds policies according to the policy implementation name)
14. findByScore(finds policies according to score they are assigned with)
15. getPolicyID (return a policy's ID number according to its name)
16. getScorePreset(returns score points assigned only from the process of finding the pareto optimal solutions)
17. increaseChosen(increases the number of times a policy implementation is chosen)
18. increaseLiked(increases the number of times a policy implementation is liked)
19. scorePreference(returns all score points assigned to policy implementations including the scores from the public preference)
20. scorePreferenceByOrder(returns the total score with public preference according to preference order)
21. setPreference(finds public preference and returns a timeline of public preference in in ranges)
22. setPreferenceByOrder(finds public preference by order of preference and returns a timeline of public preference. With data a) the most selected preference order b) most selected preference order by category [1st,2nd etc.])

5.5.4 Prototype access details

Access to the prototype is possible via a SOAP Client through the following WSDLs:

- http://147.102.19.139:8080/PreferenceDeductionService/PreferenceDeduction?wsdl
- http://147.102.19.139:8080/webMethodsService/webMethods?wsdl

These services are invoked from the Consensus Game platform, currently in order to retrieve the reward for a scenario selection from a player.

6 Visual Analytics

6.1 Introduction

As stated in Chapter 2.5 Visual Analytics combines the research area of visualization and data mining keeping the human in the loop. Therefore, we focused on three different branches of research.

1. **Visual support** of similarity judgments to help the analyst/user comparing sets of alternative policy scenarios, described by multivariate measurements each.
2. New **interaction possibilities** to support domain experts exploring the data space and comparing alternative scenarios.
3. New **automatic algorithms** to be used as a pre-calculation for improving the understanding of alternative scenarios by automatically performing data point comparisons.

Since the visualization is the interface for the human to communicate with the underlying algorithms and to understand the automatic generated results we aim at finding the most appropriate visual representation for the data. Therefore, we conducted a quantitative user study to measure the performance of different variations of glyph designs (*Chapter 6.1*). The results will later be used in our prototype by encoding data in the most effective and efficient way. This perceptional research is crucial for the further development of the visual analytics prototype.

Besides the visual output and possible interaction techniques offering strong automatic algorithms is a substantial part of our prototype. Therefore, we introduce a new technique to support the user in comparing multi-dimensional data points (Chapter 6.2).

We combine the different research areas in our **visual analytics prototype**. Besides state-of-the-art representations and algorithms we illustrate how our research is used to further improve common visualizations (*Chapter 6.3*).

Parts of this work have already been published:

J. Fuchs, P. Isenberg, A. Bezerianos, F. Fischer, E. Bertini; *„The Influence of Contour on Similarity Perception of Star Glyphs“.* IEEE Transactions on Visualization and Computer Graphics, IEEE Computer Society, 2014.

6.2 Visual Support for Similarity Perception

Since the user has to interpret the visual cues on the screen we have to find visual elements, which are most effective and efficient in communicating the underlying data and supporting the user in her task. In order to build an appropriate visual analytics prototype finding the best visual representation of data elements is crucial. Therefore, we conducted three quantitative user studies focusing on a similarity perception task with different variations of star glyphs.

Data glyphs are small composite visual representations of multi-dimensional data points. Glyphs express the dimensions of a data point by assigning them to a specific visual variable [134]. Given their small graphical footprint glyphs are very versatile, used in a variety of different application areas: Monitoring computer networks [135],[136], tracking the health of patients [137], comparing country characteristics[138], or analyzing sports games[139]. Glyphs, in contrast to general charts or other visualizations, are often used as small visual representations nested inside other visualizations such as hierarchies, networks, or geographic data---or when a very large number of data points needs to be seen in one overview. Their primarily role is typically to provide quick overviews and help detect data trends and similarities[134].

A star glyph[140] is a specific type of glyph that lays out the axes for each data dimension on a radial grid and matching the dimension's values to a position on the respective axes, typically connected with a line to the center of the glyph. There exists a great variety of alternative designs for star glyphs that differ in the amount of reference structures used, the use of additional visual variables on the ``rays,'' or whether or not the individual rays are connected to form a contour for the glyph[141]. The version of the star glyph with unconnected rays is also sometimes called *whisker or fan plot*, while the connected version also carries the name *star plot*[134]. Star glyphs are frequently used but very little advice exists on how to choose between different star glyph encodings. The question arises to what degree changes in the design of a star glyph influence its perception and, thus, the effectiveness of the glyph in certain tasks.

One important task for glyphs in small-multiple settings is the comparison of the encoded data points to one-another. Such a comparison task may be conducted to find data points that are very close over all dimensions, very different, or similar in just a subset of dimensions. We focus on the first task: finding data points encoded as star glyphs that are very similar to a target glyph. We are interested in this task because if it is well supported, it should improve people's ability to perform the other two types of comparison tasks. We hypothesized that the ability to perceive a star glyph as a coherent and closed shape would strongly influence the correctness of data similarity detection tasks---as it would potentially be easier to compare a single shape than having to compare individual rays. This hypothesis was motivated by prior research showing that a closed contour has an influence on the perception of a coherent shape[142]. As Palmer noted:

"Shape allows a perceiver to predict more facts about an object than any other property"[11]

There are many real-world scenarios where multi-dimensional glyphs can provide valuable information. Multi-dimensional data is notoriously hard to represent visually as the number of visual variables available to encode data dimensions is limited. Multi-dimensional glyphs, and more specifically glyphs where data dimensions are presented through radial axes,

[11] [5024], page 363.

provide ``hints'' of the underlying multi-dimensional structure when multi-dimensional objects are plotted on the spatial substrate. Common examples are: maps showing the geographical distribution of multi-dimensional objects (e.g., comparison of indicators such as crime rate or suicides for different regions of France[143]), multi-dimensional scaling visualizations exposing relationships between scaling algorithms and data distributions (e.g., election patterns to show political party proportions by region[144]), or data objects organized in a grid layout to show how multi-dimensional objects distribute across sets of predefined categories (e.g., food nutrients in different food categories).

6.2.1 Experiment 1: Contours for Novices vs. Experts

In our first study we were interested in the fundamental question: does contour affect people's perception of data similarity with star glyphs? Data similarity judgments are cognitive tasks, where the viewer has to judge the absolute difference in all dimension data values between two data points. This differs from other types of similarity judgments, such as detecting shape similarity e.g., under rotation or scale.

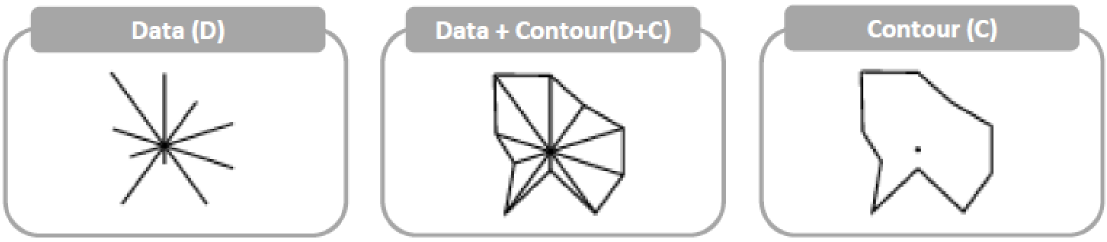


Figure 94: Experiment 1 Contour Variations: (from left to right) star glyph with rays and no contour (D); common star glyph (D+C); only the contour line of the star glyph (C) [145].

Detection of data similarity is a synoptic task according to the Andrienko & Andrienko [145] task taxonomy. Synoptic tasks are very common and important for glyphs in small-multiple settings. Analysts have to visually compare data points to detect outliers or to identify similar groups of data points, by referring to the whole data set or a subset of the data (e.g., finding countries with similar characteristics).

We were interested in the effect of contour, as we hypothesized---based on previous perception studies[142]---that a contour would impact the rapid perception of shapes and, thus, aid in tasks that require the data point to be perceived in its entirety. Finally, we hypothesized that there would be a difference between experts' and novices' ability to make accurate data similarity judgments, and thus chose to conduct a between-subjects experiment with these two groups of participants.

6.2.1.1 Design and Procedure

Glyphs: We used three different variations of the star glyph (Figure 94). The first, also called whiskers or fan plot[146],[134] uses "rays" to encode quantitative values for each dimension through the length of each ray. We refer to this variation as "Data lines only (D)". The second variation, "Data lines + Contour (D+C)", connects the end of each ray with a line to add a closed contour [140]. In the third variation the radial rays are removed, and only the contour line is presented [147]. We use the term "Contour only (C)" for this design variant. All three star glyph contour variations have been used in real-world context and in the scientific literature, thus adding external validity of our glyph choice.

Dimensionality: To investigate the effect of contours on different data densities we varied the number of dimensions shown in the glyphs. The *low* dimension density consisted of three data dimensions with corresponding data values, while the *high* density consisted of ten data dimensions. We considered ten dimensions to be *high*, as glyphs used in the literature rarely visualize more than ten dimensions; also to our knowledge there is no study investigating the maximum number of perceivable dimensions in a single star glyph to use as a basis.

Task, Procedure and Apparatus: Participants were shown a highlighted stimulus glyph surrounded by 8 more glyphs in a 3 * 3 matrix configuration (Figure 95). One of these glyphs was closest in data space (lowest absolute data distance) while the rest were distracters further away in data space. The participant had to select the glyph closest to the stimulus in terms of data value. For each contour variation, participants were given training explaining how data was encoded and the notion of similarity in data space. They were then given four practice trials where the correct answer was revealed to help learning. During the actual experiment the correct answer was no longer provided.

The three glyph variations were presented in an order randomized using a latin square. The position of the correct answer as well as the different distracters was also randomized. Similarly, the exact glyph values were randomized. Each participant repeated 4 training and 4 real trials for each contour variation. The study took place in a lab setting in the presence of an experimenter. The experiment was conducted on a 24 inch screen with a resolution of 1920 * 1200 and took around 25 minutes. The only input device was a common computer mouse to make selections.

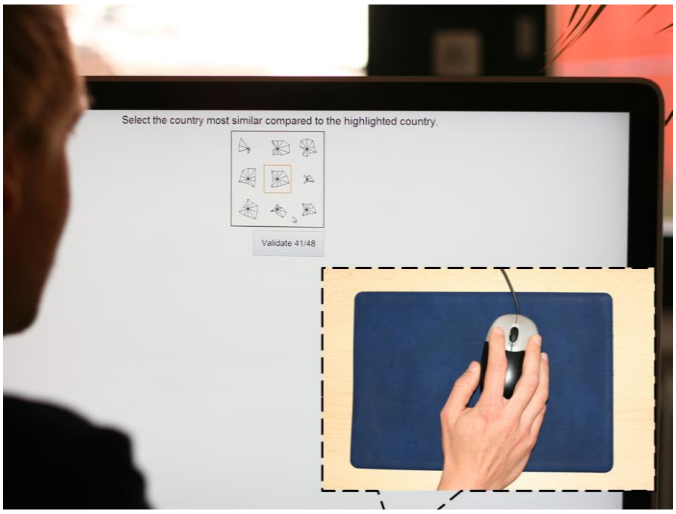


Figure 95: Experiment Setting: The participant was seated in front of a 24” screen with a resolution of 1920x1200. The only input device was a computer mouse.

Participants: Twelve novices (7 female) and twelve experts (2 female) participated in our study. The age of novice participants ranged from 18--23 years (mean & median age 20), and from 26--38 years (mean 30.3 and median 29) for experts. All participants reported normal or corrected-to-normal vision. All novice participants reported no experience in reading glyphs, but were familiar with common chart visualizations seen in print (e.g., bar and pie charts). All 12 experts were visualization researchers and students who reported a strong background in data visualization with at least basic knowledge of reading glyphs (1 Bachelor; 8 Master; 3 PhD).

6.2.1.2 Hypothesis

1. Novices are less accurate in judging data similarity than experts
2. Both experts and novices make more accurate judgments in the low dimensional than the high dimensional condition
3. For both experts and novices, contour variations (D+C, C) improve the accuracy of data similarity judgments
4. This effect will be stronger in novices who have no prior glyph reading experience.
5. Contour variations (D+C, C) lead to more accurate judgments mostly in the high dimensional condition, while the low dimensional condition is be less affected overall

6.2.1.3 Data Generation and Distracters

Our data was synthetically created: 3 dimension values for the low and 10 for the high dimensional case. For each dimension we consider data values ranging from 0 to 5, partitioned in three value categories: low [0, 1], middle [2, 3], high [4, 5]. We avoided larger value ranges as we were not interested in studying visual acuity. The stimulus (i.e., central highlighted glyph) was created randomly by assigning either a middle or a high data value to the different dimensions with an equal chance of 25\% (i.e., 50\% for each value categories and 50\% for the final data value). This was done once for all repetitions. To avoid learning effects, the stimulus was rotated between repetitions, keeping the values and the neighboring dimensions identical.

Each trial also contained a *target* glyph---the correct answer, thus the most similar to the stimulus in terms of data closeness (minimum data value distance). To generate it, we changed the data values of the stimulus randomly up to a maximum of 7 changes in data distance for the high dimensional condition, and 1 for the low. This was done by sequentially scanning the dimensions with a probabilistic function, which first decided to change the dimension or not (50%), second to increase or decrease the corresponding data value (50%) and third by how much (i.e., 1 or 2)(50%). At the end we ensured that the resulting data values fit into one of the three categories (i.e., low, middle, and high) and that the sum of all changes meet the predefined criteria.

Besides the stimulus and target glyph, we created 3 types of *distracters*. First, a rotated version of the stimulus, keeping the data values identical, but switching the dimensions one step either to the left or to the right. Second, a scaled version of the stimulus where we reduced the data values of each dimension by 1. Since the data values of the stimulus reach from 2 to 5 it is not possible to end up with negative values. Third, a close alternative of the target glyph. This alternative takes the data values from the stimulus and changes the values randomly up to a maximum of 8 changes in data distance for the high dimensional case, or 3 for the low. Values were chosen to ensure that the alternative glyph is not too different from the stimulus, while the target glyph continues to be the most similar in data distance. The remaining distracters were created randomly by assigning a data value to each dimension with an equal chance (Figure 96). For each trial we ensured that the sum of all differences between stimulus and all distracters was higher to that between stimulus and target glyph.

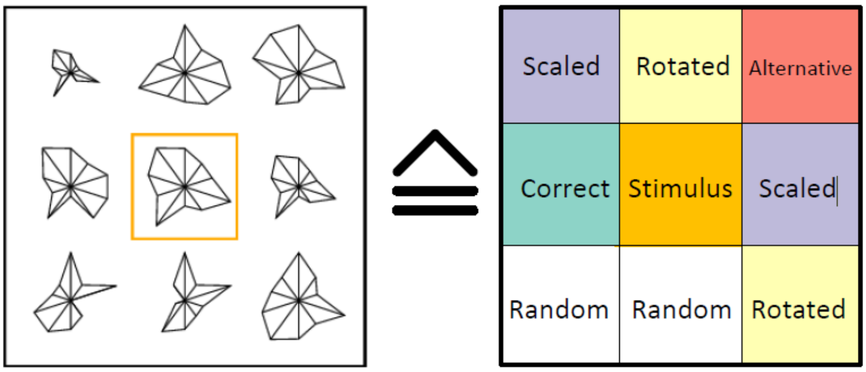


Figure 96: Experiment Setting: For each trial glyphs were arranged in a 3x3 matrix. The stimulus is highlighted and positioned in the middle to assure an equal distance to the other glyphs. This setting is used in all three experiments.

6.2.1.4 Results

We report only statistically significant results ($p < .05$) for accuracy. Given the non-normal nature of our data we used a non-parametric Friedman's test for the analysis of correct answers between glyph variations and a Kruskal-Wallis test for comparisons between expertise (between group factor). Figure 97 shows overall correct answers, and Figure 98 which type of distracters participants chose under the different experimental conditions. Although completion time was logged, we found no differences across variations and user groups, with low dimension trials taking on average 11sec (D=12.7sec, D+C=11.3sec, C=9,7sec) and high ones 18sec (D=19.7sec, D+C=16.9sec, C=16.7sec).

Overall accuracy for experts across variations was 79.1% for the low and 44.4% for the high dimensional glyphs, and for novices 74.3% and 36.8% respectively. However, there was no significant effect of expertise on accuracy. Figure 97 illustrates more high level results. **Dimensionality**: There was a significant effect of dimensionality on accuracy ($\chi2$ (1;N = 288) = 23; p < :001). Post-hoc tests revealed that participants were more accurate in the low dimensional condition (76:7%) compared to the high dimensional condition (40.6%, p < .001).

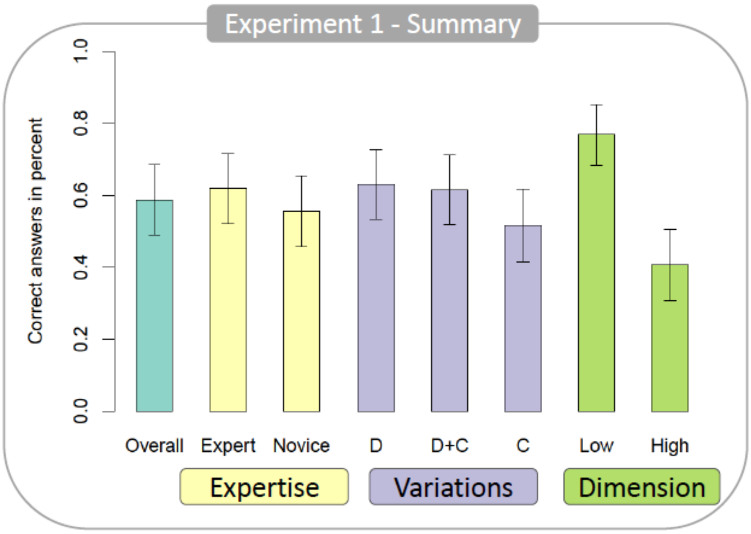


Figure 97: Experiment 1 Summary: The bar charts illustrated the percentage of correct answers and the standard deviation.

Contour variation: There was a significant effect of contour variation on accuracy ($\chi2(2;N = 192) = 7:9; p < :05$). Participants using variation C performed significantly worse (51.6%) compared to D (63%, $p < .05$) and D+C (61.5%, $p < .05$). For experts, there was a significant effect of contour variation on accuracy in the high dimensional condition ($\chi2(2;N = 48) = 12; p < :001$). A pairwise comparison revealed a significant higher accuracy with the D variation (66.7%) compared to both D+C (41.7%, $p < .05$) and C (25%, $p < .001$). No significant results were found for novice participants.

When comparing the accuracy of the two participant groups, we found that for the variation D, there was a significant effect of expertise on accuracy in the high dimensional condition ($\chi2(1;N = 96) = 5.85; p < .05$). Experts performed significantly better (66.7%) using the D variation compared to novice participants (39.6%, $p < .05$).

When selecting a wrong answer, both experts and novices most frequently selected the second closest data point to the stimulus (17.7%, 20.5% respectively), followed by a scaled version of the stimulus (16%, 16.3%) and to a lesser extent rotated versions (2.4%, 4.1%), mostly in the high dimension case of the contour variations (C+D;C).

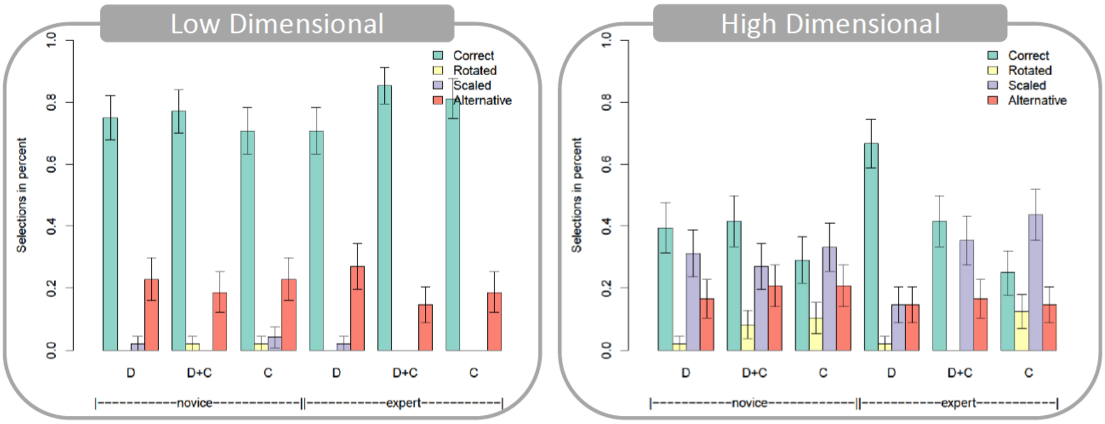


Figure 98: Experiment 1 Results: The bar charts illustrate the percentage of selections and the standard deviation for each factor. In the high dimensional condition experts using variations D+C, C were lead to judge shape similarity rather than data similarity whereas the accuracy of novices was low for all three variations.

6.2.1.5 Summary and Discussion

Overall we cannot confirm H1, our experts were not significantly more correct than novices on average. This is especially true for the low dimensional condition where both user groups had a good performance (80% correct). However, for higher dimensionalities experts using variation D were significantly more accurate compared to novices (partially confirming H1). When comparing the two dimensionalities, similarity judgments were significantly more accurate for both user groups in the low dimensional condition compared to higher dimensionalities, confirming H2. With an increasing number of dimensions more data values have to be visually compared, leading to more complex mental calculations resulting in a higher error rate.

Contrary to intuition from previous work that contour can improve similarity judgments[142],[148], we found that contour affected the accuracy of judgments negatively for experts. Thus we cannot confirm H3. As no significant effects were found for novice participants, we could also not confirm H4, however, mean accuracy for C (50%) was lower compared to D+C (59.4%) and D (57.3%). We also could not confirm H5. Contrary to expectations, the variation without a contour (D) led to significantly more correct answers for high-dimensional glyphs. The effect was not visible in the low dimensionality case where all participants were overall approx. 80% accurate with all variations. Trying to explain the unexpected negative effect of contour on experts, especially in high dimensional cases, we noted that at least half of the erroneous answers in the contour variations (C +D; C) were in the form of scaled versions of the stimulus glyph, and to a lesser extent rotated versions, i.e., glyphs that have a geometric form similar to the stimulus glyph. In retrospect, this negative effect of contour could be explained by the fact that contour, and closure in general, is one of the factors promoting the notion of unity according to Gestalt psychology[149]. In our case contours led our experts to erroneously consider glyphs as coherent shapes when judging similarity, rather than data points. This resulted in judgments and comparison of geometrical shapes rather than data, with experts being led to consider as more similar data points that were either scaled or rotated versions of the stimulus, rather than the one closest in data space.

Given the overall poor performance of novices in the high dimensional case we conjecture that due to their lack of familiarity and experience they tended to fall back to judging shape rather than data similarity for all star glyph variations. This is evidenced by the fact that at least half of their errors were a combination of scaled and rotated versions of the stimulus glyph.

6.2.2 Experiment 2: Perception of Similarity

Results from Experiment 1 indicated that in high dimensional cases contours mislead even experts to perceive rotated or scaled versions of the stimulus as more similar, rather than the one closest in data space. Based on this finding, we conducted a second experiment to better understand what type of similarity star glyphs naturally support. To this end, participants were not given any training or explanation of what similarity means, and we did not inform them that the glyphs encoded multi-dimensional data. Their only instruction was to select the most similar glyph. Our goal in this experiment was to examine what viewers naturally perceive as similar in different star glyph variations, without being instructed on how to judge similarity. Based on our results we hoped to identify the star glyph variations, if any, that naturally promote data similarity rather than shape similarity and, therefore, are more suitable for data visualization.

6.2.2.1 Design and Procedure

Glyphs: The experiment tested the glyph variations from Experiment 1, as well as a filled version of the C and D+C glyph. We wanted to examine whether variations of glyphs that are filled reinforce more strongly the notion of a closed shape, due to the strong foreground/background contrast[149]. We conjectured that fill color may lead to more shape rather than data similarity choices. The experiment was a between-subjects design with fill type as the between-subjects factor. Thus, the D glyph was included in each group as the baseline. We had a total of 2 fill types (Fill, No-Fill) with 3 glyph variations each, as illustrated in Figure 99.

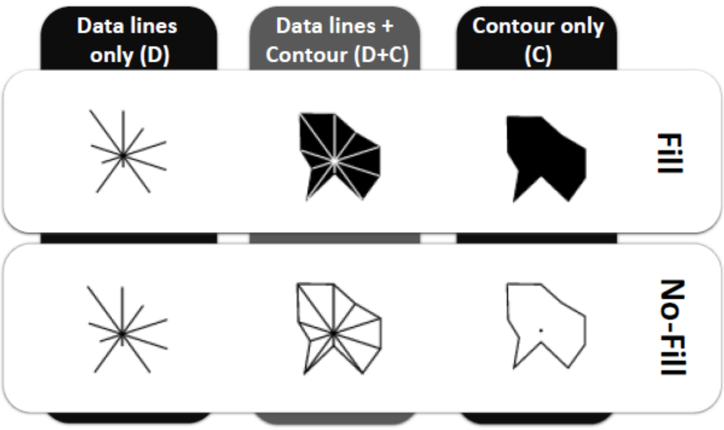


Figure 99: Experiment 2 design space: We enriched the design space from our previous study by adding a "fill" version of the star glyph. The design variations of the first study (i.e., D, D+C, C) are applied to both Fill and No-Fill.

Task: We again used a synoptic task, where participants selected the most similar glyph compared to a stimulus glyph. Participants were shown a highlighted stimulus surrounded by another 8 glyphs in a 3×3 matrix configuration. The positions of the surrounding glyphs were randomized around the stimulus. Again, we wanted to explore the notion of similarity and examine if some glyphs are naturally judged in a manner that approaches data rather than shape comparison. We thus gave no explanation as to what the glyphs represented and provided our participants with no training. Participants were free to interpret the word “similar” as they saw fit.

Data, Target Types and Dimensionality. Our data was generated as in Experiment 1, and again we tested low and high dimensionality. However, we included slightly different glyph choices to our participants, that we call "Target Types" (they are no longer distracters, as there is no correct answer). To balance the selection likelihood between each target type, we included two of each shape similarity and two glyphs that were closest to the stimulus in data space (we refer to this kind of target as “data"). As a result we had 2 data, 2 rotated and 2 scaled versions of the stimulus, and 2 randomly generated targets.

Participants and Procedure. Our study was conducted on Amazon Mechanical Turk (AMT), inspired by previous graphical perception experiments[150],[151]. We accepted 62 participants in total, and subjects were paid 0.50$ per Human Intelligence Task (HIT). Given the simple nature of our perceptual study, no qualification tests were required to complete our HITs. In accordance with AMT guidelines, however, only workers with 95% or more HIT approval rate were allowed to participate. Furthermore, we added control questions (3 in total) throughout the study, where one of the targets was identical to the stimulus and the answer was, therefore, obvious. We dismissed workers who did not get all the control questions correctly and their data was not included in the analysis. As a result we ended up with 36 participants (18 per fill type). Each participant worked on 4 trials for each variation and dimensionality, and viewed either the fill or the no-fill types. The order of presenting the glyph variations was randomized.

6.2.2.2 Hypothesis

Given the results from Experiment 1, and our conjecture on filling, we formulated the following hypothesis.

1. For the D variation, participants will choose data targets more often than rotated and scaled targets
2. Participants will choose data targets for the D variation of the glyph more often than they will for the other variations, irrespective of fill type
3. Participants will choose the scaled and rotated targets more often than the data targets for the C and D+C variations
4. For the filled D+C and C variations, data targets will be chosen less often than for the no-fill variations
5. In low dimensional conditions, data targets will be selected more often than other targets irrespective of glyph variation

6.2.2.3 *Results*

We only report statistically significant results (p<.05) for the collected quantitative data. We used a non-parametric Friedman's test for the analysis of the selections between the glyph variations (within-subjects) and a Kruskal-Wallis test for comparisons between glyph designs (between group factor). We did not log completion time, as we could not reliably control pauses during our online experiments.

There was a significant effect of target types on the selections made ($\chi2(2;N = 864) = 149$; $p < .001$). Overall, participants selected the data target type significantly more often (44.6%) compared to rotated targets (37.3%; $p < .01$), and scaled targets (17.8%; $p < .001$). For the D variation, included in both experiment groups (fill or nofill), data targets were selected more often (61.8%) compared to rotated targets (26.4%; $p < .001$) and scaled targets (11.8%; $p < .001$). For the fill designs (without the D variation), rotated targets were most commonly selected (38.3%), followed by data (35.5%) and scaled ones (25.8%) that were significantly less selected overall (all $p < .05$) A similar effect is seen for the no-fill variations (without D). Again, rotated targets were most commonly selected (47.2%), followed by data (36.5%) and scaled (16%) ones, with scaled once again being significantly less selected than the other two (all $p < .001$). In our further analysis we treat each target type as a separate dependent variable (Figure 100).

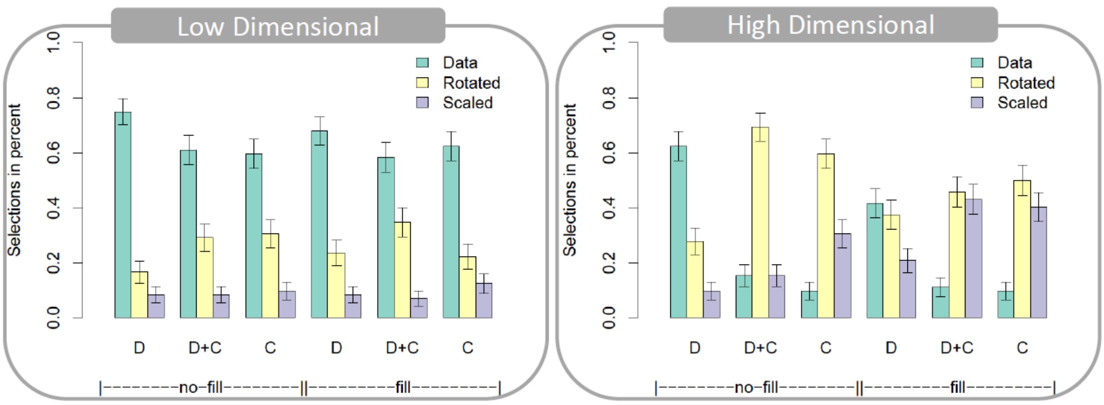


Figure 100: Experiment 2 results: The bar charts illustrate the percentage of selections and the standard deviation for each factor. The left chart represents low dimensionality, the right one the high dimension condition. Even without training or explaining the visual encoding participants using variation D judged data similarity rather than shape similarity.

Star glyph variations: There was a significant effect of contour variation on data target type ($\chi2(2;N = 288) = 32$; $p < .001$), on rotated target type ($\chi2(2;N = 288) = 12.8$; $p < .01$), and on scaled target type ($\chi2(2;N = 288) = 7.6$; $p < .05$). Post-hoc tests revealed significantly higher selection rates for data targets in variation D (61.8%) compared to D+C (36.5%, $p < .001$) and C (35.4%, $p < .001$) for both fills. Rotated targets were selected significantly less in variation D (26.4%) compared to D+C (44.8%, $p < .001$) and C (40.6%, $p < .05$), while scaled ones significantly less in variation D (11.8%) compared to C (23.3%, $p < .01$). There was also an effect of dimensionality on data target type ($\chi2(1;N = 432) = 32$; $p < .001$), on rotated ($\chi2(1;N = 432) = 26.1$; $p < .001$), and on scaled target ($\chi2(1;N = 432) = 8.3$; $p < .01$). Participants working with low dimensionalities selected the data target type significantly more often (64.1%) compared to the high dimensional condition (25%, $p < .001$) across all designs. In the high dimensional condition participants selected the rotated (48.4%) and scaled (26.6%)

target type significantly more often compared to the low dimensional condition (26.2%, $p < .001$ and 9%, $p < .01$). More details on dimensionality are reported for each fill type later on.

Fill vs. No-Fill Star Glyphs: We consider variation D neither as fill nor as no-fill (common across both experiment groups) and remove it from the analysis. Comparing the fill and no-fill variations we found a significant effect of filling types on rotated ($\chi2(1;N = 144) = 4.8$; $p < .05$), and on scaled target type ($\chi2(1;N = 144) = 8.2$; $p < .01$).

Post-hoc tests revealed a significantly higher selection rate for the scaled target type for fill designs (25.7%) compared to no-fill (16%, $p < .001$) and for the rotated target type for no-fill designs (47.2%) compared to fill (38.2%, $p < .05$).

No-Fill Star glyphs: The No-Fill star glyphs showed a significant effect of contour variation on data target type for both low ($\chi2(2;N = 72) = 8.21$; $p < .05$) and high dimensional cases ($\chi2(2;N = 72) = 28.25$; $p < .001$). Post-hoc tests revealed a significantly higher selection rate for data target type in variation D for the low and high dimensional case (75%; 62.5%) compared to D+C (61.1%; 15.3%, all $p < .05$) and C (59.7%; 9.7%, all $p < .01$). The No-Fill star glyphs also showed a significant effect of contour variation on rotated target type for both low ($\chi2(2;N = 72) = 7.7$; $p < .05$) and high dimensional cases ($\chi2(2;N = 72) = 14.6$; $p < .001$). Post-hoc tests revealed a significantly higher selection rate for rotated target types for both the low and high dimensional case in variation C (30.6%, 59.7%) and D+C (29.2%, 69.4%) compared to D (16.7%, 27.8%) (all $p < .05$).

Filled Star glyphs: The filled star glyph had a significant effect of contour variation on data target type in the high dimensional case ($\chi2(2;N = 72) = 17.33$; $p < .001$), and on scaled target type in the high dimensional case ($\chi2(2;N = 72) = 8.5$; $p < .05$). Participants working with variation D in high dimensions selected the data target type significantly more often (41.7%) compared to D+C (11.1%, $p < .001$) and C (9.7%, $p < .001$). The scaled target type was selected significantly more often with variation D+C (43%) and C (40.3%) compared to D (20.8%; $p < .01$ and $p < .05$) in high dimensions.

Variation D: We looked at variation D which is common across fill and no-fill conditions, and found that data targets were selected significantly more in the no-fill (62.5%) than the fill condition (41.6%, $p < .5$). Further analysis shows this is likely due to the order of presentation: in the fill condition, when D was the first design seen, data targets were selected more often (50%), than when they followed another fill design (35%). We explain this in our discussion.

6.2.2.4 Discussion

Independent of the fill type, participants using the D glyph variation selected the data target as more similar significantly more often than any other type, giving strong evidence that glyphs without contours promote data similarity comparison rather than shape (H1). Moreover, variation D was the one that the data target was most commonly selected compared to contour variations C; D+C irrespective of fill type (H2).

On the other hand, the most selected targets in contour variations C+D; C were indeed either rotated or scaled variations of the stimulus (H3). This reinforces our findings from the first study that factors enforcing perceptual unity of shape[149], such as contour containment lead viewers to naturally make shape judgments of similarity rather than data, while open variations of the glyphs lead to similarity choices closer to data comparisons, even without being told what similar means. Also, although not statistically significant, the C+D variation tended to have on average more data target selections than simple C.

The above effects are due mainly to the high dimensional condition. In the low dimensional condition, across all glyph designs, data targets were the ones more often select than all other target types (H5).

When comparing filling types we could not prove that filled star glyphs promote shape judgments more strongly than no-fill star glyphs. Nevertheless, in the fill condition, when the common data-lines design D appeared after fill designs, data selections dropped. We hypothesize that seeing a fill design first put participants in a frame of mind of making shape rather than data judgments, a behavior they carry on to the D design that otherwise promotes data similarity. Nevertheless, we saw no significant difference for the variations C+D;D that can actually hold fill color.

Thus, contrary to hypothesis H4, there was no difference in the selection of data targets across fill type. In our experiment the stronger figure and ground distinction that in the past has been shown to promote unity of shape[149] did not have a noticeable effect in data selections. Perhaps, this finding is also related to the fact that the brain relates surface fill color largely to edge contrast information[134]. Yet, the nature of this perceptual phenomenon does warrant further research in general as the fill type did affect which shape-related similarities people chose. Rotated target types were selected more often with no-fill star glyphs, whereas participants using fill star glyphs more frequently selected scaled target types.

We note again that in this study participants were never told that they were viewing data visualizations, they were just asked to find the most similar glyphs without further instructions. Thus, our results indicate the natural tendency of people to judge glyphs instinctively in a more “data-centric” manner in low dimensionalities, and in high ones when factors that enforce coherent shapes are absent. It is clear that with training we can further enforce data similarity judgments–but given that some glyphs and glyph variations seem to be naturally well suited for data judgments, we focus on those designs and try to further improve their performance with small design variations.

6.2.3 Experiment 3: Improvements for Star Glyph

The first experiment showed that people judge data similarity with non-contour designs more accurately while the second experiment showed that non-contour designs also lead to data similarity judgments to be made more naturally. Yet, accuracy in the high-dimensional case was quite low for all main design variations we tested previously. In this last experiment, we thus explore whether we can improve the accuracy of data similarity judgments by adding simple reference structures—tickmarks and grids—to the designs. We focused on static reference structures to learn how much these general approaches would aid data comparison before considering the design of interactive aids.

6.2.3.1 Star Glyph Reference Structures

Reference structures such as grids and tickmarks are frequently recommended for data charts to aid in relating content to axes[152]. We, thus, hypothesized that they could provide similar reading aids for star glyphs despite their smaller footprint. Tickmarks and grids use two different types of reference mechanisms. While tickmarks add information to each individual data line only, grids connect the overall glyph design. While there are many different ways to draw grids and tickmarks we settled on the following designs:

Tickmarks T: Whenever a data line exceeds a certain threshold we draw a short orthogonally oriented tickmark on the data lines using the same stroke color. Tickmarks are spaced to be 17 pixels apart. The resulting D+T glyph (see Figure 101) resembles the snowflake glyph previously mentioned in literature[90] and is also close to how tickmarks are used on axes in many data charts.

Grid G: We draw three circles in the background of the glyph using a gray value of #ccc in RGB color space chosen according to design considerations by Bartram et al.[153]. The circles are spaced 16.6 pixels apart. The resulting design resembles radar graphs or spider plots[154]. As an alternative we considered drawing a gridline at the end of each data line. Doing so would create an underlying texture that could help to identify the overall data distribution across all dimensions. Yet, we chose not to use this design as this texture can be misleading since rotated star glyphs with similar data values would have the same texture, although they have entirely different data values.

Of course, the readability of glyphs could further be improved by adding double encodings (e.g., additionally using color to distinguish dimensions or data values), dimension ordering[155], or sorting the glyphs on the display. Yet, all of these encodings have limitations: use of color is limited to glyphs with a small number of dimensions, dimension ordering may not improve legibility for a large number of variable glyphs in a small-multiple setting, and sorting glyphs may disrupt a pre-defined layout based on other meta-data such as time. We, thus, did not consider these encodings for the study.

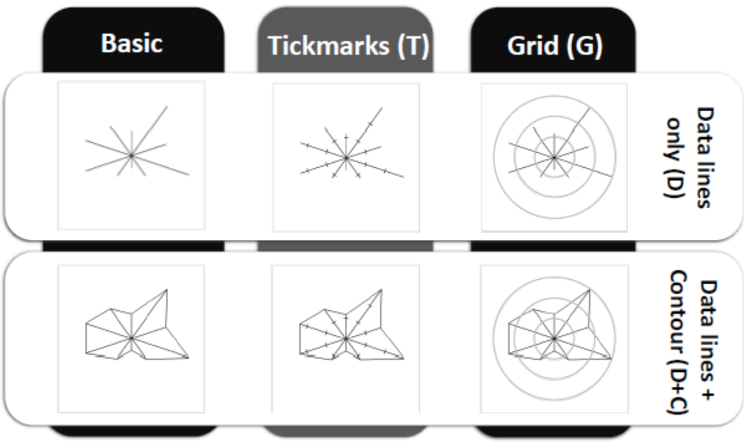


Figure 101: Experiment 3 design space: We have chosen the star glyph only with data whiskers (D) and with an additional contour line (D+C) and applied tickmarks (T) and gridlines (G) to these designs.

6.2.3.2 Design and Procedure

Glyphs: We tested the two star glyph variations that performed best in the first experiments: the data-only glyph (D) and the star glyph with data lines and a contour line (D+C). The reason for discarding the contour only design (C) is the bad performance for previous similarity judgments, the lack of ability to place tickmarks, and the minimal number of real-world examples of this glyph type in use.

For baseline comparisons we kept the originally tested versions of the star glyph (D, D+C) and added two types of reference structures (T, G). The experiment, thus, compared the six different designs (D, D+T, D+G, D+C, D+C+T, D+C+G) in Figure 101.

Participants: We recruited 12 data visualization experts (3 female). The age ranged from 23–40 years in age (mean (29.75) & median age (30)). All participants reported normal or corrected-to-normal vision. All experts focused during their studies on data visualization (4 Bachelor; 5 Master; 3 PhD) or a related topic and were familiar with reading data glyphs. They had not participated in the first study.

Task and Procedure: Participants completed data similarity search trials with all 6 designs. The order of the designs was randomized using a latin square. For each design there was a short introduction of the visual encoding and the similarity search task with 5 test questions. The participants had to complete those simple test trials with 80% accuracy in order to continue the experiment. The purpose of the test was to first check the participants' ability to read the visual encoding of the glyph and second to test their data similarity judgments. All participants passed the test section. The introduction was followed by 4 training trials to help the participants develop a strategy for solving the task. For training trials, the correct answer was shown to participants after they had made a choice. Finally the four study trials were shown without any visual feedback of the correct answer. The experiment took place in a lab setting using a 24" screen with a resolution of 1920 × 1200 pixels. The experimenter was present during the study. After the study, 11 of the 12 participants filled out a questionnaire for subjective feedback on aesthetics of the designs and strategies used to answer the questions.

Data, Distracters and Dimensionality: Since participants were already 80% correct in the low dimensional condition in Experiment 1, we only used high-dimensional glyphs in Experiment 3. We generated the data the same way as in Experiment 2 and balanced selection likelihood between distracters. To reduce the chance of a successful random guess we generated only one data point closest in data space (target) and another one second closest in data space (alternative) as in Experiment 1. The experiment included 2 rotated, 2 scaled, 2 random, 1 alternative and 1 target glyph. The stimulus was highlighted and positioned in the middle of the 3×3 matrix as in the two previous experiments. The distracters were randomly arranged around the stimulus.

6.2.3.3 Hypotheses

Based on our previous experiments and the frequent use of reference structures to aid chart reading, we tested the following hypotheses:

1. Tickmarks (T) in star glyphs improve the accuracy of data similarity judgments for both (D) and (D+C) variations compared to the variations without the tickmarks. The additional anchor points help to better read and compare line distances.

2. An underlying grid (G) in the background of the star glyph provides additional orientation and facilitates more accurate comparison of data values for both (D) and (D+C) variations than the variations without the grid.

3. The contour variation D+C benefits more from the additional reference structures than the D variation since contour has preciously shown to lead to shape comparison rather than data similarity comparisons.

4. Completion time is higher for designs enriched with reading marks (T or G). The viewer has to incest more mental effort to process the additional visual information.

6.2.3.4 Results

Similarly to Experiment 1 we used a non-parametric Friedman's Test on the data to analyze accuracy, and a one-way ANOVA for the completion time. We only report statistically significant results ($p < .05$).

The overall accuracy was 51.4%, with designs with grids (G) being more accurate (59.4%), followed by the tickmark designs (T) (47.9%) and then designs without additional marks (46.9%). There was a statistical trend for different types of reference structures on accuracy ($p<.1$), with glyphs with grids being more accurate than with tickmarks. There was no difference between designs with reference structures and the baseline design. Next, we compared the different glyph variations without contour (D) and with contour (D+C). As in Experiment 1, participants were significantly more accurate with variation D (60.4%) than when the contour was present D+C (33.3%; $p < .01$).

Reference structures on glyphs without contours (the D glyphs) did not significantly improve accuracy over the glyph without the reference structure. Participants were 60.4% accurate with D, 68.8% accurate with (D+G), and 45.8% accurate with (D+T). Nevertheless, we note that the mean accuracy of the (D+G) variation is indeed higher than for D only. We also found that for the two variations using reference structures, grids (D+G) were significantly more accurate than tickmarks (D+T) (45.8%; $p < .05$).

For the contour variations, we have a statistical trend ($p < .1$) indicating that the accuracy of both the contour variation with a grid (D+C+G) and the one with tickmarks (D+C+T) tend to be more accurate (both 50%) than that of simple glyph with contour (D+C) with accuracy 33.3% ($p = .06$ and $p = .08$ respectively).

Looking at differences across variations, we also found that D+G (68.8%), which had the highest overall mean accuracy, performed significantly better than D+C (33.3%; $p < .001$) and had a statistical trend to perform better than D+C+G ($p = .1$) and D+C+T ($p = .8$).

The mean number of selections per distracter type are shown in Figure 102. We found a significant effect of variation on distracter ($\chi2(5;N = 48) = 12.68$; $p < .05$). Participants using variations with contour lines most often selected the scaled distracter (24%) followed by the rotated (16%) and the alternative (15%) distracter. For the non-contour variations participants chose the alternative and the rotated distracter equally often (18%) followed by the scaled distracter (5%).

No significant results can be reported for the completion time, thus we cannot confirm that additional marks influenced comparison times. However, participants needed approx. 2sec longer when working with designs using additional marks. Average completion time was 22sec per trial (D = 21.7sec, D+G = 24.8sec, D+T = 26.1sec, D+C = 17.9sec, D+C+G = 21.5sec, D+C+T = 22sec).

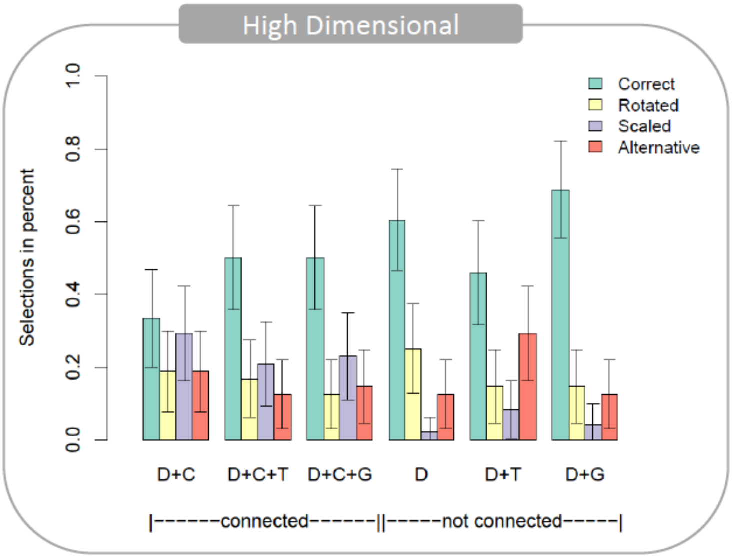


Figure 102: Experiment 3 results of the percentage of selections and the standard deviation for each factor. Design improvements (T, G) do not significantly increase the accuracy of the two star glyph variations (D+C, C).

The questionnaire showed that the glyph variations with contours ranked highly amongst participants' aesthetic preferences. The mostly strongly preferred glyph variation was D+C+G (5/11 participants), followed by D+C (3/11 participants). Interestingly, no participants preferred the D variation even though its mean accuracy (60.4%) was higher than D+C+G (50%). Participants also ranked the D variation as hard to use (median=6 on a 7-point Likert scale) with all other designs ranking at least between median 4–2. The D+C+T and D+C+G variations were both found easy to use (median=2). We report on the results of the questions regarding strategy in our discussion section.

6.2.3.5 Discussion

Adding reference structures to the star glyph did not have the effect on accuracy we were expecting for our data similarity search task. Additional anchor points on the data line (i.e., tickmarks) did not significantly improve the comparison of data points. Therefore, we cannot accept H1. Nevertheless, there was a statistical trend indicating that an overall reference in the background (i.e., gridlines) may increase accuracy, especially in the case of contour star glyphs, providing some evidence for H2.

This lack of strong significant effects is surprising, especially given that most participants mentioned in the questionnaire that for the simple star glyph D, gridlines (81%), and to a lesser extent tickmarks (72%), helped them find the most similar data point. Although the mean accuracy for the D+G variation was indeed higher, the effect was not significant, perhaps due to the already very good performance of the D variation. The value of gridlines and tickmarks in general may warrant further research. As Few notes[156], gridlines may be useful only in specific cases, e. g., when small differences have to be compared. Therefore, it is possible that for other tasks, such as direct lookup, these additional reference marks could help more strongly.

For the star glyph with contour (D+C), only 54% of our participants reported using tickmarks and 36% gridlines to complete the task. From their reports they felt (erroneously) that glyphs with contours are easier to compare and, thus, did not make conscious use of the additional improvements. Thus, in the contour case, participants were not only more error prone, but also misled to feel confident in their choices, ignoring the marks that could help them improve their performance. Nevertheless, it is highly likely that the addition of reading marks was taken into account, even if unintentionally, explaining the trend we see for both the tickmark and grid variation to be more accurate than simple contour glyphs (H3).

Finally, we could not confirm H4 due to a lack of significant results when comparing task performance time.

Even though participants using variation (D) performed very well, it is interesting that they did not like this design variation. On a 7-step Likert scale 63% of the participants rated the design with either 6 (difficult to use) or 7 (very difficult to use). Most participants (46%) preferred the star glyph with contour and gridlines, with only 1 participant rating it with a 5 (slightly difficult to use) and the others with 3 or better.

Given the results of this experiment the benefit of using reference structures for star glyphs is limited. Especially since in real world scenarios when multi-dimensional glyphs are projected to two dimensional surfaces, there is the possibility of over-plotting, and adding marks or gridlines could worsen this effect due to the additional ink introduced.

6.2.4 Design Considerations

With the results gained from the analysis and discussions we derive the following design considerations.

1. **When judging data similarity avoid contours in glyph designs.**

   Viewers have a natural tendency to judge data similarity in star glyphs without contours. In all our experiments viewers were tricked into making shape-based, rather than data-based judgments when using contours. This is especially true if glyphs in the visualization are scaled or rotated versions of each other.

2. **For low number of dimensions (around 4) any glyph variation can safely be used for data similarity judgments.**

   In the first and second experiment viewers naturally leaned towards data similarity for each glyph variation in low dimensions, even without training.

3. **When there is a need for contours, add data lines to the design to strengthen data similarity judgments**.

   Participants independent of glyph design (fill or no-fill) judged data similarity better using the D+C variation compared to C in the first two experiments. Although, there was no statistical significance, mean data comparisons for contour + data variations were always higher than contour only.

4. **When there is a need for contours, the designer can decide whether or not to use fill color.**

   Our Experiment 2 gave no indication that fill color degrades the performance of glyphs with contour.

5. **When clutter is an issue avoid reference structures in non-contour star glyphs for similarity search tasks**.

   Results of Experiment 3 illustrate that even though participants preferred using tickmarks or grids they did not perform significantly better with them, especially for glyphs without contours. Nevertheless, there is a statistical trend that shows that tickmarks and grids improve glyphs with contours.

6. **If references are required use grids rather than tickmarks**.

   Independent from the design (i.e., with or without contour) gridlines always increased mean accuracy, which is not true for tickmarks

6.2.5 Conclusion

Making use of the results and design considerations of our user studies we are able to develop visual representations most suitable and effective for similarity search tasks in multi-dimensional space. Therefore, this pre-study was an essential starting point for developing an appropriate visual analytics prototype.

As an addition to the visual comparison we would like to help domain experts in detecting similar data items by giving her strong and easy to use interaction techniques. Therefore, we make use of tangible data analysis to facilitate data comparison.

6.3 Visual Alignment: A Technique to Facilitate the Comparison of Multi-Dimensional Data Points

Identifying similar data points can be done automatically be applying clustering algorithms. However, especially in high-dimensional space it is a complicated task for the user to understand why data points have been clustered in a certain way. If the user for example tries to understand the automatic clustering a visual output of the result space is beneficial.

We have already introduced different visualization techniques to represent multi-dimensional data points. Well-known examples are scatterplot matrices, parallel coordinate plots, or various glyph designs. Visual alignment is an automatic algorithm, which can be applied to all these visualizations. As a result the user is able to compare multi-dimensional data points with each other and better reason about possible clusters or groupings.

The idea is simple. The analyst selects one multi-dimensional data point (her point of interest), which he would like to compare to all the other elements (i.e., 1 x n comparison). This data point is then considered as new baseline by adapting the visualization to automatically show the difference of all elements to this data point. In the following we are going to explain this technique in more detail for specific visualizations.

6.3.1 Visual Alignment for Scatterplot Matrices

A scatterplot matrix consists of single scatterplots arranged in a matrix layout. Therefore, multi-dimensional data points can only be compared across two dimensions. To keep track of single data points it is beneficial to identically color some of them (e.g., the points of interest) to keep track of their position throughout the whole visualization.

With the visual alignment technique the selected data point is additionally positioned at the center point of each single scatterplot making it really efficient and effective to spot data points with lower or higher data values across all dimensions. Therefore, this data point acts as the new baseline for all the other elements (Figure 103). By using animation, all elements are repositioned to fit to the new baseline. Understanding the relation between the point of interest (the new baseline) and all the other elements is facilitated. For example, all Elements in the upper right corner have higher data values compared to the baseline for both dimensions.

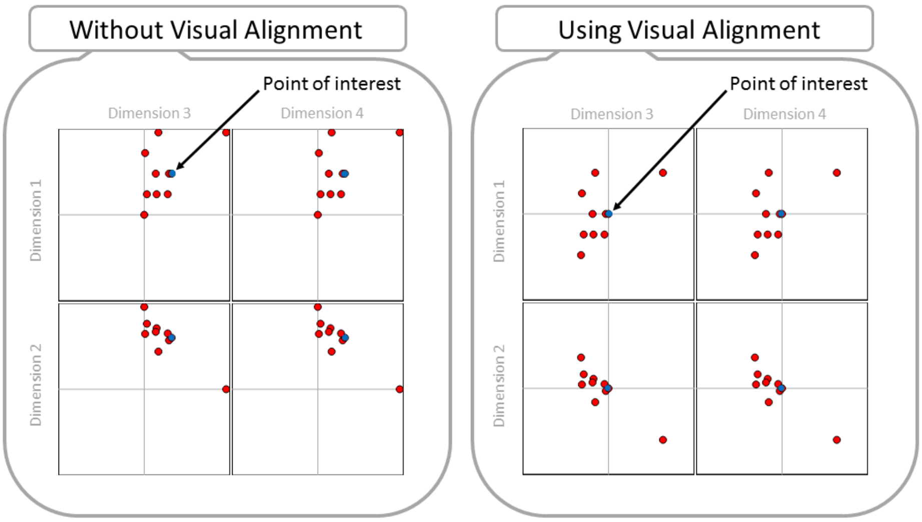


Figure 103: Visual Alignment for scatterplot matrices: Common scatterplot with one data point highlighted as point of interest (left). After selecting this data point it is considered the new baseline and smoothly moved to the center. All other elements are repositioned accordingly (right).

6.3.2 Visual Alignment for Parallel Coordinate Plots

In a parallel coordinates plot each data point consists of a poly line intersecting with the dimension axes at the corresponding value. When comparing multi-dimensional data points the analyst has to follow one or more data lines and compare the different intersection points with each other. With an increasing number of dimensions or zig-zag-patterns this task gets more and more difficult.

Applying our visual alignment technique shifts the selected data line at the center position of each dimension. The other data lines are adjusted to this new baseline. Detecting data points higher or lower to the new baseline is now a trivial task.

6.3.3 Visual Alignment for Data Glyphs

Using visual alignment in a glyph setting is a more complex but also a more interesting alternative. Basically there are two ways of arranging a glyph on the screen either data-driven by using their data values to position the glyphs (e.g., in a scatterplot), or structural-driven by showing different kinds of relationships (e.g., hierarchical relation with a treemap).

Visual alignment can now be used to position the glyphs in a structural way (e.g., on a geographic map) and additional change the data values of each single glyph to see data relations. Again, one data point (i.e., one glyph) is selected. This glyph acts as the new baseline by changing the respective values to 0. The other glyphs on the screen adjust to this new baseline by showing the difference between their raw data values and the new baseline (Figure 104). Therefore, comparisons according to each dimension between the selected glyph and all other elements can easily be done without changing the position of the glyphs.

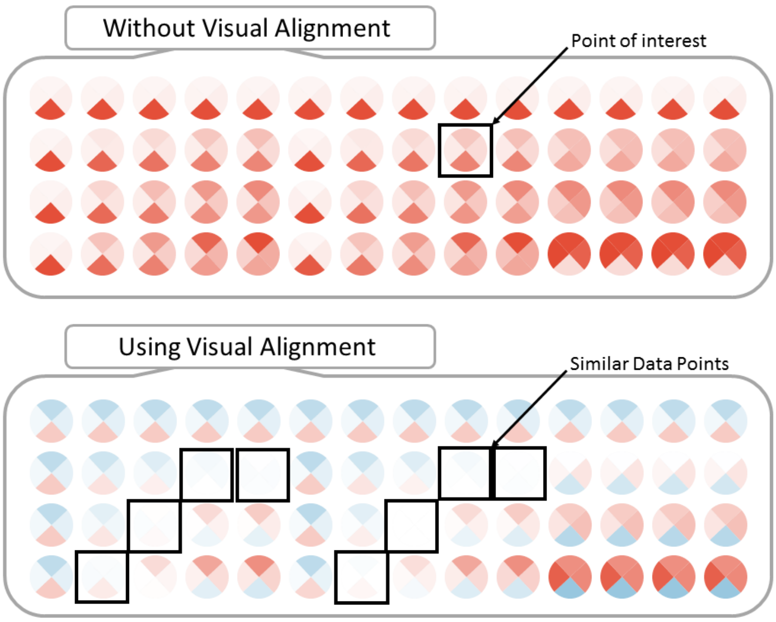


Figure 104: Visual Alignment for glyph designs: Color saturation is used to encode the data value for each dimension (top). After selecting a point of interest the values of the glyph are considered as new baseline, thus, all elements change their coloring to fit the new baseline (bottom)

6.3.4 Conclusion and Future Work

The visual alignment technique allows the analyst to do a 1 x n comparison of multi-dimensional data points by selecting one single point of interest. As a next step we would like to offer the analyst a glyph based overview visualization by showing her an overall comparison of all multi-dimensional data points making use of an extended visual alignment technique. Such an overview could be a useful start for an explorative analysis by pointing to interesting areas with similar or entirely different characteristics.

6.4 Visual Analytics Prototype

Our visual analytics prototype is web-based making use of HTML, JavaScript and D3. It can be tested online (http://consensus.dbvis.de/alternativescenario). The tool consists of 2 components.

1. **Analytic component**: This component supports the user with automatic algorithms, which can be interactively steered by the analyst.

2. **Visualization component**: This component displays the data space with multiple views and allows the user to interact with the underlying data.

The two components are tightly coupled to allow the analyst adjusting parameters and visually investigating the change. The workflow is quiet simple and need not be followed in a certain order.

As a default setting the data is visualized in a scatterplot matrix. The domain expert gets a first idea of the raw data and sees possible correlations between two dimensions. With her deep understanding of the data the domain expert can help the tool to understand the different dimensions more appropriately. For example, when visualizing country characteristics a high income or a good education level corresponds to a higher position in the scatterplot or parallel coordinate plot. However, a high crime rate has a negative impact and would be represented again with a higher position. To avoid a possible misunderstanding the analyst can inverse the scale of each single dimension individually (Figure 105). A high crime rate would then correspond to a low position. This improves the overall visualization by having a more intuitive encoding of the underlying data increasing the trustworthiness of the visualization.

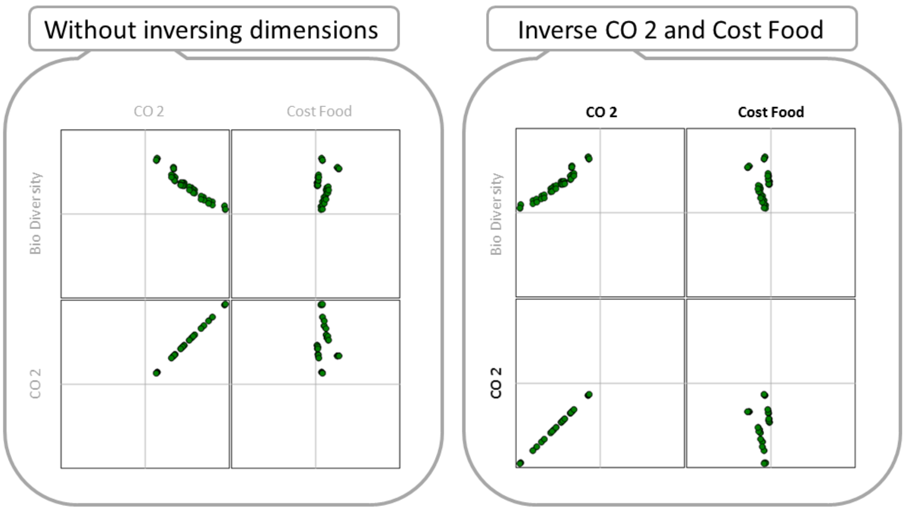


Figure 105: Inverse Functionality: The scatterplot matrix visualizes the three dimensions Bio Diversity, CO2 and Cost Food (left). CO2 and Cost Food are marked as inverse since a higher value has a negative meaning. Therefore, the visualization changes showing this fact (right).

To focus on specific data points the user can apply different filters. This is dependent on the underlying data set. For time series data the user can for example just focus on certain points in time by hiding the others. In the Consensus project these filters can be seen as different input parameters for the simulations. By adjusting those filters the visualization will only show the scenarios meeting the predefined input parameters.

To get another perspective on the data the analyst can choose between three different visualization techniques. Besides a scatterplot matrix the tool offers a parallel coordinate plot and a glyph based visualization. Depending on the task the analyst can switch between the different representations interactively. The settings or filters remain active.

The basic parallel coordinate plot is extended with additional functionality to improve the analytical process. Most important is the ordering of the dimension axes to see possible correlations. The user can do this interactively by selecting one axis and moving it to a different location. Additionally, the user can delete axis or add them as he saw fit.

Especially for the Consensus project detecting alternative scenarios, which are inferior to others is a major analysis task. Therefore, the analyst can trigger an automatic algorithm, which filters out all inferior data points. These data points are then hidden or marked in each visualization (Figure 106). In each visualization the user can easily see the reason why certain data points are inferior to others. The visual feedback helps the domain expert to better reason about the consequences of adjusting certain input parameters.

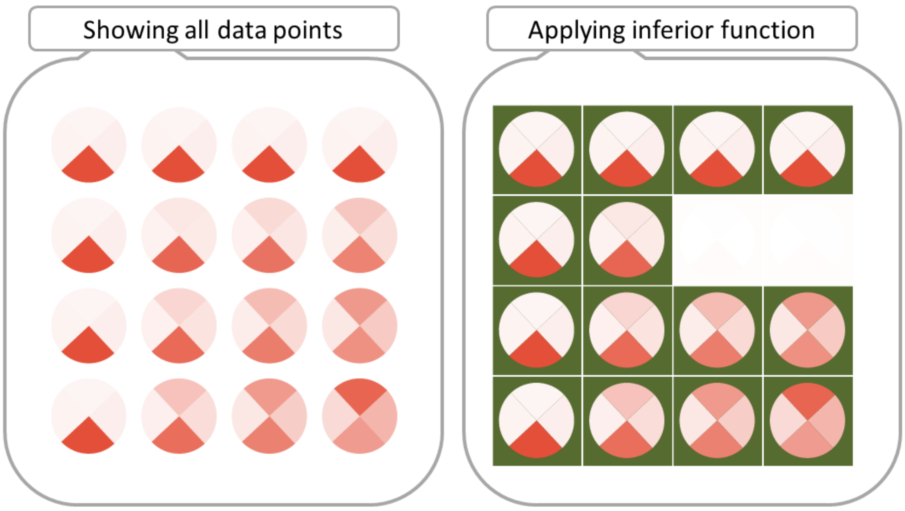


Figure 106: Inferior Function: Each glyph represents a multi-dimensional data point (left). Data points, which have lower data values for each dimension compared to the others are considered as inferior. These alternatives need not be considered in the final analysis and are, therefore, hidden (right).

In the glyph visualization the data points are arranged in a matrix layout according to their ID. However, the user can change the layout by applying a PCA projection on the multi-dimensional data points. The elements are then projected to 2d space according to their Eigenvalues (Figure 107). By keeping a detailed glyph representation the analyst can easily reason about outliers or glyphs arranged at similar positions by investigating the single dimension values of each glyph. Depending on the preferences of the analyst the user can change the glyph design from a radial color encoding to a linear length encoding of the data dimensions and values.

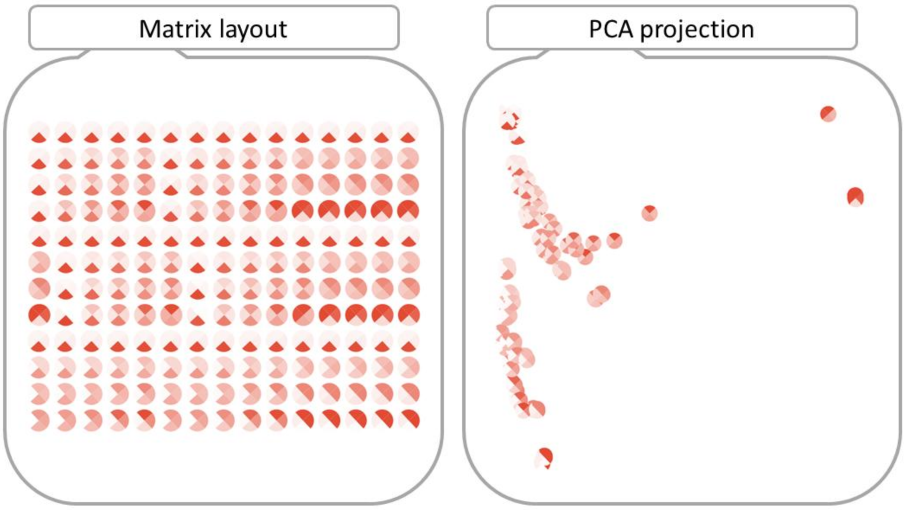


Figure 107: As a default setting glyphs are positioned in a matrix (left). By applying a PCA the analyst can detect two outliers and two main clusters (right).

To further improve the comparison of data points the previously introduced visual alignment technique can be applied to each visualization. The user has to simply click on a data point, which is then considered as the new baseline. All elements are automatically repositioned by using animation to keep track of the change.

6.5 Future Work

As a next step we would like to enhance our visual alignment technique to be able to perform an nxn comparison of data points. This would support the user in her exploratory analysis by offering her a better overview visualization of the underlying data.

Especially for the glyph representation we aim at developing additional layout algorithms like for example an arrangement on a geo-graphic map. This would allow the analyst to draw conclusions about spatial characteristics due to some visible glyph patterns. For improving the analysis of the road pricing use case we would like to implement a graph-based visualization to show connections between corridors and their relationships among each other.

6.6 Workflow

In this chapter we want to briefly show the workflow of the visual analytics prototype. The tool is a web-based application and runs with the most common browsers (http://consensus.dbvis.de/alternativescenario). The screen is divided in two areas. The settings at the top and the visualization space at the bottom. In the settings menu the user can load data files, switch between visualizations and apply different analytical features.

As a default setting a data set provided by the project partner IIASA is loaded and visualized in a scatterplot matrix.

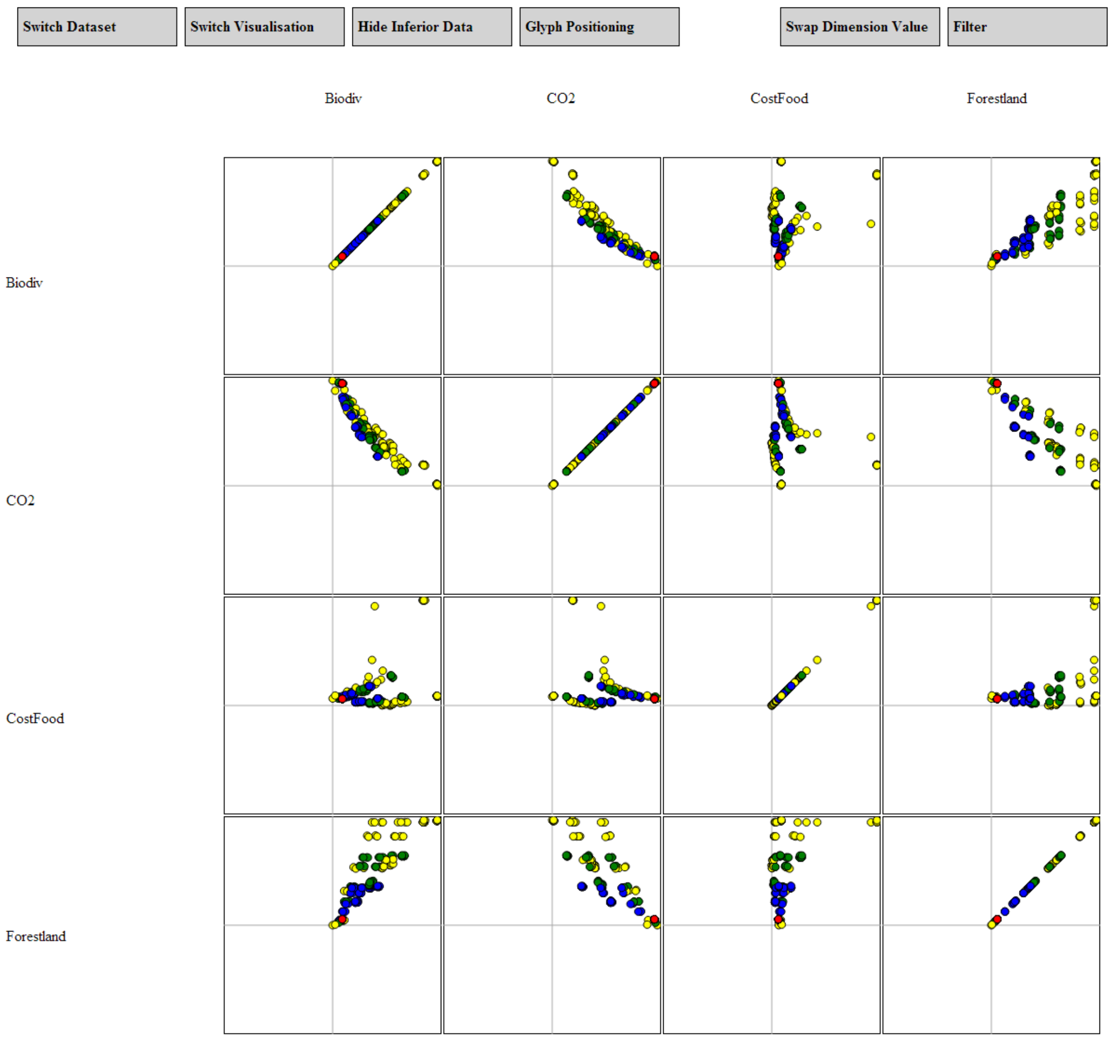


Figure 108: Visual Analytics Prototype: Default setting showing a scatterplot matrix with no further algorithms applied.

As a first step the analyst makes use of her background knowledge and swaps the dimension values of CO2 and CostFood because high data values correspond to a negative outcome. Therefore, he clicks the box "Swap Dimension Value" and selects the respective dimensions.

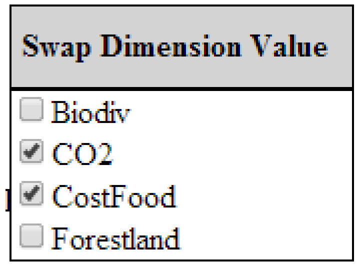


Figure 109: Dimension Swapping: Several dimensions can be selected via checkboxes to swap their values.

The visualization updates immediately and shows the new data distribution. To focus on the Pareto optimal solutions the analyst hides all inferior data points by selecting the option in the settings menu. Because he is interested in solutions supporting a high bio diversity he selects the upper right data point in the scatterplot showing two times bio diversity. The visual alignment technique arranges all data points in the whole scatterplot matrix according to the new baseline.

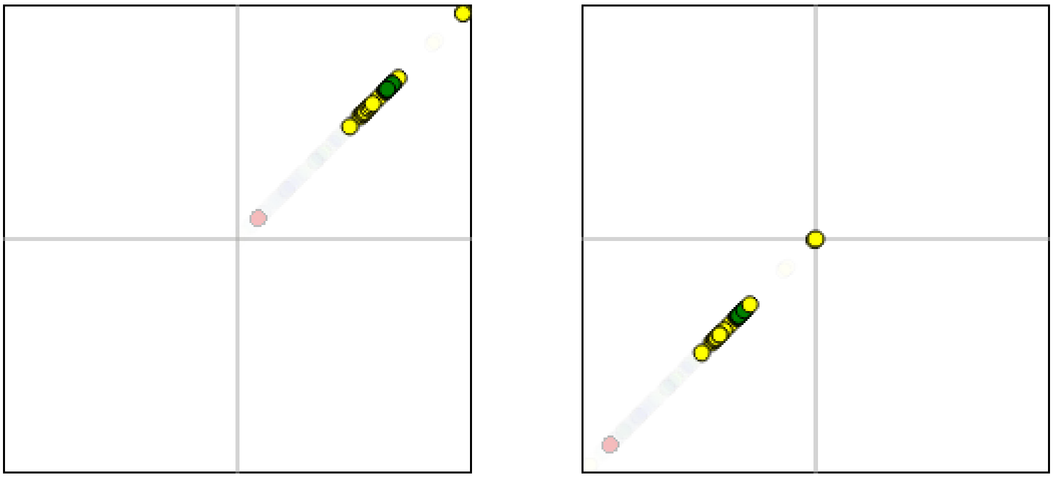


Figure 110: Visual Alignment: After selecting the upper right data point in the left scatterplot the visual alignment technique is applied arranging all data points according to the new baseline.

Keeping the mouse over the selected data point automatically highlights this data point in each single scatterplot. After scanning the visualization the analyst recognizes that selecting the data point with the highest bio diversity results in a solution not optimal for the cost food dimensions. Some data points have higher values for this dimension.

Bibliography

[1] I. Das and J. Dennis, "A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multucriteria optimization problems.," *Structural Optimization,* vol. 14, pp. 63-69, 1997.

[2] I. Das and J. Dennis, "Normal-Boundary Intersection: A New Method for Generating the Pareto Surface in Nonlinear Multicriteria Optimization Problems," *SIAM Journal on Optimization,* vol. 8(3), p. 631, 1998.

[3] S. Motta, Renato, Afonso, M. B. Silvana, Lyra and R. M. Paulo, "A modified NBI and NC method for the solution of N-multiobjective optimization problems," *Structural and Multidisciplinary Optimization,* 2012.

[4] A. Messac, A. Ismail-Yahaya and C. A. Mattson, "The normalized normal constraint method for generating the Pareto frontier," *Structural and multidisciplinary optimization,* vol. 25, pp. 86-98, 2003.

[5] A. Messac and C. A. Mattson, "Normal constraint method with guarantee of even representation of complete Pareto frontier," *AIAA journal,* vol. 42, no. 10, pp. 2101-2111, 2004.

[6] D. Muller-Gritschneder, H. Graeb and U. Schlichtmann, "A Successive Approach to Compute the Bounded Pareto Front of Practical Multiobjective Optimization Problems," *SIAM Journal on Optimization,* vol. 20, no. 2, pp. 915-934, 2009.

[7] T. Erfani and S. Utyuzhnikov, "Directed Search Domain: A Method for Even Generation of Pareto Frontier in Multiobjective Optimization," *Journal of Engineering Optimization,* vol. 43, no. 5, pp. 1-18, 2011.

[8] T. Back, Evolutionary Algorithms in Theory and Practice, New York, NY, USA: Oxford University Press, 1996.

[9] O. Shir, J. Roslund, Z. Leghtas and H. Rabitz, "Quantum control experiments as a test bed for evolutionary multi-objective algorithms," *Genetic Programming and Evolvable Machines,* vol. 13, pp. 445-491, 2012.

[10] D. V. Arnold, Noisy Optimization with Evolution Strategies. Genetic Algorithms and Evolutionary Computation., Berlin: Springer, 2002.

[11] C. Fonseca and P. Fleming, "An Overview of Evolutionary Algorithms in Multi-objective Optimization," *Evolutionary Computation,* vol. 3, no. 1, pp. 1-16, 1995.

[12] E. Zitzler, K. Deb and L. Thiele, "Comparison of Multi-objective Evolutionary Algorithms: Empirical Results," *Evolutionary Computation,* vol. 8, no. 2, pp. 173-195, 2000.

[13] B. Suman and P. Kumar, "A survey of simulated annealing as a tool for single and multiobjective optimization," *Journal of the Operational Research Society,* vol. 57, no. 10, pp. 1143-1160, 2006.

[14] D. Craft, T. Halabi, H. Shih and T. Bortfeld, "Approximating convex Pareto surfaces in multiobjective radiotherapy planning," *Medical Physics,* vol. 33, no. 9, pp. 3399-3407, 2006.

[15] N. Beume, B. Naujoks and M. Emmerich, "SMS-EMOA: Multiobjective selection based on dominated hypervolume," *European Journal of Operational Research,* vol. 181, no. 3, p. 1653, 2007.

[16] R. Battiti, M. Brunato and F. Mascia, Reactive Search and Intelligent Optimization, Springer Verlag, 2008.

[17] R. Batiti and B. Mauro, Reactive Business Intelligence. From Data to Models to Insight, Trento, Italy: Reactive Search Srl., 2011.

[18] M. Ehrgott, Multicriteria Optimization, Berlin: Springer, 2005.

[19] J. Beder and E. Zitzler, "HypE: An Algorithm for Fast Hypervolume-Based Many-Objective Optimization," *Evolutionary Computation,* vol. 19, no. 1, pp. 45-76, 2011.

[20] C. H. Papadimitriou and M. Yannakakis, "On the approximability of tradeoffs and optimal access of web sources," in *FOCS*, Redondo Beach, CA, 2000.

[21] M. Masin and Y. Bukcin, "iversity maximization approach for multiobjective optimization," *Operations Research,* vol. 56, no. 2, pp. 411-424, 2008.

[22] [Online]. Available: www.ibm.com/software/integration/optimization/cplex-optimizer/.

[23] J. Knowles, D. Corne and K. Deb, Multiobjective Problem Solving from Nature: From Concepts to Applications, Berlin: Springer, 2008.

[24] J. Knowles, "Closed-loop evolutionary multiobjective optimization," *IEEE Computational Intelligence Magazine,* pp. 77-91, 2009.

[25] D. Buche, P. Stoll, R. Dornberger and P. Koumoutsakos, "Multiobjective Evolutionary Algorithm for the Optimization of Noisy Combustion Processes," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on,* vol. 32, no. 4, pp. 460-473, 2002.

[26] S. O'Hagan, W. B. Dunn, M. Brown, J. D. Knowles and D. B. Kell, "Closed-Loop, Multiobjective Optimization of Analytical Instrumentation: Gas Chromatography/Time-of-Flight Mass Spectrometry of the Metabolomes of Human Serum and of Yeast Fermentations," *Anal. Chem,* vol. 77, no. 1, pp. 290-303, 2005.

[27] L. Bonacina, J. Extermann, A. Rondi, V. Boutou and J. P. Wolf., "Multiobjective genetic approach for optimal control of photoinduced processes," *Phys. Rev.,* vol. 76, no. 2, 2007.

[28] T. S. Kuhn, The Structure of Scienti_c Revolutions. 3rd edn., Chicago: University of Chicago Press, 1996.

[29] T. S. Kuhn, "Objectivity, Value Judgement, and Theory Choice," *he Essential Tension: Selected Studies in Scienti_c Tradition and Change,* pp. 320-339, 1977.

[30] B. Roy, Multicriteria Methodology for Decision Aiding, volume 12 of Nonconvex Optimization and its Applications, Dordrecht: Kluwer Academic Publishers, 1996.

[31] H. Ching-Lai and A. Masud, Multiple objective decision making, methods and applications: a state-of-the-art survey, Springer-Verlag, 1979.

[32] K. Miettinen, Nonlinear Multiobjective Optimization, Springer, 1999.

[33] A. P. Wierzbicki, "A mathematical basis for satisficing decision making," *Mathematical Modelling,* vol. 3, no. 5, p. 391, 1982.

[34] M. Zeleny, "Compromise Programming," *Multiple Criteria Decision Making,* pp. 262-301, 1973.

[35] S. Gaas and T. Saaty, "The computational algorithm for the parametric objective function," *Naval Research Logistics Quarterly,* vol. 2, no. 1-2, pp. 39-45, 1955.

[36] K. Miettinen, F. Ruiz and A. P. Wierzbicki, "Introduction to Multiobjective Optimization: Interactive Approaches," *Multiobjective Optimization. Lecture Notes in Computer Science,* vol. 5252, p. 27, 2008.

[37] S. Zionts and J. Wallenius, "An Interactive Programming Method for Solving the Multiple Criteria Problem," *Management Science,* vol. 22, no. 6, p. 652, 1976.

[38] P. W. Andrzej, M. Makowski and J. Wessels, Model-Based Decision Support Methodology with Environmental Applications, Springer, 2000.

[39] H. Nakayama and Y. Sawaragi, "Satisficing Trade-Off Method for Multiobjective Programming," in *Interactive Decision Analysis*, Heidelberg, Springer-Verlag Berlin, 1984, pp. 113-122.

[40] K. Miettinen and M. M. Mäkelä, "Interactive bundle-based method for nondifferentiable multiobjeective optimization: Nimbus§," *Optimization,* vol. 34, no. 3, p. 231, 1995.

[41] K. Miettinen and M. M. Mäkelä, "Synchronous approach in interactive multiobjective

optimization," *European Journal of Operational Research,* vol. 170, no. 3, p. 909, 2006.

[42] M. Luque, F. Ruiz and K. Miettinen, "Global formulation for interactive multiobjective optimization," *OR Spectrum,* vol. 33, no. 1, pp. 27-48, 2008.

[43] F. Ruiz, M. Luque and K. Miettinen, "Improving the computational efficiency in a global formulation (GLIDE) for interactive multiobjective optimization," *Annals of Operations Research,* vol. 197, no. 1, pp. 47-70, 2011.

[44] [Online]. Available: http://en.wikipedia.org/wiki/Conjoint_analysis_%28marketing%29.

[45] P. C. Fishburn, "Decision and value theory," *Operations Research,* vol. 10, pp. 202-203, 1964.

[46] P. C. Fishburn, "Utility Theory for Decision Making," *Operations Research,* vol. 18, 1970.

[47] M. Köksalan, J. Wallenius and S. Zionts, Multiple Criteria Decision Making From Early History to the 21st Century, World Scientific, 2011.

[48] K. Steele, Y. Carmel and C. Wilcox, "Uses and misuses of multicriteria decision analysis (MCDA) in environmental decision making," *Risk Anal,* vol. 29, no. 1, pp. 26-33, 2009.

[49] V. Belton and T. Stewart, Multiple criteria decision analysis: an integrated approach, Springer, 2002.

[50] I. Omann, "How can Multi-criteria Decision Analysis contribute to environmental policy making? A case study on macro-sustainability in Germany," in *3rd Biennial Conference of the European Society for Ecological Economics*, Vienna, 2000.

[51] A. Balasubramaniam and N. Voulvoulis, "The appropriateness of Multicriteria Analysis in environmental decision-making problems," *Environ Technol,* vol. 26, no. 9, pp. 951-62, 2005.

[52] F. Joerin and A. Musy, "Land management with GIS and multicriteria analysis," *International Transactions in Operational Research,* vol. 7, no. 1, pp. 67-78, 2000.

[53] G. A. Mendoza and H. Martins, "Multi-criteria decision analysis in natural resource management: A critical review of methods and new modelling paradigms," *Forest Ecology and Management,* vol. 230, no. 1-3, pp. 1-22, 2006.

[54] Y. J. Zhang, !. J. Li and T. Fung, "Using GIS and Multi-criteria Decision Analysis for Conflict Resolution in Land Use Planning," *Procedia Environmental Sciences,* vol. 13, pp. 2264-2273, 2012.

[55] T. Buchholz, V. A. Luzadis and T. A. Volk, "Sustainability criteria for bioenergy systyems: results from an expert survey," *Journal of Clearner Production,* vol. 17, pp. 586-598,

2009.

[56] J. R. Ziolkowska, "Evaluating sustainability of biofuels feedstocks: A multi-objective framework for supporting decision making," *Biomass and Bioenergy,* vol. 59, pp. 425-440, 2013.

[57] L. Turcksin, C. Macharis, K. Lebeau, F. Boureima, J. Van Mierlo, S. Bram, J. De Ruyck, L. Mertens, J. Jossart, L. Gorissen and L. Pelkmans, "A multi-actor multi-criteria framework to assess the stakeholder support for different biofuel options: The case of Belgium," *Energy Policy,* vol. 39, no. 1, pp. 200-214, 2011.

[58] A. Perimenis, H. Walimwipi, S. Zinoviev, F. Müller-Langer and S. Miertus, "Development of a decision support tool for the assessment of biofuels," *Energy Policy,* vol. 39, no. 3, pp. 1782-1793, 2011.

[59] H. Mohamadabadi, "Development of a multi-criteria assessment model for ranking of renewable and non-renewable transportation fuel vehicles," *Energy,* vol. 34, pp. 112-125, 2009.

[60] T. Myllyvitta, P. Leskinen, K. Lähtinen, K. Pasanen, S. Sironen, T. Kähkönen and L. Sikanen, "Sustainabiltiy assessment of wood-based bioenergy – A methodological framework and a case-study," *Biomass and Bioenergy,* vol. 59, pp. 293-299, 2013.

[61] T. Buchholz, E. Rametsteiner, T. A. Volk and V. A. Luzadis, "Multi criteria analysis for bioenergy systems assessments," *Energy Policy,* vol. 37, pp. 484-495, 2009.

[62] T. Hayashi, E. Van Ierland and H. Zhu, "A holistic sustainability assessment tool for bioenergy using the Global Bioenergy Partnership (GBEP) sustainability indicators," *Biomass and Bioenergy,* vol. 66, pp. 70-80, 2014.

[63] J. Sliogeriene, Z. Turskis and D. Streimikiene, "Analysis and choice of energy generation technologies: the multiple criteria assessment on the case study of Lithuania," *Energy Procedia,* vol. 32, pp. 11-20, 2013.

[64] G. Montibeller and A. Franco, "Multi-Criteria Decsion Analysis for Strategic Decision making, Chapter 2, Applied Optimization," in *Handbook of Multicriteria Analysis*, Berlin, Springer, 2010.

[65] A. Deluka-Tibljaš, B. Karleuša and N. Dragičević, "Review of multicriteria-analysis methods application in decision making about transport infrastructure," *Građevinar Journal of the Croatian Association of Civil Engineers,* vol. 65, no. 7, pp. 619-631, 2013.

[66] DCLG, "Multi-criteria analysis: a manual," Department for Communities and Local Government, London, 2009.

[67] S. Basbas and C. M. Makridakis, "A Review of the Contribution of Multi-Criteria Analysis

to the Evaluation Process of Transportation Projects," *International Journal of Sustainable Development and Planning,* vol. 2, no. 4, pp. 387-407, 2007.

[68] European Commision, "Multi-Criteria Analysis," 2005. [Online]. Available: http://ec.europa.eu/europeaid/evaluation/methodology/examples/too_cri_res_en.pdf .

[69] A. V. Jensen, *Appraisal of Transport Projects: Assessing Robustness in Decision Making. PhD Dissertation,* Technical University of Denmark. Department of Transport, 2012.

[70] I. Omann, *Multi-Criteria Decision Aid As An Approach For Sustainable Development Analysis And Implementation. PhD Dissertation,* University of Graz, 2004.

[71] World Bank, "Multi-Criteria Analysis. Annex 4D.4," [Online]. Available: http://www.worldbank.org/urban/solid_wm/erm/Annexes/US%20Sizes/New%20Annex%204D.4.pdf.

[72] PROSPECTS, *Decision Makers' Guidebook. PROSPECTS Project: Procedures for Recommending Optimal Sustainable Planning of European City Transport Systems.,* Deliverable No. 15., 2003.

[73] GUIDEMAPS, "Successful transport decision-making. A project management and stakeholder engagement handbook. Volume I: Concepts and Tools," 2004. [Online]. Available: www.civitas-initiative.org/docs1/GUIDEMAPSHandbook_web.pdf. [Accessed 2013].

[74] I. C. Schutte and A. Brits, "Prioritising transport infrastructure projects: towards a multi-criterion analysis," *Southern African Business Review,* vol. 16, no. 3, pp. 99-117, 2012.

[75] K. C. Sinha and S. Labi, ransportation Decision Making: Principles of Project Evaluation and Programming, New York: Wiley, 2007.

[76] A. Awasthi, H. Omrani and P. Gerber, "Multicriteria decision making for sustainability evaluation of urban mobility projects," *CEPS/INSTEAD Working Papers,* 2013.

[77] D. Tsamboulas and A. Kopsacheili, "A Methodological Framework for the Strategic Assessment of Transportation Policies: Application for the Athens 2004 Olympic Games," *Transportation Research Board. Transportation Research Record,* vol. 1848, pp. 19-28, 2003.

[78] D. Browne and L. Ryan, "Comparative analysis of evaluation techniques for transport policies," *Environmental Impact Assessment Review Journal,* vol. 31, pp. 226-233, 2011.

[79] S. Zanakis, A. Solomon, N. Wishart and S. Dublish, "Multi-attribute decision making: A simulation comparison of select methods," *European Journal of Operational Research,* vol. 107, pp. 507-529, 1998.

[80] M. L. Galves, "Supporting Decision Processes Related to Transport: from Cost-Benefit Analysis to Multiple Criteria Decision Analysis," in *The Association for European Transport. Proceedings of European Transport Conference*, Strasbourg, France, 2005.

[81] A. L. Bristow and J. Nellthorp, "Transport project appraisal in the European Union," *Transport Policy,* vol. 7, pp. 51-60, 2000.

[82] CUPID, *Evaluation Plan. CUPID Project: Co-ordinating Urban Pricing Integrated Demonstrations. Deliverable 4,* 2001.

[83] T. Litman, "Well Measured: Developing Indicators for Sustainable and Livable Transport Planning," Victoria Transport Policy Institute, 2013.

[84] J. Ensor, *Multi-Criteria Analysis: An Alternative Approach for the Evaluation of Road Pricing Strategies. MSc Dissertation,* Massachusetts Institute of Technology, 2003.

[85] "KonSULT- the Knowledgebase on Sustainable Urban Land use and Transport. Road User Charging," [Online]. Available: http://www.konsult.leeds.ac.uk/private/level2/instruments/instrument001/l2_001b.htm. [Accessed 2014].

[86] K. A. Small and J. A. Gomez-Ibañez, "Road Pricing for Congestion Management: The Transition from Theory to Policy," The University of California, Transportation Center. University of California at Berkeley, 1998.

[87] R. Borgo, J. Kehrer, D. H. Chung, E. Maguire, R. S. Laramee, H. Hauser, M. Ward and M. Chen, "Glyph-based Visualization: Foundations, Design Guidelines, Techniques and Applications," in *Proceedings of Eurographics*, 2012.

[88] M. Ward, "Multivariate Data Glyphs: Principles and Practice," *Handbook of Data Visualization,* pp. 179-198, 2008.

[89] L. Wilkinson, "An Experimental Evaluation of Multivariate Graphical Point Representations," in *Proceedings Human Factors in Computing Systems (CHI)*, 1982.

[90] I. Borg and T. Staufenbiel, "Performance of Snow Flakes, Suns, and Factorial Suns in the Graphical Representation of Multivariate Data," *Multivariate Behavioral Research,* vol. 27, no. 1, pp. 43-55, 1992.

[91] M. Lee, R. Reilly and M. Butavicius, "An Empirical Evaluation of Chernoff Faces, Star Glyphs, and Spatial Visualizations for Binary Data," in *Proceedings of the Asia-Pacific Symposium on Information Visualisation (APVis)*, 2003.

[92] A. Klippel, F. Hardisty and C. Weaver, "Star Plots: How Shape Characteristics Influence Classification Tasks," *Cartography and Geographic Information Science,* vol. 36, no. 2, pp. 149-163, 2009.

[93] D. E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning, Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc, 1989.

[94] C. M. Fonseca and P. J. Fleming, "Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization," 1993.

[95] N. Srinivas and K. Deb, "Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms," *Evol. Comput,* vol. 2, pp. 221-248, 1994.

[96] K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.,* vol. 6, no. 2, pp. 182-197, 2002.

[97] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach," *IEEE Trans. Evol. Comput.,* vol. 4, no. 4, pp. 257-271, 1999.

[98] J. Knowles and D. Corne, "The Pareto archived evolution strategy: a new baseline algorithm for Pareto multiobjective optimisation," in *Proceedings of the 1999 Congress on Evolutionary Computation, 1999. CEC 99*, 1999.

[99] D. W. Corne, J. D. Knowles and M. J. Oates, "The Pareto Envelope-based Selection Algorithm for Multiobjective Optimization," in *Proceedings of the Parallel Problem Solving from Nature VI Conference*, 2000.

[100] J. D. Knowles and D. W. Corne, "Approximating the nondominated front using the Pareto Archived Evolution Strategy," *Evol. Comput.,* vol. 8, no. 2, pp. 149-172, 2000.

[101] G. W. Greenwood, X. Hu and J. G. D'Ambrosio, "Fitness Functions for Multiple Objective Optimization Problems: Combining Preferences with Pareto Rankings," in *Proceedings of the 4th Workshop on Foundations of Genetic Algorithms*, San Diego, CA, USA, 1996.

[102] J. Horn, N. Nafpliotis and D. E. Goldberg, "A Niched Pareto Genetic Algorithm for Multiobjective Optimization," in *Proceedings of the First IEEE Conference on Evolutionary Computation, IEEE World Congress on Computational Intelligence*, 1994.

[103] K. Leyton-Brown and Y. Shoham, "Essentials of Game Theory: A Concise Multidisciplinary Introduction," *Synth. Lect. Artif. Intell. Mach. Learn.,* vol. 2, no. 1, pp. 1-88, 2008.

[104] M. O. Jackson, "A Brief Introduction to the Basics of Game Theory," Social Science Research Network, Rochester, NY, SSRN Scholarly Paper ID 1968579, 2011.

[105] L. S. Shapley, "Stochastic Games," *Proc. Natl. Acad. Sci,* vol. 39, no. 10, pp. 1095-1100, 1953.

[106 R. Gibbons, A Primer in Game Theory, Princeton University Press, 1992.

]

[107] L. von Ahn, "Games with a Purpose," *Computer,* vol. 39, no. 6, pp. 92-94, 2006.

[108] L. von Ahn and L. Dabbish, "Labeling Images with a Computer Game," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2004.

[109] L. von Ahn, R. Liu and M. Blum, "Peekaboom: A Game for Locating Objects in Images," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2006.

[110] v. AhnL, M. Kedia and M. Blum, "Verbosity: A Game for Collecting Common-sense Facts," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2006.

[111] E. Law, L. V. Ahn, R. B. Dannenberg and M. Crawford, "Tagatune: A Game for Music and Sound Annotation," in *Proceedings of the 8th International Conference on Music Information Retrieval*, Vienna, Austria, 2007.

[112] L. B. Chilton, C. T. Sims, N. Goldman, G. Little and R. C. Miller, "Seaweed: A Web Application for Designing Economic Games," in *Proceedings of the ACM SIGKDD Workshop on Human Computation*, New York, NY, USA, 2009.

[113] R. McCall, N. Louveton, M. Kracheel, T. Avanesov and V. Koenig, "Gamification as a Methodology: A Multi-Part Development Process," 2013. [Online]. Available: http://orbilu.uni.lu/handle/10993/5658.

[114] L. von Ahn and L. Dabbish, "Designing Games with a Purpose," *Commun ACM,* vol. 51, no. 8, pp. 58-67, 2008.

[115] R. Bartle, *Players who suit MUDs.*

[116] W. Hao and S. Chuen-Tsai, *Game reward systems: Gaming experiences and social meanings,* 2011.

[117] J. P. Gee, "What video games have to teach us about learning and literacy," *Comput. Entertain,* vol. 1, no. 1, p. 20, 2003.

[118] W. I. Zangwill and P. B. Kantor, "Toward a Theory of Continuous Improvement and the Learning Curve," *Manag. Sci.,* vol. 44, no. 7, pp. 910-920, 1998.

[119] N. Hallford and J. Hallford, *Swords and Circuitry: A Designer’s guide to computer role playing games,* Roseville, CA: Prime Publishing, 2001.

[120] L. von Ahn, B. Maurer, C. McMillen, D. Abraham and M. Blum, "reCAPTCHA: Human-Based Character Recognition via Web Security Measures," *Science,* vol. 321, no. 5895, pp. 1465-1468, 2008.

[121] *"CO2G", mobile application.*

[122] *IBM City One Game.*

[123] R. McCall, V. Koenig and M. Kracheel, "Using Gamification and Metaphor to Design a Mobility Platform for Commuters," *Int. J. Mob. Hum. Comput. Interact.,* vol. 5, no. 1, pp. 1-15, 2013.

[124] *“Simcity EDU.”.*

[125] *“intelen BIG.”.*

[126] B. A. McCarl and T. H. Spreen, "Price Endogenous Mathematical Programming as a Tool for Sector Analysis," *American Journal of Agricultural Economics,* vol. 62, pp. 87-102, 1980.

[127] R. H. Day, "On Aggregating Linear Programming Models of Production," *Journal of Farm Economics,* vol. 45, no. 4, pp. 797-813, 1963.

[128] B. A. McCarl and T. H. Spreen, "Applied mathematical programming using algebraic systems," Department of Agricultural Economics, Texas A&M University, 2004.

[129] T. Takayama and G. G. Judge, "Spatial and temporal price allocation models," 1973.

[130] S. Chen, D. Amid, O. M. Shir, L. Limonad, D. Boaz, A. Anaby-Tavor and T. Schreck, "Self-Organizing Maps for Multi-Objective Pareto Frontiers," in *IEEE Pacific Visualization Symposium*, Sydney, 2013.

[131] M. Taieb-Maimon, L. Limonad, D. Amid, D. Boaz and A. Anaby-Tavor, "Evaluating Multivariate Visualizations as Multi-Objective Decision Aids," in *Human-Computer Interaction – INTERACT 2013*, Cape Town, 2013.

[132] O. Shir, D. Moor, S. Chen, D. Amid, D. Boaz and A. Anaby-Tavor, "Pareto Optimization and Tradeoff Analysis Applied to Meta-Learning of Multiple Simulation Criteria," in *Winter Simulation Conference*, Washington, 2013.

[133 O. M. Shir, S. Chen, D. Amid, O. Margalit, M. Masin, A. Anaby-Tavor and D. Boaz, "Pareto Landscapes Analyses via Graph-Based Modeling for Interactive Decision-

] Making," in *EVOLVE - A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation V, Advances in Intelligent Systems and Computing*, Springer, 2014, pp. 97-113.

[134] C. Ware, Information Visualization: Perception for Design, Waltham: Morgan Kaufmann, 2012.

[135] F. Fischer, J. Fuchs and F. Mansmann, "ClockMap: Enhancing Circular Treemaps with Temporal Glyphs for Time-Series Data," in *Short Paper Proceedings of the Eurographics Conference on Visualization (EuroVis)*, 2012.

[136] C. Kintzel, J. Fuchs and F. Mansmann, "Monitoring Large IP Spaces with Clockview," in *Proceedings Symposium on Visualization for Cyber Security*, 2011.

[137] J. Meyer-Spradow, L. Stegger, C. Doring, T. Ropinski and K. Hinrichs, "Glyph-based SPECT Visualization for the Diagnosis of Coronary Artery Disease," *IEEE Transactions on Visualization and Computer Graphics,* vol. 14, no. 6, pp. 1499-1506, 2008.

[138] OECD Better Life Index, "Your better life index," 2011. [Online]. Available: http://www.oecdbetterlifeindex.org/. [Accessed 2013].

[139] P. A. Legg, D. H. Chung, M. L. Parry, M. W. Jones, R. Long, I. W. Griffiths and M. Chen, "MatchPad: Interactive Glyph-Based Visualization for Real-Time Sports Performance Analysis," *Computer Graphics Forum,* vol. 31, pp. 1255-1264, 2012.

[140] J. Siegel, E. Farrell, R. Goldwyn and H. Friedman, "The Surgical Implications of Physiologic Patterns in Myocardial Infarction Shock," *Surgery,* vol. 72, no. 1, p. 126, 1972.

[141] M. O. Ward, "A Taxonomy for Glyph Placement Strategies for Multidimensional Data Visualization," *Information Visualization,* vol. 1, no. 6, pp. 194-210, 2002.

[142] J. Elder and S. Zucker, "The Effect of Contour Closure on the Rapid Discrimination of Two-Dimensional Shapes," *Vision Research,* vol. 33, no. 7, pp. 981-991, 199.

[143] M. Friendly and A. M. Guerry's, "Moral Statistics of France": Challenges for Multivariable Spatial Analysis," *Statistical Science,* pp. 368-399, 2007.

[144] M. Stefaner, "Wahlland," 2013. [Online]. Available: http://moritz.stefaner.eu/projects/wahlland/. [Accessed 2014].

[145] N. Andrienko and G. Andrienko, Exploratory Analysis of Spatial and Temporal Data, Berlin: Springer Verlag, 2006.

[146] R. M. Pickett and G. G. Grinstein, "Iconographic Displays for Visualizing Multidimensional Data," in *Proceedings of the Conference on Systems, Man, and*

Cybernetics, 1988.

[147] Y. H. Chan, C. Correa and K. L. Ma, "The Generalized Sensitivity Scatterplot," *IEEE Transactions on Visualization and Computer Graphics,* vol. 19, no. 10, pp. 1768-1781, 2013.

[148] P. Garrigan, "The Effect of Contour Closure on Shape Recognition," *Perception,* vol. 41, no. 2, pp. 221-235, 2012.

[149] K. Koffka, Principles of Gestalt Pschology, London: Lund Humphries, 1935.

[150] N. Boukhelifa, A. Bezerianos, T. Isenberg and J. D. Fekete, "Evaluating Sketchiness as a Visual Variable for the Depiction of Qualitative Uncertainty," *IEEE Transactions on Visualization and Computer Graphics,* vol. 18, no. 12, pp. 2769-2778, 2012.

[151] J. Heer and M. Bostock, "Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design," in *In Proceedings Human Factors in Computing Systems (CHI)*, 2010.

[152] S. Kosslyn, Graph Design for the Eye and Mind, New York: Oxford University Press, 2007.

[153] L. Bartram, B. Cheung and M. C. Stone, "The Effect of Colour and Transparency on the Perception of Overlaid Grids," *IEEE Transactions on Visualization and Computer Graphics,* vol. 17, no. 12, pp. 1942-1948, 2011.

[154] G. Von Mayr, Die Gesetzmässigkeit im Gesellschaftsleben, Munich: R. Oldenbourg, 1877.

[155] W. Peng, M. O. Ward and E. A. Rundensteiner, "Clutter Reduction in Multi-Dimensional Data Visualization Using Dimension Reordering," in *Symposium on Information Visualization (InfoVis)*, 2004.

[156] S. Few, "Grid Lines in Graphs are Rarely Useful. In DM Review. Perceptual Edge," 2005. [Online]. Available: http://www.perceptualedge.com/articles/dmreview/grid_lines.pdf.

[157] N. Bueme, B. Naujoks and M. Emmerich, "SMS-EMOA: Multi-objective Selection based on Dominated Hypervolume," *European Journal of Operational Research,* vol. 181, no. 3, pp. 1653-1669, 2007.

[158] C. Coello, G. B. Lamont and D. A. van Veldhuizen, Evolutionary Algorithms for Solving Multi-objective Problems, Berlin: Springer, 2007.

[159] K. Deb, Multi-Objective Optimization Using Evolutionary Algorithms, New York: Wiley, 2001.

[160] D. A. van Veldhuizen and G. B. Lamont, "Multiobjective evolutionary algorithms: Analyzing the state-of-the-art," *Evolutionary Computation,* vol. 8, no. 2, pp. 125-147, 2000.

[161] ERF, "User requirements," Consensus Project, 2014.