Degree Project in the Field of Technology Media Technology and the Main Field of Study Computer Science and Engineering

Second cycle, 30 credits

# Decision-Making and Usability in Power Grid Systems

## A Study of Interaction Design and Data Visualization

**WILLIAM O'GRADY**

# Decision-Making and Usability in Power Grid Systems

## A Study of Interaction Design and Data Visualization

WILLIAM O’GRADY

Master’s Programme, Interactive Media Technology, 120 credits
Date: April 27, 2026

Supervisor: Rémy Rey
Examiner: Tino Weinkauf
School of Electrical Engineering and Computer Science
Swedish title: Beslutstagande och användbarhet i Kraftnätssystem
Swedish subtitle: En studie i interaktionsdesign och datavisualisering

# Decision-Making and Usability in Power Grid Systems

A Study of Interaction Design and Data Visualization

The research and development of data visualizations to support power grid operation faces a number of challenges. Increasing data complexity place high demands on both operators and the visual tools used to achieve their tasks. While new visualization approaches continue to emerge, many are insufficiently evaluated in terms of how they function in practical contexts. This thesis presents a user-centered approach to evaluating power grid data visualizations by investigating how users' decision-making process is shaped and supported by interface layout and interaction design. A prototype simulating a power grid system was developed and controlled through two contrasting interfaces. A mixed methods approach was used by conducting a user study with 16 participants, tasked with solving power redispatching tasks in a constrained network grid. Results found that while user strategy was in large part the same across interfaces, key design elements such as layout complexity, control, and visual mapping supported or stifled the decision-making process of users, and in turn the overall usability of the system.

## SAMMANFATTNING

Utveckling av datavisualiseringar för att stödja övervakning och kontroll av elektriska kraftnät står inför ett antal utmaningar. Den växande datakomplexiteten kräver mer av de verktyg som operatörer använder för att utföra deras arbete. Trots att forskare presenterar nya visualiseringar för att bättre spegla kraftnätet, utvärderas de inte i hur de möter användarnas krav i en praktisk kontext. Den här uppsatsen tar ett användarcentrerat perspektiv i utvärderingen av kraftnätsvisualiseringar, genom att undersöka hur användares beslutsprocesser formas och stöds av gränssnittens layout och interaktionsdesign. En prototyp utvecklades för att simulera ett kraftnätssystem, som kunde styras via två motstående gränssnitt. En användarstudie med både kvantitativa och kvalitativa metoder genomfördes, där 16 deltagare fick i uppgift att lösa omdirigeringsuppgifter i en begränsad nätstruktur. Resultaten visar att användarnas övergripande strategier inte skiljde sig mellan gränssnitten, men att centrala designelement såsom layout, kontroll, och koder stödde eller försvårade användarnas beslutsfattande, vilket i sin tur påverkade systemets övergripande användbarhet.

CCS Concepts: • **Human-centered computing** → **Visualization**; *User studies*; • **Applied computing** → Power and energy systems.

Keywords: data visualization, power grid systems, user interfaces, decision-making, interaction design, user-centered design, human-computer interaction

Nyckelord: datavisualisering, kraftnätssystem, användargränssnitt, beslutstagande, interaktionsdesign, användar-baserad design, människa-datorinteraktion

---

Author's address: , ogrady@kth.se, KTH Royal Institute of Technology, School of Electrical Engineering and Computer Science, Stockholm, Sweden, SE 100 44.

---

Document date: April 27, 2026
Swedish title: Beslutstagande och användbarhet i Kraftnätssystem
Swedish subtitle: En studie i interaktionsdesign och datavisualisering

## 1 INTRODUCTION

### 1.1 Background

The *power grid* is an interconnected electrical network that transmits electricity from generators to consumers. It is a highly complex system composed of transmission lines, substations, power plants, and control equipment which, when stable, can reliably meet electricity demand, even across national borders. Ensuring this stability is the responsibility of grid operators, who rely on data visualizations and digital control tools to interpret the large volumes of numerical data produced by the system. Data visualizations are a fundamental part of grid operation in supporting time-critical operational decisions, preventing failures like blackouts. These visualizations also serve as shared communication tools, enabling operators to coordinate with one another and with external stakeholders in order to better plan and forecast future threats to grid stability [23].

Power grid visualizations are driven by data collected from sensors distributed across the power grid. Advances in sensor technology have resulted in a dramatic increase in the volume, resolution, and availability of grid data. In today's smart grid environments, where system-wide monitoring and control are standard, visualization design has become a central challenge in ensuring that operators can effectively interpret and use this data in their tasks [6] [26] [27]. As Zheng et al. note, "how to intuitively display [the information] to power dispatching personnel has become an important problem." [33]

Researchers describe this challenge as achieving a higher level of *situational awareness* [10] [12] [14] [22]. Following Endsley's definition, this refers to a cognitive model of user perception of the system state [8]. Through this lens, data visualizations have evolved from simple charts and bar graphs into complex geographical, multivariate, and spatio-temporal representations, ranging from large overview displays to specialized control systems. These are seen today in commercial programs, such as GIS and SCADA systems. Data visualizations have allowed grid operators to not only maintain the stability and demands of the power grid but to enhance its potential [16] [23] [26].

### 1.2 Problem Statement

Survey literature indicates a persistent lack of evaluative methods used to verify newly proposed visualization approaches. Fischer and Keim's 2022 survey notes that "especially for new techniques, the approaches were untested." When evaluative methods are employed, they rely on weak or insufficient methods, such as exploratory or illustrative tasks, that neither reflect the operational tasks of grid operators nor meaningfully challenge the usability of the proposed systems [9] [30] [31].

A recurring assumption in the literature is that increasing data complexity primarily constitutes a technical challenge. In *Big Data and Monitoring the Grid*, Phan and Chen argue that the growing volumes of smart grid data "necessitate new architectures, algorithms, and procedures." [23] Previous approaches overlook the user's cognitive relationship to the data visualizations. Before introducing increasingly sophisticated visual representations, it is necessary to understand the limits, strategies, and needs of the user.

Taken together, these tendencies have led to a research landscape that has been described as increasingly stagnant or underdeveloped [9] [16] [26]. This study addresses the identified gap through a scenario-based evaluation in which users are required to solve constrained redispatching problems, where they make trade-offs and operate under pressure, thereby challenging how the interface supports user reasoning rather than merely how efficiently isolated tasks are performed. This is examined by comparing two contrasting designs: a list-based interface reflecting interaction patterns commonly found in conventional control systems, and a map-based interface, representing a more exploratory approach motivated by previous research and interaction design principles. A mixed-methods approach is employed, combining quantitative task performance measures with a qualitative thematic analysis of interview data, to examine users' strategies, reasoning, and experiences. Together, these methods are used to assess how interface design choices shape system decision-making support and usability as a whole.

### 1.3 Research Questions

This study addresses the following research questions:

(1) **RQ1:** How does the choice of interface layout (map-based vs list-based) influence users' decision-making process when solving power-grid control tasks?
(2) **RQ2:** What usability challenges and affordances do users experience when interacting with a map-based versus a list-based power-grid control interface?

### 1.4 Delimitations

A key delimitation of this work is that the prototype is not intended to be a high-fidelity representation of an existing power grid control system, nor to be directly compared with real-world operational tools. Instead, the prototype is designed as a research instrument, informed by prior work and real-world practices, best answering the research questions. The abstractions and reductions applied to the grid topology are therefore intentional and serve the research purpose rather than operational realism. This is discussed further in Chapter 3. In addition, the study is limited by a small participant sample and a restricted interaction time of 45–60 minutes per participant. As a result, the findings should be interpreted as exploratory and indicative rather than conclusive.

### 1.5 Structure of the Thesis

The report begins with Chapter 2, which summarises how power grid data has been visualized, as well as identifying key gaps in current research. Chapter 3 describes the prototype developed for this thesis, including its design rationale, a walkthrough of the two interfaces, and methods employed for evaluation purposes. Chapter 4 outlines the evaluation methodology. Chapter 5 presents the quantitative and qualitative results of the evaluation. Chapter 6 discusses these results in relation to the research questions, as well as the limitations of the evaluation. Finally, the report concludes by outlining opportunities for future research in the domain.

## 2 RELATED WORK

### 2.1 Visualization of Power Grid Data

Power grid data is at its core numerical, describing quantities such as power generation and consumption, voltage levels at buses, and power flows on transmission lines. Each value represents a local condition, while the behavior of the grid emerges from the relationships between values across the network. The system therefore cannot be understood by reading numbers in isolation. In the 1990s, early work by researchers Mahadev and Christie questioned whether numerical representations alone were sufficient for understanding power system data, asking "when and why graphics improve understanding, and when numbers are better." [13] In a separate study, Becker focused on the growth of power system data itself, observing that "tools for analyzing network data have not kept pace with the data volumes." [2]

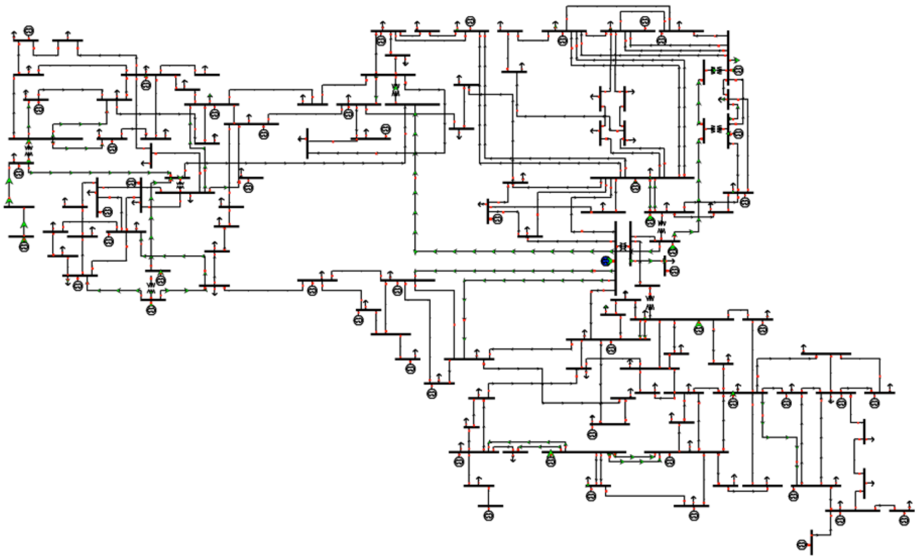


**Figure 1: Single-line diagram of the IEEE 118-bus test system. Source: Illinois Center for a Smarter Electric Grid (ICSEG) [11].**

Visualization became a way to express power system data in terms of patterns, trends, and outliers, rather than isolated values. It also enabled relationships between multiple parameters to be examined simultaneously, mitigating the complexity of large, interconnected systems. As visualization became established in power system analysis, a wide range of techniques were explored to cope with increasing data complexity. These include low-dimensional views for individual parameters (charts, graphs, or diagrams), higher-dimensional visualizations for multivariate relationships (scatter plots, Andrew plots, and parallel coordinates), and three-dimensional approaches intended to integrate multiple data layers into a single representation [15] [26]. Notably, works by Overbye and Weber developed a range of two-dimensional techniques, and also considered a broader set of users, including planners, analysts, and market actors such as power brokers, when designing visual representations of power system data [18] [19] [20] [21] [22] [26] [28].

Visualization was used not only to represent data values, but also to represent the grid itself as a network structure. This led to a fundamental distinction between geographical and topological representations. Geographical views preserve the physical locations of each network element, while topological views abstract away geography to emphasize connectivity and electrical relationships in diagram form, most clearly expressed through single-line diagrams (Figure 1) [12] [16].

Development of data visualizations has been considered through cognitive theory. Current research, such as Cheng et. al. 's “GridViz”, is motivated by strengthening *situational awareness* for the user [7]. Situational awareness was proposed by M. R. Endsley in 1995 as a three-level cognitive model. From lowest to highest, the levels are: perception of elements in the environment, awareness of the current state, and prediction of future states [8] Research has also framed the intuitiveness of power grid visualizations in terms of mapping quality. This concerns how naturally different system properties are translated into visual form, such as whether power flow is most intuitive as encoded by colour, or which chart types best convey quantities like phase angle or frequency [14].

In practice, these visualization techniques are used to provide users with an overview of system state, support monitoring tasks, and enable reasoning about grid behavior under changing conditions. Abstract topological views, multivariate encodings, and overview displays are commonly integrated into operational tools that allow users to assess system state, anticipate potential issues, and make informed decisions. Such visualizations are widely deployed in industry-standard systems, including commercial simulation environments such as PowerWorld Corporation's PowerWorld Simulator, as well as in control room platforms based on SCADA and Advanced Metering Infrastructure (AMI) [23]. Geographic information systems (GIS) are also frequently used to visualize grid assets and contextual data, particularly for planning and asset management tasks.

## 2.2 Current Research

Survey literature provides a consistent diagnosis of the current state of power grid visualization research. Three recurring issues are highlighted: a lack of innovation in visualization techniques, a lack of evaluation, and the growing challenge of representing ever-expanding datasets in the transition toward smart grids. [9] [16] [26].

Several studies have proposed innovative visualization tools accompanied by evaluation. Systems such as GreenGrid and GreenCurve introduce alternative representations and assess their impact using quantitative performance measures: task completion time, solution accuracy, and satisfaction [30] [31]. Similarly, work by Daniel A. Keim compares topological and geographical views and evaluates how these affect user performance (in tasks such as identifying parts of a network affected by an outage). Importantly, Keim's analysis moves beyond raw performance metrics by drawing qualitative insights about user behavior, such as identifying "the smart filtering concept is not useful for the task" or when "the focus of the users is on the topology view." [12] These findings demonstrate the value of examining how users engage with a visualization, rather than only how fast or how accurately they complete predefined tasks. However, such insights remain grounded in exploratory and comprehension-focused tasks, leaving open questions about how visualization design supports decision-making in constraint-driven control scenarios.

A more thorough evaluative perspective is offered by Mikkelsen et al., whose work systematically examines visual mapping, identifying a range of design-related challenges within single-line diagram overview displays, including perceptual issues such as visual clutter, occlusion, scaling, and color overload, as well as cognitive challenges related to mismatches between visual encodings and user expectations. These factors directly affect users' ability to maintain situational awareness and make sense of system state in complex grid overviews. At the same time, questions of how users actively engage with the system remain outside the scope of the analysis. They conclude, "The only proposed solution for this problem is to animate historical values… since it would require interaction, it is not a good solution for an overview display." [14]

Interaction has nevertheless long been recognized as a fundamental component of effective visualization [9] [32]. Overbye explicitly emphasized the importance of interaction in early work, stating that "to be truly effective, computer visualization must be interactive and must be fast." [20] However, interaction remains underevaluated in power grid research in how it relates to the user's situational awareness. This motivates the use of interaction design as a lens for shaping and evaluating power grid control interfaces, focusing on how layout and functionality are structured around user behaviour.

## 2.3 Purpose of This Study

Findings from prior work point toward a common limitation: development has been focused on intuitive mapping rather than user behaviour. Previous studies have relied heavily on quantitative performance measures collected in low-stakes or exploratory scenarios, not putting enough strain on the functionality and the visualizations in relation to how users make use of them. As Burnay et al. note, "producing a graph or a map and plotting data on it is one thing, but designing visualizations that actually support decision-makers is something completely different." [4] This study seeks to widen evaluative approaches by investigating how users' decision-making processes relate to the system's interface layout. To fill the identified gap, a mixed approach analysis is used to capture how users interact and experience the system while solving constrained control tasks.

# 3 DESIGN

This chapter describes the design and development of the *Power Grid System* prototype, including its two interfaces—*Map View* and *List View*—the mapping and application of power grid data, and the interaction design choices that shaped the system. It also outlines how the prototype was instrumented and structured to support the user evaluation study.

## 3.1 Grid Logic

The prototype is a steady-state simulation of the IEEE-118 power grid, a standardized benchmark model used in power-system research, that allows users to solve scenario-based tasks by manipulating generator output [1]. Each scenario requires users to meet power demand and maintain line stability under a time limit, while balancing trade-off parameters, cost and emissions, which affect their score. The two interfaces provide fundamentally different modes of control: Map View uses the grid topology itself as the primary interaction surface, while List View separates control from the map, using it instead as a visual reference.

The prototype is a web-based JavaScript/HTML application, using D3.js for visualization, and a Flask backend. Grid data was derived from PYPOWER case118 and the IEEE-118 operating manual [1].

The system was designed from a set of design goals as constraints of the evaluation study. The most important goal was for the prototype to be understandable and usable by participants with no prior power-grid expertise within the one-hour study time constraint. To achieve this, the simulation deliberately abstracts away many aspects of real grid operation, including ramp times, detailed cost models, frequency control, component failures, environmental effects, and advanced controllable elements such as shunt devices. Only core components—generators, buses, transmission lines, and loads—are retained.

Loads impose a fixed power demand, while generators inject active power at specific buses up to their maximum capacities. Whenever generator outputs are changed, the backend recomputes the system state by running a power-flow calculation. One generator is assigned as the slack generator to absorb any remaining imbalance so that total generation and total demand remain equal. This is needed to compute power flow, but does not equate to the system being operationally solvable. Power is then distributed across the network according to the grid topology and generator locations.

Each transmission line has a fixed capacity limit. After each power-flow calculation, the loading of every line is computed as a percentage of this limit. A line is considered overloaded when its loading exceeds 100%, meaning that the power flowing through it is greater than its rated capacity. Overloads are not automatically prevented or corrected by the system. Instead, they are mitigated indirectly by changing generator outputs: reducing generation near stressed areas, increasing generation elsewhere, or redistributing generation so that power flows take alternative paths through the network. When these adjustments reduce a line's loading below 100%, the overload condition is resolved.

### 3.2 Topology Mapping

Topology mapping involved deciding which grid elements were essential. Once core elements were identified, the challenge was to present them so that each component was distinct and easy to recognize. The design was informed by established principles of visual variables as described by Roth, as well as the findings of Mikkelsen et. al. [14] [25].

The grid topology was adapted from an existing image and recreated as an SVG in Figma. This allowed the map to be imported into the system and manipulated through code.

Colour coding was the main visual mapping technique. Green indicated active, while white and gray indicated inactive. Line load was represented using a white-to-yellow-to-orange-to-red colour scale applied to chevrons. Cost used the same colour scale, while emissions were represented using a hatch-pattern density scale. Generators were represented as rectangles. Output loads were shown as blue house icons. The colour blue was also used as a selection outline for overview cards and generator cards. Transmission lines and buses were drawn as thin gray rectangles.

Locked and disabled generators were introduced to impose meaningful constraints while maintaining usability in the prototype. In real power-grid operation, operators typically have limited control over generation assets; however, fully replicating this restriction would have reduced users' ability to explore the system and solve tasks within the study context. To balance realism and learnability, most scenarios allow users to adjust approximately 60–90% of generators.

The resulting prototype consists of two interfaces, a scenario and scoring system, a user test structure, and a tutorial slide deck.

### 3.3 *List View* Interface Layout

The List View interface, as shown in Figure 2, is designed to mirror the philosophy of current power-grid control tools used by operators. This meant separating overviewing grid topology from manipulating the data. Consequently, the interface emphasizes displaying precise numerical values and de-emphasizing the spatial grid map. A notable inspiration for the List View was Microsoft Excel, reflecting its focus on tabular data.

The final List View interface can be divided into four main elements. From top to bottom and left to right, these are: the overview bar, the inactive generator list, the active generator list, and the grid map. The user manipulates the system state by adding, removing and toggling the power of individual generators through their corresponding generator card in the lists, while using the grid map as a reference for their location in the topology.

*3.3.1 Overview Bar.* The overview bar is divided into two main sections: system state, shown on the left, and scenario requirements, referred to as the task section, shown on the right.

*System State Cards.* The System state is presented using a set of cards, each corresponding to a different parameter. From left to right, these represent power, line stability, cost, and emissions, respectively. The cards are intentionally ordered and sized to reflect the priority list of current power grid goals, *Safest* → *Cheapest* → *Cleanest*; as stated in Svenska Kraftnät's plan for 2024-2033: "The aim of operational development is to find a balance between the technical status of infrastructure, the potential for outages, safety, environmental impact and cost." [29] Power and line stability are given equal prominence, followed by cost and emissions, with cost placed visually above emissions. This ordering was a deliberate design choice intended to guide user decision-making.

Each card includes a header describing the displayed metric. The Total Power card shows the ratio of current total generation to target demand (in MW). This is presented both numerically and in an abstracted form using the same green and gray (on and off) 100 MW "power blocks" associated with individual generators.

The Overloaded Lines card displays the number of currently overloaded lines in the system. Below this, two rows of smaller cards represent individual transmission lines. The top row displays up to three overloaded lines, while the bottom row shows up to three of the most heavily loaded lines when no overloads are present. If a line's load reaches or exceeds 100%, it is promoted to the top row, whose cards are larger to emphasize critical conditions.

Each line card in turn displays the current load percentage, a small chevron icon indicating load direction (as described in the Grid Topology Mapping subsection), and a delta value that updates after each system change. For example, a red "+4" indicates that the most recent action increased the line's load by four percentage points, bringing it closer to overload. Line cards are interactive: by triggering a brief scaling animation that highlights the corresponding line chevron on the grid map.

The Total Cost and Total Emissions cards follow the same visual logic. Each displays the ratio of current value to the scenario budget (normalized per MW), along with a colored pill indicator showing whether the user is under budget, near the limit, or over budget. These states are represented using green, yellow, and red colors, respectively.

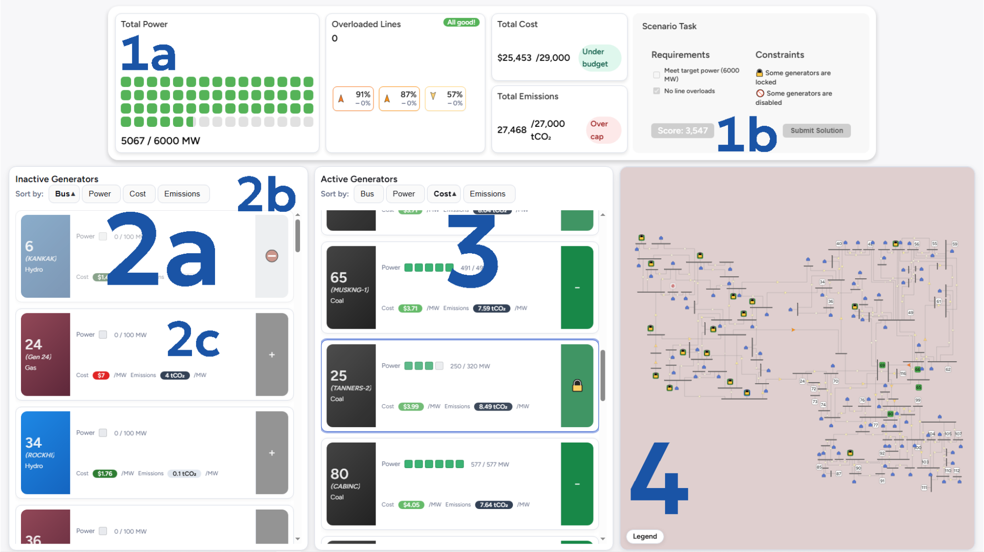


**Figure 2: List View interface. The top section shows the Overview Bar with the system state (1a) and the task section (1b). The lower section contains the inactive generators list (2a) with sorting controls (2b) and generator cards (2c), the active generators list (3), and the map element (4).**

*Task Section.* The task section displays the requirements and constraints specific to the current scenario. It also shows the user's current score and the submit button. Scenario requirements are presented as a checklist; as each requirement is fulfilled, it is marked as completed. When all requirements are satisfied, both the score indicator and the submit button turn from gray to green, signaling that the scenario is ready to be submitted.

*3.3.2 Inactive and Active Generators List.* The inactive and active generator lists separate generators that are offline and online, respectively. This separation provides users with a clear overview of the initial system configuration for each scenario. It also allows the List View to accommodate a higher density of interface elements than the paired-down Map View. From a design perspective, the increased complexity of the List View was intentional rather than a consequence of poor usability. The interface was designed to reflect the interaction philosophy of existing power grid control systems, in which a significant portion of operator activity involves navigating menus, lists, and parameter panels rather than directly interacting with the grid topology.

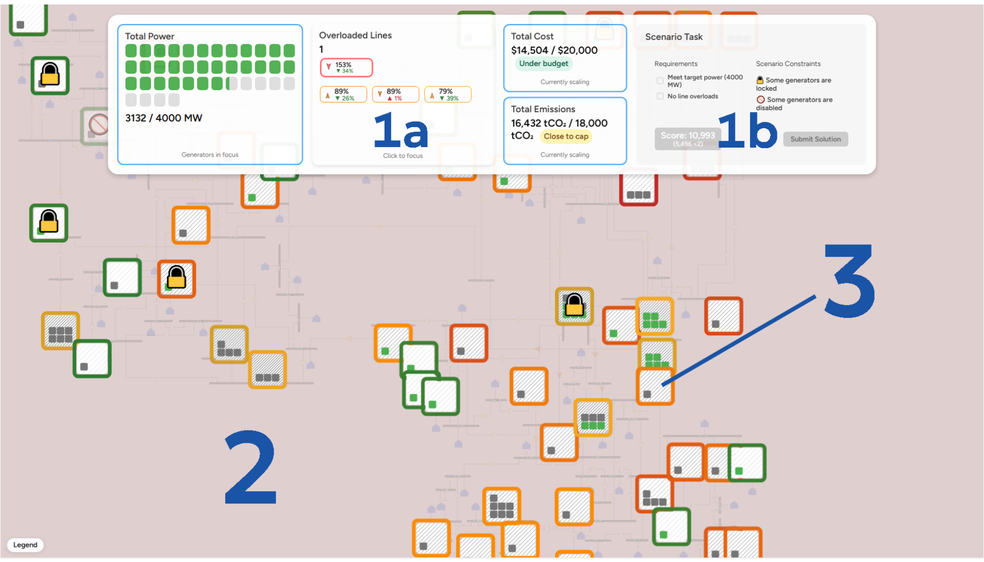
The inactive and active generator lists share the same layout and interaction logic. Both lists provide sorting controls that allow generators to be ordered by bus number, power, cost, or emissions, in either ascending or descending order. In both cases, sorting by power refers to a generator's maximum capacity rather than its current output. Multi-parameter sorting and sorting by current generation were not implemented, as these features were deprioritized during development in favor of core interaction functionality.

*3.3.3 Generator Cards.* Each list consists of generator cards, with each card representing a single generator. The card is visually divided into three sections: a left, middle, and right section. Selecting a generator card will cause the card and corresponding generator element on the grid map to be highlighted. The left section displays the generator's associated bus number, name, and fuel type, set against a background color corresponding to that fuel type (coal are shown as dark gray, combined cycle are teal; hydro plants are blue, and gas are burgundy). The bus number serves as the default sorting key and functions as an indexical identifier. It is also used to locate the generator within the grid map.

The middle section presents the generator's power, cost, and emissions information. These values are arranged in two rows, with the top row displaying a series of power blocks alongside a numerical fraction indicating current generation relative to maximum capacity. For example, a generator with a maximum capacity of 300 MW is represented by three power blocks. When the generator is inactive, all blocks appear gray. The number of power blocks is computed by dividing the generator's maximum capacity by 100 and rounding the result up to the nearest integer.

Users can activate a generator by selecting one or more power blocks, which immediately sets the generator's output to the corresponding level, moves the card to the active generator list, and updates the selected blocks to green. The same interaction is used to adjust the output of generators that are already active, allowing users to increase or decrease generation in discrete steps.

The right section displays a symbol indicating the generator's status. Inactive generators show a plus symbol; selecting this section activates the generator at full capacity and moves it to the active list. Active generators display a minus symbol, which deactivates the generator when selected. Generators that are locked or disabled display corresponding symbols and cannot be toggled by the user; locked generators cannot be turned off, and disabled generators cannot be turned on.



Total Power
3132 / 4000 MW
Generators in focus

Overloaded Lines
1
153% ▼ 34%
89% ▼ 26%
89% ▲ 1%
79% ▼ 39%
Click to focus

Total Cost
$14,504 / $20,000
Under budget
Currently scaling

Total Emissions
16,432 tCO₂ / 18,000 tCO₂
Close to cap
Currently scaling

Scenario Task
Requirements
☐ Meet target power (4000 MW)
☐ No line overloads
Scenario Constraints
Some generators are locked
Some generators are disabled
Score: 10,993 (5,496 ×2)
Submit Solution

**Figure 3: Map View interface. The top section shows the Overview Bar containing the system state and filter control (1a) and the task section (1b). The lower section presents the map element (2), which includes interactive generators (3).**

*3.3.4 Grid Map.* The grid map element is used exclusively for navigation and visual reference. Navigation is limited to panning and zooming, with interaction logic borrowed from common web mapping tools such as Google Maps. This design choice was made to align with established user expectations and to avoid introducing unnecessary interaction complexity. The interactive elements within the grid map are the generators and line load chevrons. Selecting a generator highlights the corresponding generator card in the active or inactive list. A legend is provided to explain the visual encoding of grid map elements. Generators are labeled by their associated bus number. If a generator is locked, this label is replaced with a lock symbol; if a generator is disabled, the label is replaced with a disabled symbol and the generator is rendered with reduced opacity. Hovering over a transmission line chevron reveals a tooltip displaying the line's identifier (e.g., `Line_Bus5_Bus6`) and its current load percentage.

### 3.4 *Map View* Interface Layout

The Map View interface, as shown in Figure 3, was developed in response to findings reported by Mikkelsen et al. in their evaluative study of overview displays. These findings emphasize, among other principles, the avoidance of explicit numerical values, keeping the focus as "clean as possible", and the mapping of system parameters to strong visual patterns. For the design, this resulted in omitting everything but the map and overview bar elements [14].

To maintain consistency and reduce unnecessary cognitive overhead, the Map View intentionally shares the overview bar and underlying grid map navigation logic with List View. However, their role and functionality differ in key ways, as described below. The resulting interface allows users to manipulate the system state by directly interacting with the grid map's generator elements.

The overview bar in Map View retains the same structure and visual encodings as in List View but introduces additional interactivity. This was motivated by the interactive, user-friendly power plant map introduced by Rodrigues et al. in their 2017 paper [24]. Each card can be toggled on or off, directly influencing how information is mapped onto the grid map. Activating the Power card focuses the view on generator elements by increasing their size and visually de-emphasizing the surrounding grid, reducing the need for zooming when assessing generator capacities. Selecting the Total Cost and Total Emissions cards maps these parameters onto generator elements. Cost is encoded using a green–yellow–orange–red color scale, representing cheap to expensive generation, while emissions are represented using an internal diagonal hatch pattern, where increasing density indicates higher $CO_2$ emissions per MW.

The goal was to abstract numerical data away from the grid map and instead represent all relevant parameters through visual metaphors directly embedded in each generator element. This approach was intended to keep the user's attention entirely on the spatial representation of the grid, without requiring auxiliary views or additional panels. Including the overview bar therefore represents a deliberate compromise within this design philosophy. While the Map View emphasizes local, spatial reasoning, the overview bar provides a persistent global reference. Regardless of zoom level or map position, it anchors the user's actions to overall system goals and constraints.

Map View inherits the 100 MW "power block" logic from List View and integrates it directly into the generator elements on the grid map, using a dice dot metaphor. Generator output is adjusted by toggling these blocks from the bottom row upward. Deactivating a generator is done by selecting the first block (the bottom-left block) when the generator is already active.

### 3.5 Scenario and Scoring Design

*Scenarios* are the format in which problem-solving tasks are presented to the user. Each scenario sets the parameters and conditions in which certain requirements must be met by the user. They were developed for two reasons: to simply change the grid's constraints/settings, as well as to create the evaluation tool in order to investigate the research questions. Scenarios were developed so that they could be used regardless of interface.

Scenarios act like "puzzles" rather than real-life cases where users must find the best possible solution. Each scenario sets these constraints: requirements (*meet target power* and *ensure no line overloads*), target values for power, cost and emissions, line sensitivity, locked and disabled generators (a AGC pool of slack generators), and initial outputs. Cost and emissions budget caps were set by running pilot tests and taking the average cost and emissions results of found solutions by users multiplied by 120%.

A scoring system was added to each scenario in order to steer the priorities of the user. This introduces an element of gamification intended to engage participants in their decision-making process. Each scenario awarded one point for every dollar below the cost budget cap, with the score doubled if the emissions budget cap was also met. The score was only unlocked for valid solutions, meaning the user had to satisfy both power and stability constraints in order to receive a score. Achieving higher scores therefore required balancing cost and emissions while maintaining system stability. Together with the available generators, these constraints defined the difficulty of each scenario.

### 3.6 Design Iterations

Iterative design was applied to developing the prototype and user evaluation. These iterations helped refine mapping of system elements and interaction design. The goal of this process was to ensure that the design goals were being met. A key outcome of this process was the inclusion of a scoring system, suggested during pilot testing as a way of deepening the decision-making process. Participants involved in pilot testing were excluded from the subsequent user study to avoid bias from prior exposure to the

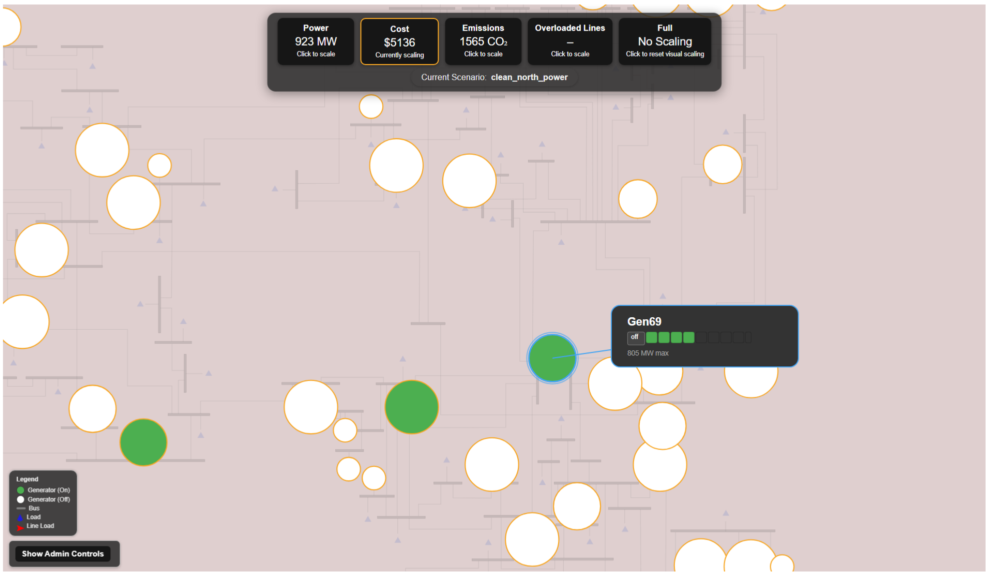


**Figure 4: An Early Version of the Map View Interface.**

system. An early version of the Map View interface is shown in Figure 4, in which generator elements were represented as circular nodes. Generator output was controlled through a single-parameter scaling mechanism, visualised by variations in node radius. Power toggling and configuration were handled through a separate modular control window, rather than being directly integrated into the map-based interaction.

# 4 EVALUATION

## 4.1 Participants

16 participants were recruited for the user study. The participants consisted of fourteen KTH students and two Stockholm University students. None of the participants had any prior experience with power grid systems, approaching both interfaces with equal novelty. As opposed to using experienced grid operators or domain experts, this study creates the opportunity for the operators of the grid system's decision-making process to develop by the nature of the interface design itself.

## 4.2 Procedure

Each user study session was planned to last approximately 50 minutes, comprising a tutorial, a user test, and a post-test survey and interview. In practice, sessions ranged from 45 to 70 minutes depending on participant reflection and interview length.

The tutorial which began the study session was designed to establish baseline familiarity with the system through an introductory script and a small set of exploratory tasks in each interface (Appendix A). Pilot testing showed that asking participants to solve scenarios without this preparation was too demanding. Before the tutorial, participants signed a consent form and audio recording began. Participants were encouraged to use the *think-aloud method* throughout the session and were free to speak Swedish, their native language [5]. The user test consisted of four scenario tasks, each limited to six minutes, with an optional six-minute break between scenarios.

After the scenarios, participants completed a Google Forms survey (Appendix B), which also served to initiate reflection for the post-test interview. The session concluded with a semi-structured post-test interview, guided by pre-established questions while allowing additional topics to be introduced based on insights from earlier participants (Appendix C).

### 4.3 Setting

The user study was conducted in a MIDDLA Multi Studio space at KTH in Stockholm. Each session involved one participant and the interviewer. The prototype was displayed on a large screen connected to a laptop via HDMI, with participants interacting using a mouse and completing the survey on the laptop, while the interviewer monitored the session, recorded audio, and took notes.

### 4.4 Conditions

The study employed a within-subject design with a balanced Latin square over interface type and scenario. Scenario sequences were predefined and assigned to participants during the user test.

- **Independent variables:**
  - Interface type (*Map View*, *List View*)
  - Scenario (A–D)
- **Dependent variables:**
  - Scenario performance
  - Self-reported experience (Likert-scale items)

### 4.5 Data Collection

Data collection primarily consisted of audio recordings of the full session, captured using a smartphone placed between the participant and the interviewer.

In addition to qualitative data, the study collected quantitative measures for each scenario, including completion status, completion time, score, highest line load percentage, and post-test Likert-scale responses. The use of post-test Likert-scale surveys was informed by prior evaluations of GreenGrid and GreenCurve, where participants rated task satisfaction on a five-point scale [30] [31]. This study extends the scale to seven points and applies it to assess each interface as a whole rather than individual tasks. Rather than serving as primary indicators of performance, the metrics were used to contrast and complement the interview results, enabling a more focused discussion of where measured outcomes aligned with, or diverged from, participants' reported experiences.

Completion time was measured in seconds (0–360), corresponding to the six-minute time limit per scenario. Completion status was binary and dependent on whether the scenario was submitted before timeout. Scenario score reflected cost performance relative to the budget cap, with a binary bonus applied if emissions were also kept under the emissions cap, as described in Chapter 3. The post test survey answers were recorded digitally through Google Forms.

Audio recordings were transcribed and analysed using Braun and Clarke's *Thematic Analysis* method [3]. Transcription was performed using the European Union Council's DIGITAL Language Tools Speech Services' AI transcription tool, followed by manual correction and speaker attribution through repeated listening. Quotes included in the results section were translated from Swedish to English by the author without the use of automated translation tools. Scenario results and post-test survey data were digitized by transferring handwritten notes to Google Sheets.

### 4.6 Data Analysis

Thematic analysis was conducted by reviewing each transcript and highlighting quotes or exchanges relevant to user experience, impressions, and decision-making, while less relevant or purely procedural dialogue was excluded. Highlighted excerpts were transferred to a Miro board as individual notes and iteratively refined to ensure each captured a coherent idea. These notes were then coded following Braun and Clarke's thematic analysis approach, allowing codes to be split, merged, or discarded before being grouped into higher-level themes.

Methods used for quantitative analysis depended on the nature of the data. Quantitative metrics analysed were time to completion, task completion, score, and the Post-Test Survey Results.

## 5 RESULTS

### 5.1 Quantitative Results

*5.1.1 Time Usage.* To examine whether interface layout influenced time allocation, completion time was compared between Map and List conditions. This measure reflects time until the user submitted their solution. Analyses are restricted to completed runs when interpreting time-to-submit behaviour; timeouts were described separately as completion outcomes.

Completion times between interfaces were compared using a Wilcoxon signed-rank test (Figure 5). Participants 1 and 16 only completed scenarios in the Map interface and were thus omitted from this test. The results show that there were two cases (P4 and P11) who submitted their solutions at the last possible moment, thus listed as completing the scenarios with the time 360 seconds.

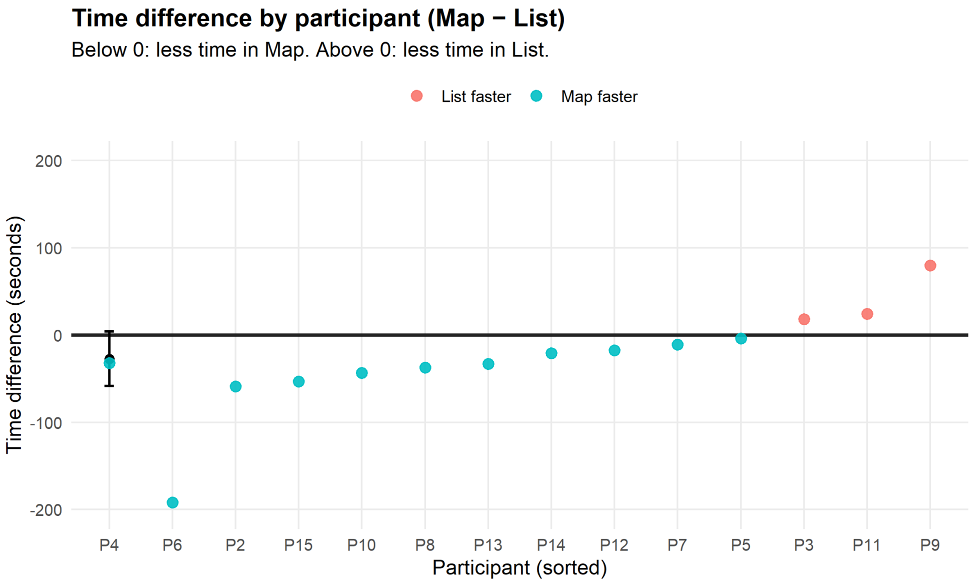


**Figure 5: Time Difference by Participant.**

Participants tended to spend slightly less time before submission in the Map interface than in the List interface. Across completed runs, mean time to submission was 275 seconds in Map and 303 seconds in List, However, this difference did not reach statistical significance when using a Wilcoxon signed-rank test (p = .068).

*5.1.2 Score.* Score was used as an outcome measure to reflect how far participants chose to optimise their solutions within the constraints of each scenario. Because no maximum possible score was calculated and scoring was influenced by scenario-specific conditions and trade-offs, score is not interpreted as an absolute measure of solution quality or correctness. Instead, it is treated as a behavioural indicator of participants' willingness to continue adjusting the system to improve performance under the given constraints.

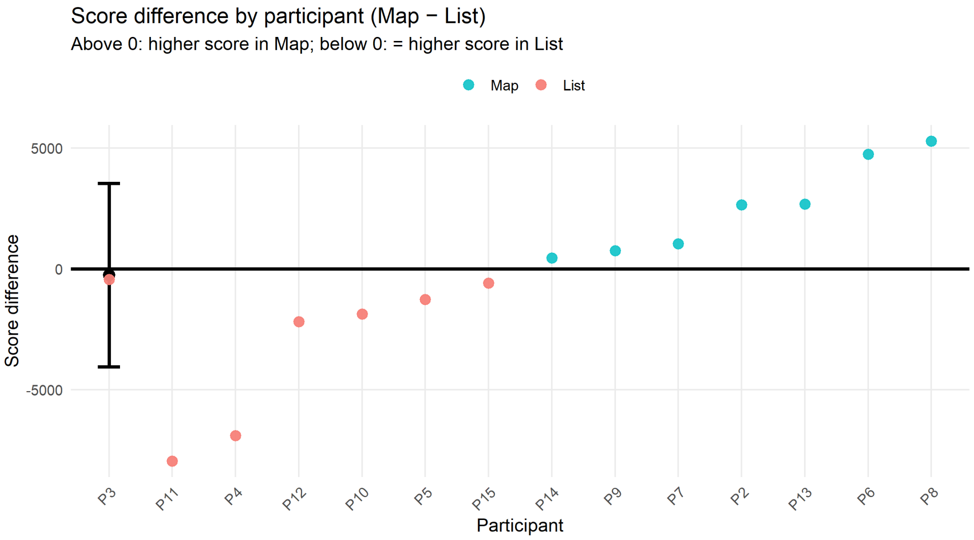


**Figure 6: Score Difference by Participant.**

Scores varied substantially across participants and scenarios (Figure 6). There was only one participant (Participant 10) whose score increased for every subsequent scenario. There were no cases of a participant getting a lower score for each subsequent scenario. When comparing interfaces at the participant level, no systematic difference in score was observed between the Map and List interfaces. A Wilcoxon signed-rank test indicated no significant difference in score between Map and List interfaces, (p = 0.55).

Score differences between interfaces were heterogeneous across participants. Several participants achieved substantially higher scores in one interface than in the other, while others showed minimal differences or the opposite pattern. This lack of a consistent directional effect suggests that interface layout did not reliably influence the degree to which participants engaged in score optimisation.

*5.1.3 Task Completion.* Across all runs, completion rates were higher in the Map interface than in the List interface (Figure 7). Of the 32 Map View scenario runs, 31 were completed before timeout, compared to 25 of 32 List View scenario runs.

*5.1.4 Post-Test Survey Results.* Participants completed a post-test survey in which they rated their experiences with the Map View and List View interfaces across five Likert-scale items: perceived stress points, confidence in control, understanding of system state, mental effort, and perceived support for decision-making. Responses were analysed using paired Wilcoxon signed-rank tests, treating the Likert items as ordinal data.

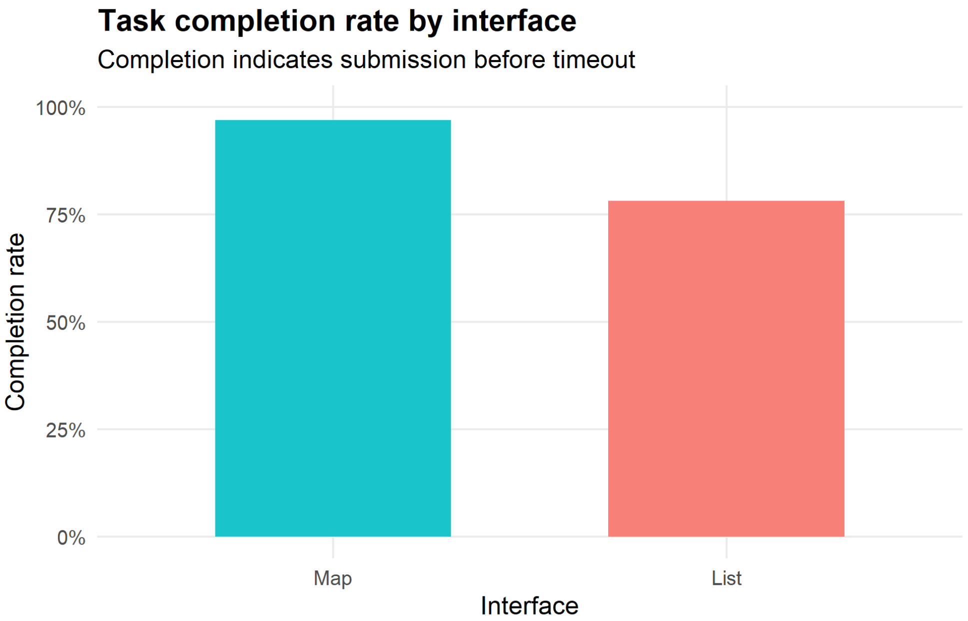


Figure 7: Task Completion by Interface.

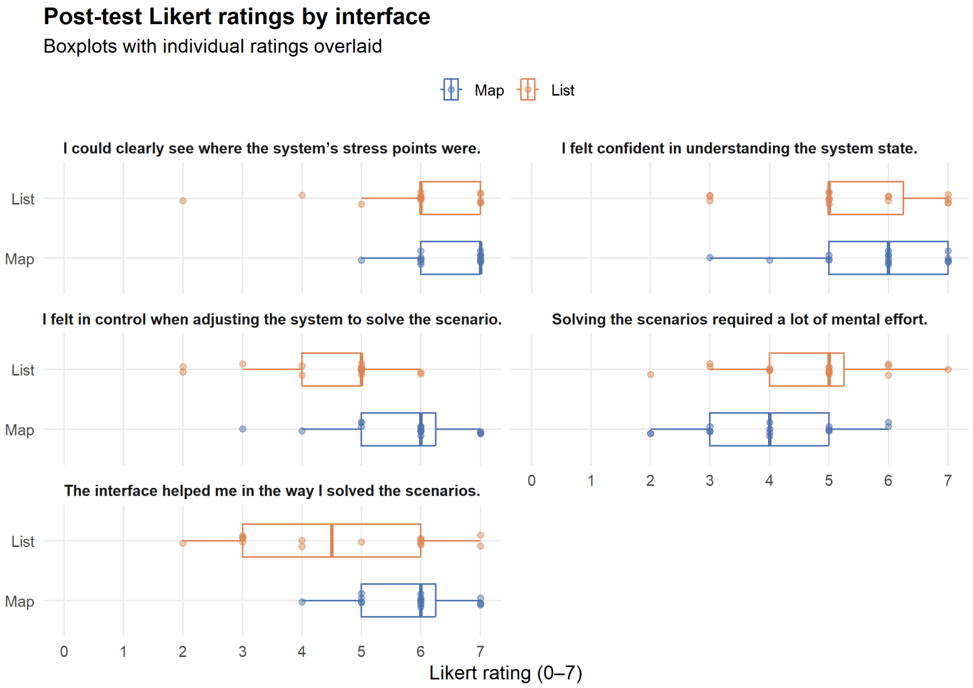


Figure 8: Post-Test Survey Results.

Significant differences between interfaces were observed for three of the five items (Figure 8). Participants reported a clearer perception of system stress points ($W = 28.0, p = .018$), greater confidence in control ($W = 74.5, p = .005$), and stronger perceived support for decision-making ($W = 86.0, p = .037$) in the Map View interface. In contrast, no statistically significant differences were found between interfaces for perceived understanding of the system state ($p = .14$) or mental effort ($p = .10$).

### 5.2 Qualitative Results

The reflexive thematic analysis method identified four key themes, describing how participants explored data, interacted, evaluated their decisions, and formulated strategies in the two interfaces.

*5.2.1 Data Exploration.* Participants discussed their decision-making process by how and why they explored the data. This part of the process was described as the initial stage to identify issues and compare options.

In Map View, participants felt that the data was best explored once generators were filtered by both costs and emissions. This way, all of the generator's parameters, as well as its location, were shown at the same time. Participants found it a fast and intuitive way to overview the available options.

In List View, data exploration was divided between sorting the Active and Inactive Generators Lists and referencing generator locations on the grid map. Several participants stated that the sorting function was especially useful for identifying extreme values, such as the most expensive generators. Comparing multiple parameters simultaneously was described as difficult, as sorting by one parameter led some participants to disregard the others. This was expressed as Map View enabling "multi-tasking".

In Map View, generators were not contained in a structured list, only visualized by their location. Participants described this as hurting confident exploration, since options could be overlooked if they were outside the current view. The same problem was reported when finding overloaded transmission lines, which if out of view, had no feedback on the map when highlighted in the overview bar.

A majority of participants described understanding the data in terms of spatial relationships. They described comparing options based on their relevance to stressed regions of the grid. Participants expressed that this approach was better supported in Map View, as attention could remain on the grid map. Participants used the direction of an overloaded transmission line's chevron to guide them towards the relevant area of options. In contrast, participants thought there was a disconnect between location and parameter information in List View. A number of participants stated that they largely ignored the map element when using List View, explaining that cross-referencing generators between the map and list felt slow and disruptive to their decision-making.

> "What helped me the most with Map View was how it showed where the generators were located. It was clearer than to compare between [the generator cards]. Everything was collected in one spot— costs and emissions— and without having to show any numbers, made it easier to compare them." (P7)

Some participants described how using the numerical values displayed for every generator in List View were the most intuitive way of comparing options. They contrasted this visualization with Map View's abstract encoding of these values into colour shades, which were perceived as more difficult to use when comparing values.

> "What I liked about List View was that it was very easy to see specific generators, how much emissions they have and how much they cost. In Map View it's more guesswork, because I'm mostly going by how 'dirty' they look." (P11)

Several participants described experiencing information overload in List View which prevented them from making confident decisions. Participant 1 remarked how List View's detailed design would be beneficial in getting more perspectives on the data, but after having used Map View realising he "didn't need the information in two places; it made more sense to have it in one place." The layout led participants to hesitate or second-guess their choices, even when several attributes were expressing the same idea.

> "There's an overwhelming feeling in some of this [List View]. It felt like being on the stock market. And I just thought, okay, I'm not built for this." (P16)

Rather than favouring one interface over the other, most participants proposed combining elements of both layouts. A common suggestion was to allow users to expand the simplified Map View by selecting elements to reveal specific details in modular windows. Another recurring suggestion was to combine the Inactive and Active Generators' lists into one list that could be sorted by applying tags, allowing the user to sort by several parameters at once. Similarly, participants suggested the Map View include more filters, hiding options above or below a desired value.

*5.2.2 Action.* Participants discussed how they interacted with the system and how immediate feedback informed their actions. A common experience in List View was that the increased number of interface elements encouraged a slower and more deliberate process. Participants explained that performing the same action required more individual steps than in Map View, as they needed to shift attention between the generator lists and the map. In addition, when generators were switched on or off, cards moved between lists, which several participants described as confusing and disruptive to their actions. Participants contrasted this with Map View, describing its layout as simpler to act upon due to its more minimal layout and the direct manipulation of generators on the map.

> "List View I found myself fighting with, especially when generator cards kept jumping between lists. I couldn't follow how the interface reacted." (P5)

Early in the scenarios, participants relied closely on the visible effect of individual actions to guide subsequent steps. Over time, some participants reported becoming less dependent on immediate feedback from single actions, expressing greater confidence in the overall direction of their decisions.

Participants expressed List View's sorting function as an intuitive way to control and act on the system. Many participants felt List View was faster when adding or removing many generators at once according to a certain parameter. A recurring approach to solving scenarios in List View was by removing all active generators to begin with a "clean slate". The colour coding associated with each generator's fuel type was described as a helpful indicator and was used as a shorthand for higher or lower values, leading participants to add all hydro generators first, followed by generators with the lowest costs and emissions, until the demand requirement was satisfied.

*5.2.3 State Evaluation.* Participants described another stage of their decision-making process as evaluating the system state, where they determined their decisions' effect on the system state, and identified possible next steps. Participants used the overview bar's cards to determine the effect of their interactions on the overall system.

Participants noted that the only card which expressively addressed the comparison of states were the red or green arrows below the top transmission lines, which visually indicated increases or decreases in load percentage. In contrast, changes related to power output, cost, and emissions were described as requiring mental comparison, as these parameters were not directly contrasted within the interface. Participants stated that Map View's singular focus on the map helped them anticipate how a decision affected the system state based on changes in the colour of nearby line chevrons.

For decisions involving multiple generators, the system allowed no way to save, undo, or directly compare solutions. Participants described this as a central challenge of solving scenarios, given the large number of possible combinations of active generators and generator power levels.

> "I was mostly frustrated… I can't save a particular setting, and so I don't know what I am comparing." (P4)

Participants connected evaluating the system to understanding the logic of power grids. They discussed the abstracted nature of the data in Map View in contrast to the perceived transparency of List View's layout. Several participants described Map View as easier to understand due to its simplistic presentation. At the same time, participants noted that abstraction could obscure important factors, making the interface feel opaque. Participant 1 referred to this experience as a form of "black boxing," where mechanisms influencing system behavior were not directly visible.

> "I was thinking about the concept of black boxes. That the user isn't shown vital information. [Map View] abstracts so many, even important factors. [The risk is that] this allows you to stop thinking about them, when you ought to be." (P1)

Participants reported that, while they learned how to solve scenarios, they did not gain a clear understanding of the underlying logic of power grid behavior. Several participants described initially attempting to resolve overloaded transmission lines by tracing them to individual generators perceived as a "root cause." Participants who used this approach reported either that it was ineffective or that the time constraint led them to adopt a different strategy. When asked about potential improvements to the interface, several participants suggested that the system lacked visual support for identifying which generators contributed to a specific overloaded transmission line.

*5.2.4 Strategy.* Lastly, participants described how they approached solving scenarios and how this changed over time; how they developed their decision-making process. Across the user test, participants reported that their approach to solving scenarios was the same in both interfaces. However, they felt that one interface supported this process better than the other. This came down to if they felt more confident making decisions based on the visual encodings of parameters in Map View or numerical values displayed in List View.

Most participants described their initial strategy as one of trial-and-error. This involved activating or deactivating individual generators, assessing the resulting system state, and reverting actions when outcomes were unfavourable. According to participants, this approach also functioned as their way of learning to use and understand the system. By observing changes in system indicators following each interaction, participants said that they gradually formed an understanding of how generator activation affected power balance, line loading, and other parameters. Learning was also described as having a gut feeling about the system's limits. However, not all participants found trial-and-error equally productive. Some reported difficulty forming a coherent picture of the system through this approach, describing their interactions as lacking clear direction. Under time constraints, participants appeared willing to continue experimenting in order to reach a valid solution, even if this meant proceeding without a clear understanding of the underlying system logic.

> “I’m not seeing it. I don’t get the whole picture of how things fit together. I could click around at random, but I won’t get closer to anything I find logically a better option.” (P1)

Participants described how over time, their strategy was formed around a set of priorities. For most participants, ensuring system stability— avoiding overloaded transmission lines— was treated as the primary concern. Participants described this as establishing a solvable system before optimizing trade-off parameters for a better score. Participants adopted a self-imposed strategy of using as much of the available time as possible to refine their solutions.

> “Line overloads were the first thing I looked at. Especially with this kind of challenge I had to identify, ‘What’s the situation?’ I would tend to look at overloads and the budget requirements.” (P16)

Over time, several participants reported taking power demand for granted, describing it as a parameter they felt fully in control of. This was contrasted with line stability, which was experienced as more indirect and less predictable. As one participant explained, power adjustments fell under their control, while line load was understood primarily as the outcome. Participants also noted how the visual mapping of the parameters influenced how they judged their importance.

> “The colour scale [cost] popped out more than the hash pattern [emissions]. In that sense, cost seemed more important than emissions.” (P7)

Participants also described their learning of the system as when they began to see the grid on a global scale rather than a local scale. Many participants described how they approached problem solving from a local viewpoint at first, attempting to understand grid behavior by focusing on individual generators or transmission lines. Participants would adjust a single generator to see if it affected a nearby stress point. Several participants described this approach as limited or misleading, as they came to realize that local changes rarely explained system-wide behavior.

> “Switching to Map View, I came to understand that turning a lot of generators on in particular clusters, the load would even out. And if a problem persists, I could then start adjusting on a more granular level.” (P7)

This shift was described as zooming out on the map to keep all generators visible, relying on higher-level patterns to guide their decisions, and zooming in when addressing a specific transmission line that remained overloaded. However, participants did not believe the layout fully supported a global view, as it was harder to discern and interact with elements on this scale: colour coding and hash patterns were harder to identify; the generators were smaller and harder to adjust, and the line chevrons were barely visible.

## 6 DISCUSSION

### 6.1 Research Question 1: Decision-Making Across Interface Layouts

The first research question concerns how the decision-making process of users is impacted by the interface layout used when solving tasks. Results show that the process of users is largely the same across interfaces. Participants used both the List View and the Map View to explore and compare data in terms of different relationships, evaluate the impact of each decision in terms of how the system state has changed, and develop strategies. These strategies were a combination of experimental, trial-and-error interactions through which participants understood the system's behaviour and limits. Strategies also concerned a sequence of priorities, where base requirements were met before optimizing trade-off parameters. Lastly, participants began seeing the system from alternating global and local perspectives when solving general or specific issues.

Participants largely saw the main problem of each scenario to be one of solving the problem of overloaded lines. In turn, the goal was to understand how power should be distributed throughout the grid map in relation to the location of stressed lines. This made participants mainly make decisions based on the spatial relationships of generators. In Map View, comparing options could therefore be filtered out mentally by knowing which cluster of generators would fix a particular overloaded line. This process was not as supported in List View, where the layout did not support a simultaneous comparison of options based on location and parameter values; generators could not be sorted in the list based on location because the lists were structurally linear.

The non-linear visualization of data in the Map View layout also allowed participants reason in terms of several parameters simultaneously. List View's sorting functionality not only limited the user to sorting the lists by one parameter at a time, but participants also saw this as limiting their ability to compare options, while Map View allowed for comparison across parameters.

Another insight is how interaction is required in order for users to gain full situational awareness. Both interfaces were sufficient for participants' understanding of the available options and the current state of the system. However, in order to predict their decisions' impact on the system, participants had to compare the effect of adjusting different generators. Gradually, the participants formed a mental model of the system's limits, and which options would have the best effect on solving a certain problem.

Lastly, results highlight how layout impacted the pace, or fluidity of making decisions. In the interviews, participants described Map View as encouraging a fast, general assessment of the data, comparing all parameters at once, while List View was better suited for a slower and more deliberate way to make changes to the system.

### 6.2 Research Question 2: Usability Across Interface Layouts

Taken together, qualitative and quantitative results highlight key affordances and limitations of the List View and Map View interface layouts.

Map View was perceived as having high learnability, intuitive controls, and best suited for general judgments. For most participants, Map View was the tool that helped them understand how to approach solving scenarios. Interview data and survey results highlight how Map View also exhibited high controllability. Participants note direct manipulation and interactive filters as best supporting their decision-making process. These methods helped users not only make fluid choices but also understand their choices' impact on the system state.

While simple design through abstraction of numerical values made it easier for participants to get an overview of the available options, it limited the ability for precise comparison. In addition, the visualization of grid elements were not equally intuitive to compare or control on different scales.

By comparison, the List View was perceived as a more complex system. Participants described the layout as fragmenting their decision-making process in several ways, by separating generator cards across two lists, and by separating the generators' location on the map from their details. This required frequent cross-referencing and increased mental effort. The density of information in List View caused some participants to experience information overload, hesitating or second-guessing their decisions, and slowing down their process.

At the same time, participants found confidence in List View's structured visualization of the available options. List View's sorting method allowed for users to activate or deactivate groups of generators efficiently according to specific criteria, such as the identification of extreme values. The List View, which displayed each generator's exact numerical values, better supported precise comparison when dealing with trade-off parameters.

The main limitation across interfaces was the lack of support for comparing system configurations. Participants could not save or contrast alternative configurations; as a result, decisions were often limited to adjusting single generators or focusing on one or two criteria at a time. Without a way to compare the combined effect of multiple adjustments, participants struggled to evaluate broader strategic differences between solutions.

*6.2.1 Abstraction and Accuracy.* The interview data highlighted a trade-off between abstraction and accuracy. In the case of this study, the prototype took a simplistic, abstract approach in representing a complex system. Despite power-grid control systems being complex systems difficult to approach without expert knowledge [23], participants in this study were able to quickly engage with the prototype, understand the task structure, and complete scenarios. This suggests that abstraction succeeded in lowering the threshold for participation, allowing novice users to reason about constraints, trade-offs, and system state within a short time frame. At the same time, this learning was primarily procedural rather than conceptual. While users became increasingly confident in operating the system and improving their solutions, many expressed uncertainty about how the underlying grid dynamics functioned, relying instead on interface feedback and scoring cues to guide their actions. Thus, results suggest that abstraction enables quick learning, but at the cost of deep understanding.

*6.2.2 Shifting Between Levels of Detail and Scale.* Results show that usability depended on the ability to shift between levels of detail. When evaluating the system state, participants shifted between seeking a general overview and examining precise details. An ideal system should give the user more control over data visualizations by different levels of detail. Participants suggested a modular system that is rooted in a wide overview but can be subsequently expanded by focusing on the details of specific elements.

A similar challenge concerned how to visualize the map topology on different scales. While users were given local control over elements, neither interface adequately supported a stable, zoomed-out overview of the system state. Participants were required to zoom in to interpret small visual elements, such as line-load chevrons. Data visualization tools, such as Overbye's Geographic Data Views, address this by presenting simplified global patterns at a distance and revealing detail progressively when zooming in. This principle was not fully realised in the prototype, limiting usability [22].

*6.2.3 Interaction Methods.* Results also showed how interactive methods mitigated previously held visualization challenges, such as visual obtrusion [14]. The prototypes in this study employ interactive methods and thus mitigate these issues by allowing users to add or remove information. Users gain confidence through different visual cues such as numerical, visual or spatial relationships. Interactive methods allow the user to break down the information freely according to their preferred reasoning. Future research could investigate the limits of where added user control reduces perceived usability, when visualizations should or shouldn't be interactive. Finally, the functionality of interfaces should also inherit those of conventional systems, so that the user with less experience can predict how the system will react based on prior experience with similar tools. Participants reported uncertainty when selecting overloaded lines that were outside the current view, as the system provided no visible feedback unless the highlighted element was already on screen. This led users to question whether their action had resolved the issue or simply moved it out of view. As noted by Participant 3, this could be mitigated by inheriting functionality from tools such as Google Maps, where the view automatically re-centers on selected elements.

## 6.3 Limitations

*6.3.1 Quantitative Metrics.* There was no statistical significance found in the scenario performance metrics—time difference, task completion, or score. For the time difference metric, the values were uniformly close to the maximum time allotted per scenario, which reflect the goal of the scenarios, which were based on optimization rather than speed. In regards to score, the scoring system served primarily as a design mechanism to motivate optimization, rather than as a bounded performance metric. Participants rarely referred to the score explicitly when describing their reasoning. Lastly, task completion rates were high in both interfaces, suggesting that reaching a solvable solution was generally straightforward, while the challenge of each scenario lay in optimization. These results could not deepen the insights provided by the interview data.

The time metric used in the evaluation expressed *time to submission* instead of *time to the first stable solution*, reflecting the focus on optimization. If the focus had been shifted to speed, capturing how quickly participants first reached a stable system state, the evaluation could have provided additional insight into learning effects over time or differences between interfaces. For example, Participant 6 remarked that they felt faster and less stressed after their initial use of the Map View, and later also in the List View once interaction became more familiar ("I felt faster… I wasn't stressed in the same way"). Such comments suggest a perceived ease-of-use or growing confidence that is not reflected in the recorded time-to-submission data. Several participants also reported that they felt they reached stable solutions more quickly in later scenarios, even when their recorded completion times remained similar.

In addition, the post-test survey questions were sometimes more effective in generating conversation points than as qualitative metrics. For example, the statement "*Solving the scenarios required a lot of mental effort.*" was interpreted differently by participants—some experienced Map View as requiring less effort because it made scenarios easier to solve, while others rated it as higher because Map View allowed for deeper reasoning than List View; for example, allowing them to account for multiple parameters simultaneously.

*6.3.2 Within-Group Study Design.* Because the study was a within-group design, the participants' understanding of the system and strategy was developed from using both interfaces. Perhaps Research Question 1 would be more accurately answered if the study was a between-group design, where participants learned only one of the two interfaces.

*6.3.3 Limitations Related to Scenario Design and Dataset Simplification.* In addition to this, the data itself might have hindered participants from developing nuanced decision-making strategies. Hydro-power generators were both the cheapest and least emissive options, making them immediately identifiable as preferable choices. This led to many participants starting off each scenario immediately turning all of the hydro generators on, before taking a closer look at the specific target requirements. While this was the grid data tied to the IEEE-118 bus system operation manual, it could have led to the evaluation being less effective in its goal to investigate how users reason under different constraints.

*6.3.4 No Video Recording.* Keim's evaluation used video recording while users performed tasks, which captured another side of the user experience, that which the user does not acknowledge themselves [12]. Using video recording technology could have served as another aspect with which to compare results, such as monitoring a user's visual attention patterns, or recurring misclicks when executing an action. This study instead takes the risk of treating the perceived experience of the user too seriously. The consideration here was to use the quantitative results as a counter-weight.

## 7 CONCLUSION

This study investigated decision-making and usability in power grid control systems by having users use and compare two opposing interfaces.

Results showed how the layout and functionality of the interfaces shaped the users' sense of control, confidence, and understanding of power grid data. Users developed their decision-making processes in the same way regardless of the interface, but their strategies were supported differently in the two interfaces. Users learned the systems through interaction, and approached decisions in terms of spatial or numerical relationships. They also benefitted from being able to shift between levels of data complexity in alignment with different phases of their decision-making process. The evaluation supports a shift from focusing on whether users understand the visualizations to how well functionality supports users' reasoning about the system.

The study design was limited by a lack of robust quantitative metrics that could challenge the findings of interviewing study participants. To address the limitations identified in this study, future work could vary key elements of the study design by applying more targeted quantitative and qualitative evaluation methods, including: (1) orienting the task around alternative goals to optimization, such as speed or system stability; (2) introducing scoring systems with clearly defined bounds, and (3) using more varied datasets to allow finer-grained differentiation of user performance. Further work may also isolate how gamification elements or mapping could shift user priorities. Lastly, the same study design, as well as established qualitative methods of evaluation, such as Nielsen's Heuristics, may be applied to evaluate control systems used in real-world grid operation [17].

## 8 GENERATIVE AI DISCLAIMER

The generative AI tool ChatGPT-4 was used for prototype development in a debugging context. The prototype code has been edited and reviewed by the author, William O'Grady. For more information about the nature of the usage of AI in the project, please contact the author.

## ACKNOWLEDGMENTS

I would like to give very special thanks to my supervisor Rémy Rey, for making this long year spent working on the project so rewarding. To Felix Blomkvist and Leonard Lund, whose friendship made every day a little brighter. To my father, an invaluable partner in getting back in the saddle, day in and day out. Finally, my gratitude to all the participants who took time out to play my little game.

## REFERENCES

[1] Alexander A Anderson, Slaven Kincic, Brett A Jefferson, Blaine J Mcgary, Corey K Fallon, Danielle K Ciesielski, John E Wenskovitch, and Yousu Chen. 2022. *A real-time operations manual for the IEEE 118 bus transmission model*. Technical Report. Pacific Northwest National Laboratory (PNNL), Richland, WA (United States).

[2] Richard A. Becker, Stephen G. Eick, and Allan R. Wilks. 2002. Visualizing network data. *IEEE Transactions on visualization and computer graphics* 1, 1 (2002).

[3] Virginia Braun and Victoria Clarke. 2012. *Thematic analysis*. 978-1-4338-1003-9

[4] Corentin Burnay, Fatima Dargam, and Pascale Zarate. 2019. Data visualization for decision-making: an important issue. *Operational Research* 19, 4 (2019), 853–855.

[5] Elizabeth Charters. 2003. The use of think-aloud methods in qualitative research an introduction to think-aloud methods. *Brock Education Journal* 12, 2 (2003).

[6] Peiguang Chen, Yiwen Jiang, Yuchun Liu, Xuanduo Yu, and Shunjie Han. 2022. Research and Implementation of Power System Data Visualization. In *2022 5th International Conference on Robotics, Control and Automation Engineering (RCAE)*. IEEE, 169–173.

[7] Kai-Wen Cheng, Yize Chen, and Yuanyuan Shi. 2023. Gridviz: a toolkit for interactive and multi-modal power grid data visualization. In *2023 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE.

[8] Mica R Endsley. 1995. Measurement of situation awareness in dynamic systems. *Human factors* 37, 1 (1995).

[9] Maximilian T Fischer and Daniel A Keim. 2021. Towards a survey of visualization methods for power grids. *arXiv preprint arXiv:2106.04661* (2021).

[10] Mirjavad Hashemi Gavgani and Sara Eftekharnejad. 2017. A graph model for enhancing situational awareness in power systems. In *2017 19th International Conference on Intelligent System Application to Power Systems (ISAP)*. IEEE.

[11] Illinois Center for a Smarter Electric Grid. n.d.. IEEE 118-Bus System. https://icseg.iti.illinois.edu/ieee-118-bus-system/

[12] Alexander Jäger, Sebastian Mittelstädt, Daniela Oelke, Sonja Sander, Axel Platz, Gies Bouwman, and Daniel A Keim. 2016. Lessons on combining topology and geography: Visual analytics for electrical outage management. (2016).

[13] Pramod M Mahadev and Richard D Christie. 2002. Envisioning power system data: concepts and a prototype system state representation. *IEEE Transactions on Power Systems* 8, 3 (2002).

[14] Christine Mikkelsen, Jimmy Johansson, and Matthew Cooper. 2012. Visualization of power system data on situation overview displays. In *2012 16th International Conference on Information Visualisation*. IEEE, 188–197.

[15] Federico Milano. 2009. Three-dimensional visualization and animation for power systems analysis. *Electric Power Systems Research* 79, 12 (2009).

[16] Dao Viet Nga, Ong Hang See, Chee Yung Xuen, Lai Lee Chee, et al. 2012. Visualization techniques in smart grid. *Smart Grid and Renewable Energy* 3, 03 (2012).

[17] Jakob Nielsen. 1995. How to conduct a heuristic evaluation. *Nielsen Norman Group* 1, 1 (1995).

[18] Thomas J Overbye, Raymond P Klump, and Jamie D Weber. 1999. A virtual environment for interactive visualization of power system economic and security information. In *1999 IEEE Power Engineering Society Summer Meeting. Conference Proceedings (Cat. No. 99CH36364)*, Vol. 2. IEEE.

[19] Thomas J Overbye and Jamie D Weber. 2000. New methods for the visualization of electric power system information. In *IEEE Symposium on information visualization 2000. INFOVIS 2000. Proceedings*. IEEE.

[20] Thomas J Overbye and Jamie D Weber. 2000. Visualization of power system data. In *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*. IEEE.

[21] Thomas J Overbye and James D Weber. 2002. Visualizing the electric grid. *IEEE spectrum* 38, 2 (2002).

[22] Thomas J Overbye, Jessica L Wert, Komal S Shetye, Farnaz Safdarian, and Adam B Birchfield. 2021. The use of geographic data views to help with wide-area electric grid situational awareness. In *2021 IEEE Texas Power and Energy Conference (TPEC)*. IEEE.

[23] Sonal K Phan and Cathy Chen. 2017. Big data and monitoring the grid. In *The power grid*. Elsevier.

[24] Nils Rodrigues, Rudolf Netzel, Kazi Riaz Ullah, Michael Burch, Alexander Schultz, Bruno Burger, and Daniel Weiskopf. 2017. Visualization of time series data with spatial context: Communicating the energy production of power plants. In *Proceedings of the 10th International Symposium on Visual Information Communication and Interaction*.

[25] Robert E Roth. 2017. Visual variables. *International encyclopedia of geography: People, the earth, environment and technology* (2017).

[26] Maria-Angeles Sanchez-Hidalgo and Maria-Dolores Cano. 2018. A survey on visual data representation for smart grids control and monitoring. *Sustainable Energy, Grids and Networks* 16 (2018).

[27] Maria Stefan, Jose G Lopez, Morten H Andreasen, and Rasmus L Olsen. 2017. Visualization techniques for electrical grid smart metering data: A survey. In *2017 IEEE third international conference on big data computing service and applications (bigdataservice)*. IEEE.

[28] Yan Sun and Thomas J Overbye. 2004. Visualizations for power system contingency analysis data. *IEEE transactions on power systems* 19, 4 (2004).

[29] Svenska kraftnät. 2024. Svenska kraftnät – Grid development plan 2024–2033. https://www.svk.se/en/about-us/sustainability/. Accessed: 2026-01-28.

[30] Pak Chung Wong, Harlan Foote, Patrick Mackey, George Chin, Zhenyu Huang, and Jim Thomas. 2011. A space-filling visualization technique for multivariate small-world graphs. *IEEE Transactions on Visualization and Computer Graphics* 18, 5 (2011).

[31] Pak Chung Wong, Kevin Schneider, Patrick Mackey, Harlan Foote, George Chin Jr, Ross Guttromson, and Jim Thomas. 2009. A novel visualization technique for electric power grid analytics. *IEEE Transactions on Visualization and Computer Graphics* 15, 3 (2009).

[32] Ji Soo Yi, Youn ah Kang, John Stasko, and Julie A Jacko. 2007. Toward a deeper understanding of the role of interaction in information visualization. *IEEE transactions on visualization and computer graphics* 13, 6 (2007).

[33] Fenglei Zheng, Yunfeng Xia, Huafeng Su, and Wanwei Wang. 2021. Research on key technologies of power system visualization platform. In *2021 13th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*. IEEE.

## 9 APPENDICES

## A TUTORIAL SCRIPT

### A.1 Introduction

First off, thank you for taking part in this study.

### A.2 Task Overview and Power Demand

- The task you are given is that you are an ENERGY GRID MANAGER who must ensure power demand is being met and that the grid is stable. The system state is shown in the OVERVIEW BAR.
- You will achieve this by turning on and off POWER STATIONS/GENERATORS across the GRID.
- Each power source can be fully or partially enabled. e.g. 200 MW of a possible 400 MW, yielding different results.

### A.3 Grid Stability

- One of the issues you will be faced with is that the power flows through TRANSMISSION LINES—and too much power from one side of the Grid flowing to another side through a particular LINE will OVERLOAD it. Think of this as traffic congested from one direction.
- The interface will indicate to you when lines are OVERLOADED.
- To achieve grid stability, you must balance and spread out power evenly throughout the grid. Sometimes this is more complex than solving a particular overload; the whole grid works as one big system.

### A.4 To Complete a Scenario

- To successfully complete a SCENARIO you must not only supply the required amount of Power but also ensure that no lines are Overloaded. When and only when this is achieved, will the SUBMIT button become Active and you can successfully finish the scenario.

### A.5 List View

- The first interface is called LIST VIEW. This is loosely based on what power grid operators currently use. Here, the map of the power grid is just a reference. Actually controlling the grid happens in lists detailing the ”Stats” of each generator.

- Along the top of the interface you have the OVERVIEW BAR. Here is displayed the SYSTEM STATE: how close you are to meeting the target power, what lines might be overloaded, the cost and emissions totals, as well as the task section where you can see your score and submit the task.
- You can add generators by interacting with their cards. Turning a generator on moves it from being ”INACTIVE” to being ”ACTIVE”.
- Each GENERATOR CARD shows the generator’s power capacity, represented in rounded blocks. Each block represents 100 MW. You can turn the generators on by clicking the desired amount of blocks, or turn on the generator FULLY by pressing the PLUS SIGN.
- Likewise, you can remove a generator from the active section by pressing the MINUS SIGN.
- The reference map is still used to navigate the power grid. To navigate, you drag and zoom, much like GOOGLE MAPS.

### A.6 Map View

- The second interface is called MAP VIEW. This interface instead extends the reference map to be the MAIN METHOD in which you control the power grid.
- The OVERVIEW BAR is also more interactive. By pressing the overview buttons, you can scale the generators by cost and emissions, as well as focusing on generators or lines specifically.
- Regulating POWER PER GENERATOR is as easy as pressing the individual blocks on the generator nodes. Each block represents 100 MW.

### A.7 Score

- To make scenarios more interesting you can try to achieve a HIGH SCORE.
- How is the SCORE calculated? - Each Power Source has different costs and emissions associated with them - Hydro, Coal, Gas, Combined.
- As you make selections for the Power settings you will be able to see the associated Cost and Emissions level change to reflect your choices. Note: If a Generator has a Max output of 400MW then a setting of half capacity i.e. 200MW will cost half and emit half of the full 400MW. It’s a linear scale!
- Each Scenario has a recommended budget. For every dollar under that budget you will get a point added to your SCORE—e.g. if the budget is $10,000 and you SUBMIT a solution costing $9000, your SCORE will be 1000 points.
- Each Scenario also has a recommended Emissions level. If you manage to meet or be below the Emissions level then you get a Multiplier BONUS x2—which would mean 2000 points!
- TO REITERATE: A Scenario is submittable even if the cost and/or emissions cap is not met. This is still legit as a stable system. Achieving a high score is up to you.

Remember to THINK ALOUD!

Let’s move on to getting comfortable with the interfaces!

### A.8 Easy Tasks to Learn the System

*Map View:*

(1) Find the ”biggest” generator (the generator with the highest capacity).

(2) Apply ”Focus” on the transmission lines.

(3) Apply ”Focus” on generators, and ”scale by emissions”.

(4) Identify the line with the highest load percentage.

*List View:*

(1) Identify the ”biggest” generator through ”Sort”.
(2) Add a generator with power $>$100MW by pressing the ”$+$” button.
(3) Decrease that generator’s output power.
(4) Identify that generator on the map by pressing the generator card.
(5) Press a generator on the map to jump to it in the list.
(6) Identify the line with the highest load.

## B POST-TEST SURVEY

To be rated on a seven-point Likert Scale (1: disagree completely → 7: agree completely).

(1) I could clearly see where the system’s stress points were.
(2) I felt in control when adjusting the system to solve the scenario.
(3) I felt confident in understanding the system state.
(4) Solving the scenarios required a lot of mental effort.
(5) The interface helped me in the way I solved the scenarios.

## C INTERVIEW QUESTIONS

(1) What were your initial thoughts for each interface before and after the test?
(2) Tell me about your experience navigating and interacting with each interface.
(3) What strategies did you use for solving tasks? Did they differ between interfaces?
(4) Did you prefer one interface over the other? Why?
(5) If you could combine features of each interface into one, ideal interface, what aspects would you lift from each?
(6) What made you believe a decision was a good decision?
(7) Did you feel overwhelmed at any point during the test?

TRITA-EECS-EX-2026:86
Stockholm, Sweden 2026

www.kth.se