

Matches, Mismatches, and Methods: Multiple-View Workflows for Energy Portfolio Analysis

Matthew Brehmer, Jocelyn Ng, Kevin Tate, and Tamara Munzner, *Member, IEEE*

Abstract—The energy performance of large building portfolios is challenging to analyze and monitor, as current analysis tools are not scalable or they present derived and aggregated data at too coarse of a level. We conducted a visualization design study, beginning with a thorough work domain analysis and a characterization of data and task abstractions. We describe generalizable visual encoding design choices for time-oriented data framed in terms of *matches* and *mismatches*, as well as considerations for workflow design. Our designs address several research questions pertaining to scalability, view coordination, and the inappropriateness of line charts for derived and aggregated data. We also present guidelines relating to *familiarity* and *trust*, as well as methodological considerations for visualization design studies. Our designs were adopted by our collaborators and incorporated into the design of an energy analysis software application that will be deployed to tens of thousands of energy workers in their client base.

Index Terms—Design study, design methodologies, time series data, task and requirements analysis, coordinated and multiple views.

1 INTRODUCTION

Consider a university campus containing about a hundred buildings. For a university operations worker looking for opportunities to conserve energy, visualization can be helpful when analyzing and monitoring the energy performance of this large portfolio of buildings.

In this design study, we collaborated with a team of people at EnerNOC, a company that develops energy analysis and reporting software for organizations such as commercial business chains, universities, and utility companies. Our collaborators' goal was to deploy a redesigned version of Energy Manager, their energy analysis software tool; in doing so, they hoped to retain their existing client base encompassing thousands of organizations, attract new clients, and increase engagement with their software. Meanwhile, our goal as researchers was to successfully integrate our research process into our collaborators' software development practice. We designed and evaluated potential visualizations with a variety of stakeholders in an industry setting, which included the collaborator's colleagues as well as their clients. This paper documents a success story, where our collaborators committed software development resources and adopted our visualization designs that resulted from our research.

Visualization researchers and practitioners working in domains unrelated to energy analysis will find several transferable aspects of this paper, beginning with our characterization of data and task abstractions. This project required that we design visualizations and promote sophisticated visual analysis by individuals accustomed to unsophisticated visual idioms. We needed to identify scalable visual encoding and interaction idioms that can accommodate dozens to hundreds of concurrent time series. To complicate matters further, we could not rely upon the visual idiom of a line chart due to a combination of data semantics and domain convention. Once we identified appropriate mappings between visual idioms and individual tasks, we next addressed the question of accommodating task sequences: determining which visualizations ought to be juxtaposed in the same display, and which ought to be presented sequentially. Coordinating these visualizations also posed a challenge; specifically, we sought to reduce the amount of manual navigation between visualizations, as this is an issue with the existing Energy Manager tool.

Contributions: Our primary contribution is a set of generalizable design choices and guidelines framed in terms of **matches** and **mismatches** between abstractions and visual encoding idioms for time-oriented data, guidelines that transfer beyond the energy domain. We also present guidelines relating to the themes of **familiarity** and **trust**. The former refers to the spectrum between ubiquitous visual encodings and those that a prospective user may have never seen before, while the latter refers to the appropriate display of derived and aggregated data, as well as giving users control over these data transformations. Finally, we contribute **methodological advice** for visualization design projects. This includes considerations for designing workflows that incorporate multiple views; while prototyping the visual encoding design of a single view has received considerable attention in the literature [24], workflow prototyping has received far less.

Outline: We begin by describing our research and design methodology in Section 2. We then characterize the data and tasks in Section 3. We consider how the tasks we characterized are supported by the existing Energy Manager tool in Section 4, and how previous work addresses similar data and tasks in Section 5. We then briefly describe our prototyping environments in Section 6. The visual encoding design of individual views is discussed in Section 7, where we characterize these designs in terms of matches and mismatches with regards to combinations of data and task. The design of workflows incorporating multiple views is discussed in Section 8. We report on which of our designs were adopted by inclusion into our collaborators' product development cycle in Section 9. In Section 10, we reflect on familiarity, trust, and on methodological considerations for visualization design studies. Section 11 summarizes our contributions.

2 METHODOLOGY

In this section, we focus on *how* we conducted our research; we reflect upon our methodological decisions and provide methodological advice in Section 10.2.

Analyzing the work domain: Over the course of five months in 2013, we conducted 9 in-depth interviews with current users of Energy Manager. Eight of those interviewed were employed by client organizations, which included three North American universities, an engineering consulting firm, and a school board. The final interviewee was a colleague of our collaborators who regularly consulted with new clients. Although we use the term *energy worker* to describe these individuals, we encountered a wide range of roles, job titles, skill sets, and levels of training. The use of energy analysis tools such as Energy Manager also varied in terms of context and frequency of use. Despite these differences, we did identify several common goals and activities relating to the energy analysis of building portfolios, which we characterize in terms of data and task abstractions in Section 3.

Validating the abstractions: We validated our abstractions by check-

• Matthew Brehmer and Tamara Munzner are with the University of British Columbia. E-mail: {brehmer,tmm}@cs.ubc.ca.

• Jocelyn Ng and Kevin Tate are with EnerNOC, Inc. E-mail: {jocelyn.ng,kevin.tate}@enernoc.com.

Manuscript received 31 Mar. 2015; accepted 1 Aug. 2015; date of publication xx xxx 2015; date of current version xx xxx 2015.

For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

ing back with a subset of the people that we had previously spoken to during the work domain analysis phase: our primary collaborators and two “power user” energy workers. To do so, we consolidated our thoughts and findings into a slide deck that contained screenshots, examples, mockups, and notes. These slides were a living research artefact: we used them to present previous findings to collaborators and energy workers, and we recorded their feedback as further annotations.

Eliciting feedback on visual encoding designs: We developed an interactive sandbox prototyping environment that allowed us to rapidly explore the design space of visual encodings, which we describe in Section 6. This phase of the project lasted about four months. In addition to weekly design feedback sessions with our collaborators, we conducted two 60–90 minute design feedback sessions with the two power users mentioned above, as well as two feedback sessions with energy workers that we had not previously spoken to.

We continued with the method of slide decks as research artefacts. For each session with an external energy worker, we created a personalized slide deck that included screenshots from our sandbox environment along with explanations; these slides were sent to energy workers in advance. Recognizing the importance of showing project stakeholders their own data [24], these slides featured data from the energy worker’s own portfolio of buildings. Since many of the energy workers that we consulted with are located in other cities or countries, many methods for participatory design and evaluation methods that depend on the researchers and stakeholders being co-located [15, 26] could not be implemented [4]. During these sessions, we shared our screen and conducted *chauffeured workflows* using our visualization sandbox: we would present the energy worker with data from their own portfolio and ask them to step through their energy analysis workflow with our alternative visual encoding designs. These sessions were recorded for further analysis and their feedback was later transcribed as annotations on the session’s slide deck. The result of this phase was the identification of a set of *matches* and *mismatches* between visual encoding idioms and individual tasks, which we discuss in Section 7.

Prototyping workflows: Based on feedback collected during the chauffeured workflows with energy workers, we prepared storyboards of these workflows using sandbox screenshots and mockups. We then continued our design process by considering how to best juxtapose, link, and sequence multiple visualizations, all while continuing to consult with our collaborators and energy workers. This phase of the project resulted in the workflow described in Section 8 and realized in the redesigned Energy Manager described in Section 9.

Example artefacts: We generated 11 slide decks containing a total of 302 slides over the course of this project. We include example slides in the supplemental material, along with other research artefacts.

3 ABSTRACTIONS

Following the work domain analysis phase, we recast activities from domain-specific terminology to data and task abstractions.

3.1 Data Abstraction

The energy workers to whom we spoke oversee the energy performance of dozens to hundreds of buildings, which are referred to as *portfolios*. We now abstract a portfolio of buildings and its associated time-oriented energy data, which we summarize in Table 1.

Building metadata: Consider a university, where buildings vary by area, age, and number of occupants. They can also be differentiated using any number of categorical *tags*, such as the type of building (“lecture hall”, “laboratory”), its campus or department (“chemistry”, “physics”), or the name of its building operations manager.

Building groups: Given all of this building metadata, we can have groups of buildings with shared attribute values or ranges. A portfolio may have many building groups, and they may overlap.

Multiple time series per building: The energy performance of each building in a portfolio is monitored over time, and each building has multiple time series associated with it. A building may consume several forms of energy, such as electricity, natural gas, or steam. Many

Term	Abstraction	Example
<i>Building metadata</i>		
Building ID	unique categorical	#123
Building area	quantitative	450 m ²
Building age	quantitative	20 years
# occupants	quantitative	50 people
Location	spatial	49.25° N, 123.10° W
Tag	categorical	“restaurant”
<i>Temporal data for each building</i>		
Energy demand	quantitative	200 kW
Outdoor temperature	quantitative	18° C
Open / closed	categorical	Open Mon–Fri, 08–18h
<i>Derived temporal data for each building</i>		
Consumption	quantitative	800 kWh
Energy intensity	normalized quantitative	1.78 kWh / m ²
Weather-independent performance [†]	normalized quantitative	50 kWh / HDD [‡]
Predicted perform. [†]	quantitative	190 kW
% Savings	normalized quantitative	40%
Rank	ordinal	1st, 2nd, 3rd

Table 1. Data abstraction summary. [†]Performance could be assessed in terms of *demand*, *consumption*, or *intensity*. [‡]Heating Degree Day is one of several approaches to normalizing energy performance using weather data; a full discussion of them is beyond the scope of this paper.

non-residential buildings are equipped to report raw energy demand values every 15 minutes, along with outdoor air temperature. Finally, building opening and closing hours are also recorded.

Derived data: Raw continuous energy *demand* data is typically examined during detailed investigations of single buildings at fine granularities of time; for instance, a building might exhibit an unexpected spike in demand on a Sunday morning. However, for an energy worker overseeing a large portfolio of buildings, analysis tasks typically revolve around *derived* and *aggregated* data, rather than raw continuous data. These may include *averages*, *minimums*, and *maximums* for different temporal granularities of interest, such as the average weekday electricity demand in January. *Consumption* is the most common of these derived attributes, which is an accumulated amount of energy. *Intensity* is consumption normalized by a building’s area, which allows the energy worker to directly compare the energy performance of buildings of different size. Similarly, it is possible to normalize energy performance values by considering outdoor temperature, which allows the energy worker to compare buildings at different times of the year, or buildings in locations with different climates. *Predicted energy performance* based on statistical models is also considered, though predicted values are problematic for reasons we describe in Section 4. *Relative* and *absolute differences* between the observed energy performance and the predicted or historical performance are also considered; for example, the energy worker can determine how a building is performing this year relative to its performance last year. Finally, given any of these derived values, the energy worker can compare multiple *rankings* of buildings, allowing them to identify, for instance, the buildings most in need of an energy conservation measure.

In summary, we have a multidimensional table of buildings and building attributes, along with quantitative energy-related attributes with values changing over time.

Domain convention: The energy workers to whom we spoke are accustomed to interpreting any encoding that incorporates a continuous line graph as raw energy *demand*, and thus it is inappropriate and potentially misleading to use such encodings to display derived and aggregated time series values such as *energy consumption* or *intensity*. In the existing Energy Manager tool, derived and aggregated values are encoded in bar charts or listed in tables; we will discuss in Section 4 how these encodings do not scale, and in Section 7, we examine alternative encodings.

3.2 Task Abstraction

Energy workers need to balance reducing costs and conserving energy while ensuring the comfort and safety of people who use buildings in

their portfolio. To achieve these goals, the energy workers need to: (i) assess the performance of buildings in their portfolio following the implementation of an energy conservation measure, such as installing new windows or lighting; (ii) determine which buildings in their portfolio require energy conservation measures; and (iii) find and diagnose anomalous energy performance such as spikes, outages, surges, or otherwise erratic and inconsistent behaviour.

We characterize these activities as abstract tasks according to a typology that distinguishes between *action* and *target* terms [33], one that builds upon a previous typology [5]; it can be helpful to think of actions as verbs and targets as nouns. Each of the tasks can be described by a need to *discover*: to generate and verify hypotheses. While energy workers also occasionally *present* their findings to colleagues and other stakeholders, our current focus is predominantly on the discovery process.

T1 / Overview: An individual will *lookup* and *summarize* time-oriented data from all the items spanning a coarse period of time. Of interest are trends, outliers, distributions, extreme values, and similarities. This abstract task corresponds to the domain activities of assessing overall energy performance and determining which buildings in a portfolio require energy conservation measures. A concrete example of this task would be an energy worker who asks: “how did the entire university perform this past year? Do any buildings stand out?”.

T2 / Drill Down: The second task is the result of drilling down from the portfolio to a group of buildings and to a finer temporal resolution, examining energy performance in detail: a person *locates* and *compares* trends, outliers, and features in time-oriented data for items in the group. This abstract task corresponds to the domain activities of assessing building performance following an energy conservation measure, or finding and diagnosing anomalous energy performance, wherein a feature of interest could be a spike or sudden outage. One concrete example is “are my restaurants in Vancouver performing better this January than they did last January?”.

T3 / Roll Up: The third task is the converse of drilling down from the portfolio to a group and from the group to individual buildings. This task is one in which a person *explores* and *locates* trends, outliers, and features in time-oriented data to *identify* the proportion of individual behaviour relative to the group’s behaviour, or

the dependency between the aggregate amount and the individual contribution. Concrete examples of this task include “what proportion of a university’s energy consumption is consumed by its computer science building over time?” or “which science faculty building consumes the largest proportion of energy?”.

Task sequences: These three tasks are not performed in isolation, but in sequential energy analysis workflows, and some task sequences occur more often than others. We know from our interviews that the frequency of this work varies, so the Overview (T1) task may be more prevalent among energy workers who infrequently analyze their portfolio’s energy performance. Others may skip the Overview task because they have a specific question about a group of buildings, proceeding directly to the Drill Down (T2) or Roll Up (T3) tasks. An energy worker will alternate between the Drill Down and Roll Up tasks, as the completion of one task may prompt new questions.

4 EXISTING TOOL

We now analyze the existing Energy Manager tool and determine the extent to which the tasks are supported.

Energy Manager, shown in Figure 1, is a multiple-page web-based application, one that uses a small number of visual idioms adhering to the domain convention mentioned in Section 3: bar charts and tables are used for derived aggregate values such as *consumption*, while line charts are used for continuous values such as *energy demand*. The main page presents a summary dashboard for a portfolio of buildings (Fig. 1a). At the bottom of this dashboard is a sortable table listing *consumption*, *intensity*, as well as aggregate *savings* values for the currently selected time period.

Line and bar charts such as those in Figures 1b and 1c are indexed on the other pages, similar to how a Microsoft Excel workbook has sheets, which can in turn contain multiple charts. Unlike Excel, Energy Manager’s visualizations provide some interactivity: the user can zoom, pan, and reveal values upon mouseover. However, none of the visualizations are directly linked to one another or to the dashboard, so the energy worker must navigate between them manually.

Task support: Energy Manager only partially supports the Overview (T1) task: with the portfolio dashboard (Fig. 1a), the energy worker can observe the aggregate *consumption* for a portfolio over time, or alternatively she can observe single aggregate values for individual

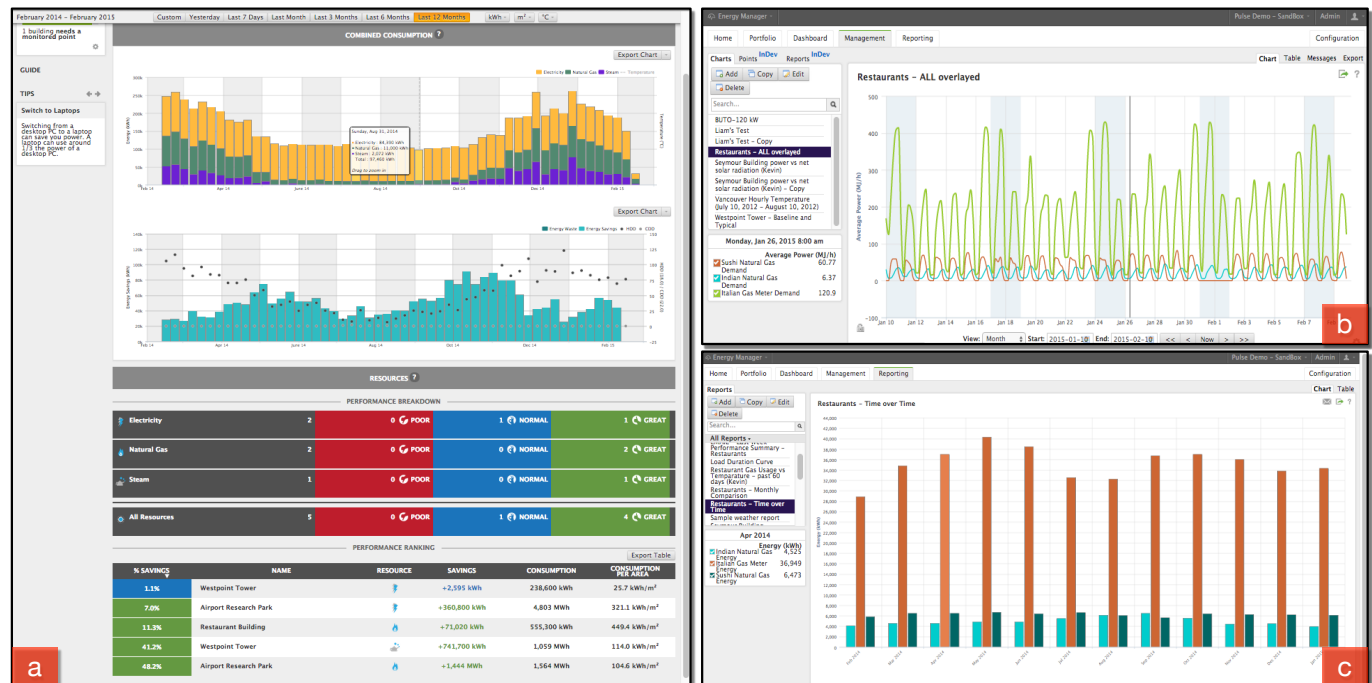


Fig. 1. Energy Manager, our collaborators’ existing energy analysis tool. (a) A dashboard for a portfolio of buildings. (b) A superimposed line chart of energy demand and (c) a grouped bar chart of energy consumption (bottom) for a group of three restaurant buildings within this portfolio.

buildings in the sortable table, but she will not be able to directly compare how individual buildings vary over time. As a result, the energy workers that we interviewed essentially ignored this dashboard, and none of them had found the sortable table to be useful.

The Drill Down (T2) task is supported but the current approach used by Energy Manager does not scale. The examples in Figures 1b and 1c respectively display *energy demand* and *consumption* performance for three restaurants. Since superimposed line charts and grouped bar charts are limited by the number of discriminable colours, these visualizations are inappropriate for the energy worker who needs to consider more than a handful of buildings.

The Roll Up (T3) task is not explicitly supported. Though it is possible to estimate the proportion of a building's energy performance relative to its group with bars and lines, this process is error-prone.

Task sequences not supported: Because the visualizations in Energy Manager are not coordinated or linked in any way, it is difficult and tedious to alternate from Overview (T1) to Drill Down (T2) and Roll Up (T3) Tasks. As in Excel, the energy worker will have to locate an existing visualization or specify a new visualization using a wizard dialog; if the bar or line chart for a set of buildings does not already exist when the energy worker needs it, she has to create it. By the time she has created it, she may have forgotten her objective.

Limited filtering and aggregation: The bar and line charts in Energy Manager allow an energy worker to hide or show individual buildings. However, the energy worker cannot filter buildings with shared attributes, such as filtering a portfolio of buildings to only show restaurants. Similarly, it is impossible to aggregate buildings together when they share attributes: for instance, the energy worker cannot compare the aggregate energy performance of restaurants in one city to those in another city.

No faceting: Aside from the seldom-used portfolio dashboard shown in Figure 1a, there is no faceting or juxtaposition of visualizations in Energy Manager: the energy workers' workarounds included opening multiple browser windows, adjusting the line or bar charts to display the same scale and ranges, and tiling these windows manually across their monitor. Similarly, some energy workers printed and taped charts together to accomplish the same result.

Summary: Due to these limitations, energy workers can only observe narrow slices of their portfolio data, or they are presented with aggregate data that is too coarse to be useful. In addition, they did not trust Energy Manager's derived predicted values based on statistical models, and would have preferred to compare observed energy performance to historical values. In other words, the derived and aggregated values currently shown may hide information such as extreme values and unusual distributions. As a result of this loss of detail, energy workers routinely export tabular data from Energy Manager and import it into Excel, with which they would conduct a custom analysis. Finally, many of the energy workers to whom we spoke remarked that energy analysis software tools built by our collaborators' competitors [10, 34, 37] have limitations similar to those of Energy Manager. Altogether, these problems may help to explain why the number of energy workers who actively use Energy Manager is substantially less than the number of user accounts.

5 RELATED WORK

We now review relevant previous work, beginning with work in the energy domain. We also discuss visualizations of time-oriented data in other domains, as well as evaluation studies that assess the effectiveness of visualization idioms for similar data and task abstractions.

Visualization in the energy domain: Technology that allows for the continuous measurement of a building's energy demand is becoming increasingly available, and several techniques to monitor and present this data have recently been proposed, especially for residential buildings [9, 11, 15, 35]. Erickson et al. [11] developed a web-based residential energy dashboard for homeowners, allowing them to compare against their neighbours with familiar bar and line charts. However, such a dashboard would be unsuitable for the analysis work of an energy worker who oversees a portfolio of many buildings.

Though bar and line chart depictions of energy data are most common, other visual encodings have also been employed, from abstract and artistic ambient visualizations [35] to a compelling calendar-based visualization [42], in which calendar dates with similar energy behaviour are visually associated using a common colour. We also explore visual encodings beyond bar and line charts, and in Section 7.3, we consider how to encode data from multiple buildings using calendars. Another approach to summarizing the energy behaviour of multiple buildings is to use map-based visualizations [18, 25]; we discuss the limitations of that approach in Section 7.4.

More closely related to our work is that of Goodwin et al. [15], who presented several prototype visualizations of modelled residential energy use across thousands of households at the scale of individual household appliances. The domain activities they address overlap partially with those performed by the energy workers we spoke to, such as the need to find anomalous energy performance across many buildings; another activity they address, in which energy modelers perform energy load-shifting simulations to estimate potential savings, is not an activity performed by our energy workers. Their designs incorporated several visual encodings not typically seen in the energy domain: horizon charts [19], boxplots [45], and matrix-based encodings. However, the focus and main contribution of their paper was on creative methods for visualization requirements analysis, rather than on a thorough analysis of the visualization designs themselves. In our work, we reexamine some of these visual encodings, among others, and evaluate their effectiveness in the context of our data and task abstractions.

Visualizing multiple time series: Many techniques for visualizing time-oriented data have been proposed, and Aigner et al.'s 2011 survey [1] of these visualizations has provided us with a structured way to think about this design space. In their terminology, our designs incorporate *linear* and *cyclic* encodings of time, depicting *abstract multivariate interval* data.

Another axis on which we can analyze existing techniques is the number of time series being considered. At the low end of this continuum, superimposed line charts or grouped bar charts are appropriate for a small number of time series. In the middle of this continuum, faceting techniques such as small multiple line charts, horizon charts [19] and matrix-based encodings [17, 40] are appropriate. At the high end of this continuum, dense multi-form faceting techniques and those that aggregate time series together are appropriate when dealing with thousands of time series, such as in LiveRAC [27] or Line Graph Explorer [22]. Since we are addressing portfolios of dozens to hundreds of buildings, we position our designs toward the middle of this continuum, and we evaluate faceting and matrix-based approaches in the following section.

Evaluation of time-oriented data visualizations: We also situate our work with regards to experimental studies [2, 8, 14, 21] that have examined the viability of alternative visual encodings for abstract tasks similar to those that we characterized: identifying and comparing averages, trends, extreme values, and outliers. Some studies address the viability of encodings for a single time series [2, 8], while others consider multiple concurrent time series; one study considered up to sixteen time series [21], while another considered forty-eight [14]. We too assessed the viability of different visual encodings for multiple concurrent time series; however, our approach involves a qualitative evaluation (see Section 2), as opposed to a controlled experiment.

Moreover, while these studies considered continuous time series data, we must consider alternative scalable encodings, since in our case domain conventions dictate that continuous line graph encodings would be misleading for the display of derived and aggregated time series values, as discussed in Section 3.1.

6 PROTOTYPING ENVIRONMENTS

Shiny sandbox: We developed an interactive browser-based visualization design sandbox¹ to produce *data sketches* [24], shown in Figure 2. The sandbox allowed us to rapidly prototype different visual en-

¹<http://mattbrehmer.shinyapps.io/PortfolioSandbox>

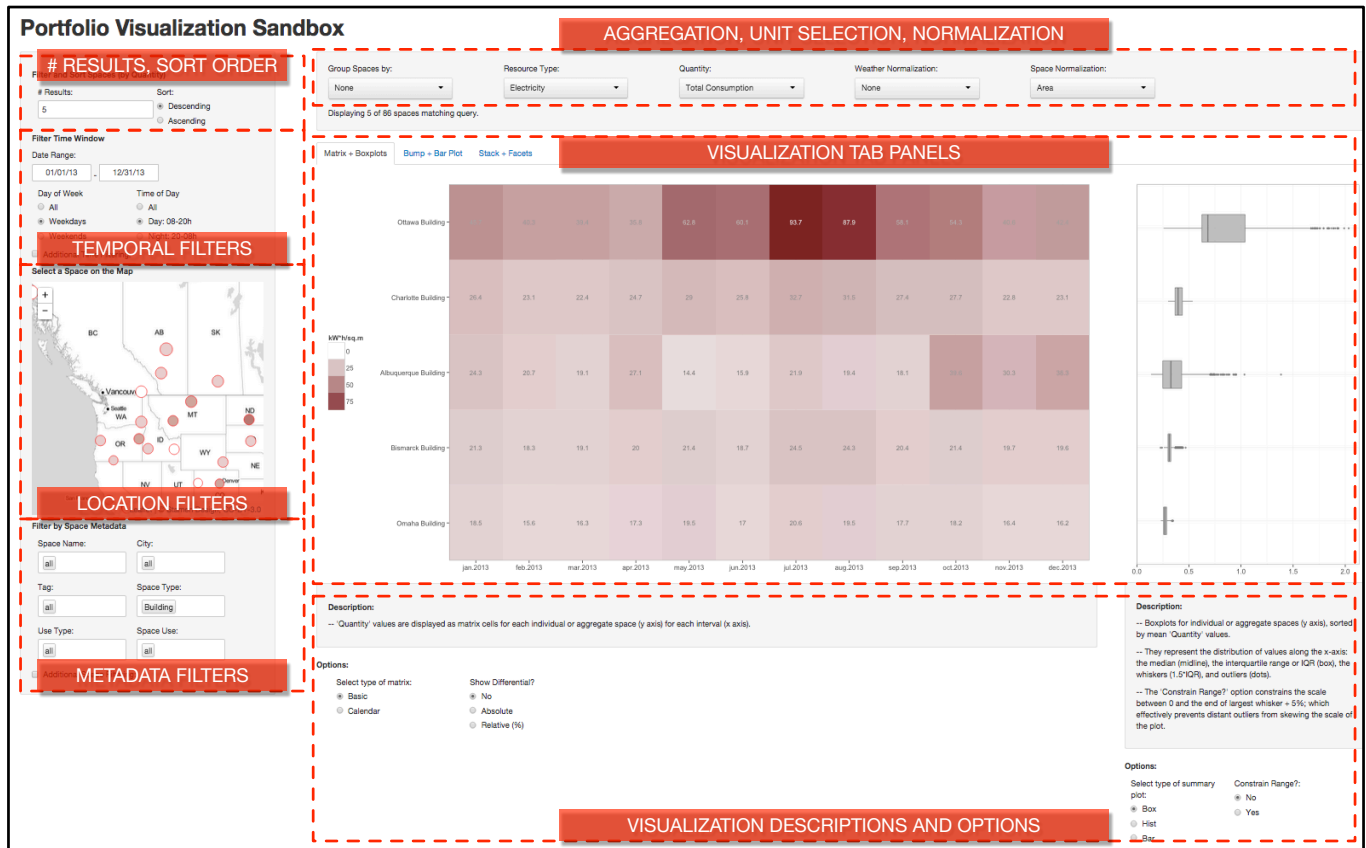


Fig. 2. A sandbox design environment for visualizing energy data from a portfolio of buildings. Designs depicted in Figures 3, 4, 5, and 7 were also produced within this environment. A matrix of aggregate *energy intensity* values with auxiliary boxplots is shown for 5 (of 86) buildings, those with the highest *intensity*. Client portfolio data has been anonymized by changing building names and location; all other data is real.

coding designs and conduct chauffeured workflows with energy workers, a process that we described in Section 2. All of the designs discussed in Section 7 were produced within this environment, which was developed² using the Shiny web application framework [36].

Central to this sandbox environment are interactive controls for sorting, filtering, and aggregating buildings, controls for selecting units of interest such as *demand* or *consumption*, as well as controls for area and weather normalization; recall that there were no such controls in the Energy Manager interface. Whenever these controls are adjusted, we sort the filtered set of buildings according to the currently selected energy unit of interest and time span. The sandbox operator can choose the sort order and select the number of results to show. For instance, Figure 2 shows a view described in Section 7.3, and in it we show 5 buildings from a geographically anonymized portfolio of 86 buildings, those with the highest *energy intensity* in 2013.

D3 interactive prototypes: Since our Shiny-based sandbox implementation did not allow us to directly experiment with interactions involving coordinated selection across juxtaposed views, we developed several interactive prototypes using D3 [3] that specifically address this coordination; these prototypes are discussed in Section 8 and one of them³ is shown in Figure 6.

7 VISUAL ENCODING MATCHES AND MISMATCHES

The Nested Blocks and Guidelines Model [30] describes a need for guidelines that relate the domain, abstraction, idiom (technique), and algorithm levels of visualization design [32, 33]. In Section 3, we described the relationship between domain activities and the data and task abstractions. In this section, we consider the space of visual encoding idioms and present guidelines for matching idioms to abstractions. Since the space of possible visual encoding idioms for

time-oriented data is large [1], we undertook a typical design study approach [39]: considering several idioms, implementing a subset of them, and selecting only the few good matches.

We identified five *matches* {✓} between visual encoding idioms and the combination of data and task abstractions, based on evidence resulting from our process described in Section 2. We also identified four *mismatches* {✗} and two *potential matches* {?}. These matches and mismatches, summarized in Table 2, serve as generalizable guidelines that are transferable beyond the energy domain, especially when we consider the similarity between our abstract tasks to those addressed in domain-agnostic evaluation studies [2, 21]. Furthermore, these matches and mismatches fill a gap with regards to identifying suitable visual encoding idioms for multiple time series in which values are not continuous, but derived and aggregated values that should *not* be encoded as line charts.

Task	Visualization Idiom	Match?
T1: Overview	Faceted bar chart	✗
	Bump plot	✗
	Bar + bump plot	?
	(Calendar) matrix	?
	Map	✗
T2: Drill Down	Juxtaposed matrix and boxplots	✓
	Faceted bar chart	✓
	Faceted boxplot	✗
	Faceted line graph	✓
T3: Roll Up	Stacked area graph	✓
	Stacked bar chart	✓

Table 2. A summary of the *matches* and *mismatches* between abstract tasks and visual encoding idioms.

²<http://github.com/mattbrehmer/PortfolioSandbox>

³<http://bl.ocks.org/mattbrehmer/287e44c9a12151967874>

7.1 Faceted Visualizations for Overview and Drill Down

We initially thought that faceted “small multiple” visualizations would be a good match for *both* the Overview (T1) and Drill Down (T2) tasks, in that they provide a scalable alternative to grouped bar charts and superimposed line charts.

Faceted bar charts: a mismatch {✖} for the Overview (T1) task, yet a match {✓} for the Drill Down (T2) task. Faceted bar charts were among the first designs that we considered, especially after one energy worker provided us with his own mockup of such a design. However, if an energy worker has dozens or hundreds of buildings in their portfolio, faceting is unlikely to scale [21]. We determined that it was a poor match for the Overview (T1) task, though a match for the Drill Down (T2) task, provided that the energy worker has already filtered down to a smaller group of buildings, such as filtering a university portfolio to show only the “laboratory” buildings. In addition, one of the power user energy workers lamented that bar charts only show coarse aggregate values, typically an average or a sum, and as a result of this loss of detail, they do not show other aggregate values of interest, such as ranges or extreme values.

Faceted boxplots: a mismatch {✖} for the Drill Down (T2) task. We expected that faceted boxplots would allow energy workers to compare ranges, distributions, and extreme values for multiple buildings at different points in time, such as in Figure 3. However, despite the long history of boxplots [45] and support from influential visualization practitioners [12], we found that most energy workers are not familiar with boxplots, except for a minority who had taken a post-secondary statistics course. Furthermore, comparisons in faceted boxplots are more difficult than in faceted bar charts, where positions are aligned to each facet’s baseline; with faceted boxplots, the observer must compare multiple positions and widths across separate facets. Our design was therefore a daunting introduction to boxplots for those unfamiliar with them and a poor match for the Drill Down task.

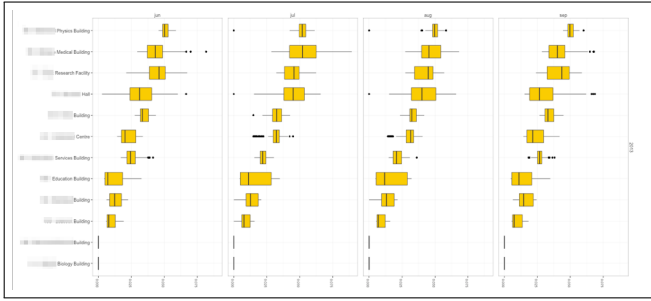


Fig. 3. Faceted boxplots that encode aggregate area-normalized energy demand distributions for 12 buildings across four months, sorted in descending order according to the average demand value for this four month period. A mismatch {✖} for the Drill Down task (T2). Building names are blurred to sanitize real client portfolio data.

Faceted line charts: a match {✓} for the Drill Down (T2) task. Faceted line charts are a good match when observing non-derived continuous quantitative time series values such as energy demand; an example is shown in Figure 7 (bottom). They are a scalable alternative to superimposed line charts [21] and the line chart encoding is already very familiar to energy workers. As mentioned above in Section 3, line charts are *not* appropriate for derived and aggregated values such as energy consumption or intensity.

7.2 Rank-Based Overview Visualizations

As faceting seemed unlikely to be effective for the Overview (T1) task, we considered non-faceted visualizations of aggregate values. Recall how the sortable table in Energy Manager’s portfolio dashboard (Figure 1a, bottom) was never used for the Overview task; it contained only coarse aggregate values for each item, providing little detail about temporal variation. We therefore experimented with encodings for displaying rank as well as rank change over time.

Bump plots: a mismatch {✖} for the Overview (T1) task. Bump plots encode rank and rank change; they incorporate a familiar line encod-

ing across equally-spaced temporal intervals [41]. However, as with superimposed line charts, it becomes difficult to distinguish individual items using colour. One possible solution is to highlight items that vary in rank, rather than requiring the observer to locate these items. Another problem is that bump plots only show relative rank and rank change, whereas the absolute values that produce these ranks are not shown. As a result this a loss of detail, the bump plot is also a poor match for the Overview task.

Bump + bar plots: a potential match {?} for the Overview (T1) task. We next considered an encoding that incorporates relative rank, rank change, and absolute value, by adding familiar bars to each series in the bump plot, as shown in Figure 4. This approach is similar to two recently proposed techniques that encode both relative rank and absolute value [16, 20]. With this design, we still face the scalability problem associated with colour discriminability. A combination of interaction and highlighting rank variation may facilitate this discriminability; in Figure 4, rank variation is encoded using the alpha channel, so the pink series that varies considerably over time is most salient.

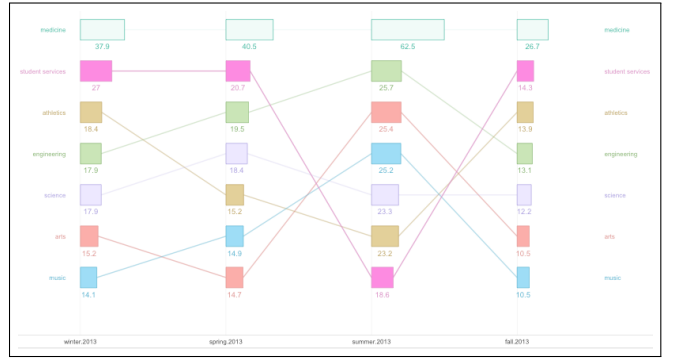


Fig. 4. A bar + bump plot of energy intensity, encoding rank change for the top 7 building groups (buildings aggregated by tag) across four seasons. The alpha channel encodes rank variation to highlight inconsistent buildings. A potential match {?} for the Overview task (T1).

Energy workers responded positively to this visualization, as it is comprised of familiar bar and line encodings. However, despite this positive response, we discovered that *rank*s as derived values are actually infrequently considered during energy analysis, and that they tend to be more appropriate for annual planning and presentation activities, such as determining how to prioritize energy conservation projects, and less so for recurring analysis and monitoring activities. Thus, the hunt for a match for the Overview (T1) task continued.

7.3 Matrix-Based Overview Visualizations

Time series matrix: a potential match {?} for the Overview (T1) task. Matrix encodings are scalable and space-efficient [15, 17], as can be seen in the center of Figure 2. Matrix encodings allow us to display observed as well as differential values, allowing an energy worker to review energy savings relative to predicted or historical values; a matrix displaying differential energy data is shown in Figure 5. Most of the energy workers that we interviewed were unfamiliar with this form of encoding, except one who routinely made such visualizations in Excel. As a result, it took more effort to convince our collaborators of the value of these matrix-based encodings for the Overview task.

We initially referred to these visual encodings as “heat maps”, but this visualization term led to considerable confusion because of conflicting domain conventions with the energy-related meaning of *heat* and expectations raised by the word *map*: this encoding does not show energy solely used for heating, nor does it show the geographic location of buildings⁴. We also learned that energy workers found matrices with diverging colour scales easier to interpret than those with unidirectional colour scales. Finally, we found that while red is fine for use in diverging colour scales, as it has a negative connotation, it is inappropriate for unidirectional colour scales in this context. As a

⁴This confusion is not unique to the energy domain [13, 46].

result of this mixed response to matrix-based encodings, we realized that more work needed to be done.

Calendar matrix: a potential match {?} for the Overview (T1) task. We altered our matrix encoding by partitioning the cells corresponding to months into calendars (Figure 5), a design decision inspired by previous work [23, 42]. Energy workers responded positively to this encoding, which helped to resolve the unfamiliarity of the more generic matrix encoding. However, months and days are not the only time granularities of interest, so this encoding may not be appropriate for all time ranges.

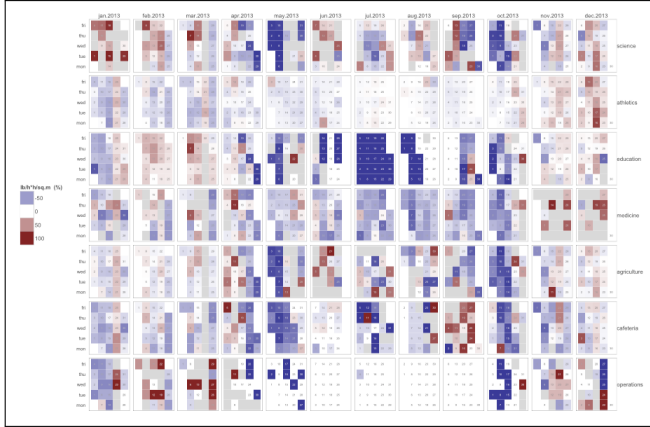


Fig. 5. A time series calendar matrix of *energy intensity savings* for 7 building groups (buildings aggregated by shared categorical tag), relative to historical values (blue = *savings*, red = *higher than historical intensity*). A potential match {?} for the Overview task (T1).

7.4 Map-Based Overview Visualizations

Maps: a mismatch {✖} for the Overview (T1) task. We explored the use of maps based on their popularity in previous work [18, 25]. We conjectured that maps may be appropriate for buildings in a shared neighbourhood, such as a university campus, even though that encoding may be less appropriate for building portfolios spanning large geographic areas. However, after interviews with energy workers, we realized that maps are better suited for presenting coarse aggregate summary values of energy performance to a casual observer, and they are less appropriate for recurring analysis work; to view energy behaviour varying over time, animating or faceting the map would be necessary. Furthermore, an energy worker overseeing a portfolio of buildings is already likely to be familiar with the locations of buildings in her portfolio, and their relative locations are not informative. While using a map to encode energy data was found to be inappropriate for the tasks that we characterized, an interactive map may be an effective means to filter a portfolio by building location, which we considered in our sandbox environment shown in Figure 2.

7.5 Stack-Based Roll Up Visualizations

Stacked area charts & stacked bar charts: matches {✓} for the Roll Up (T3) task. The obvious visualizations of stacked area [6] and stacked bar charts were indeed matches; an example of the former is shown in Figure 7 (top). Stacked area charts are appropriate when considering *energy demand* values, while stacked bar charts are appropriate for derived and aggregated values such as *energy consumption* or *intensity*. For both of these encodings, differentiating individual time series can be accomplished by interactive highlighting [44], as using hue to differentiate stack elements will not scale.

8 WORKFLOW DESIGN WITH MULTIPLE VIEWS

Our design discussion up to this point has focused on visual encoding choices for single views; we also want to stress the importance of interaction and workflow design, which involves juxtaposing and linking multiple views.

Juxtaposed matrix and boxplots: a match {✓} for the Overview (T1) task. One reason to juxtapose views is to support a single task with complementary data. None of the encodings discussed thus far are a clear match for the Overview task, although the time series matrix designs described in Section 7.3 showed promise. A problem with the matrix encodings is that they display coarse aggregate values only, such as averages or sums for each matrix cell. Recall from Section 7.1 that an energy worker made the same remark about faceted bar charts. Partitioning a matrix into calendars is one way to show a finer resolution in the same amount of space, however this encoding will not always be appropriate: for instance, an energy worker may be interested in a time span shorter than a month, or longer than several years. The alternative that we developed involves complementing and reinforcing the aggregate values in the original matrix design by juxtaposing single boxplots that encode ranges and distributions for each time series, as shown in Figure 2. Though boxplots remain unfamiliar to energy workers, these auxiliary boxplots are easier to interpret than faceted boxplots (c.f. Figure 3), as they require no comparison of length or width across separate facets. We reflect further upon the balance between familiarity and the use of auxiliary charts to combat information loss in Section 10.1.

We then sought a better way to coordinate and link the matrix and its juxtaposed auxiliary boxplots. We created several prototypes, and one is shown in Figure 6. The interactive linked highlighting in this prototype served both to promote engagement with these juxtaposed visualizations and to facilitate the learning of these visual encodings, which were previously unfamiliar to energy workers.

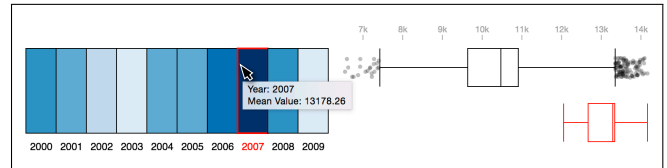


Fig. 6. An interactive auxiliary boxplot prototype: boxplots corresponding to the brushed time period are shown alongside the boxplot for the entire time series.

Juxtaposed stack and facets: Another reason to juxtapose views is to support fast alternation between tasks. Recall that the Drill Down (T2) and Roll Up (T3) tasks are often performed in alternation, and we were concerned about the loss of context when switching between stacked bar or area charts and faceted bar or line charts. To prevent this loss of context, we juxtaposed the stacked chart with the faceted charts, and provide linked highlighting between elements in the stack and those in the facets, as shown in Figure 7; as a result, both the Drill Down and Roll Up tasks are supported in a single display. Currently, four facets are shown in a row, with additional facets wrapping to subsequent rows, sorted in the same order as the elements in the stack.

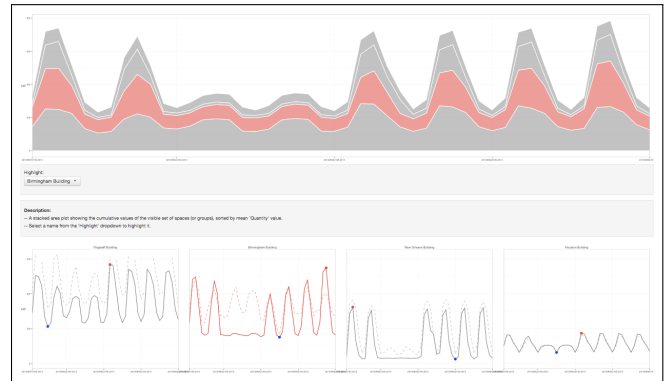


Fig. 7. A stacked area chart of *energy demand* data for 4 library buildings, juxtaposed alongside faceted line charts of the same data. The same building is highlighted in red in both the stack and the facets.

Sequenced view navigation: Recall that the Drill Down (T2) and Roll Up (T3) tasks involve fewer buildings and finer temporal resolutions than the Overview (T1) task, which has broader scope; thus, it

would be inappropriate to juxtapose the Overview visualizations with the Drill Down and Roll Up visualizations in a single display. Instead, we considered how an energy worker would navigate between these views shown on separate displays. Beginning with the matrix and auxiliary boxplots, the energy worker can perform the Overview task, select units of interest, and filter or aggregate buildings in the portfolio. If the currently selected unit of interest is *consumption* or *intensity*, selecting a column of the matrix directs the energy worker to juxtaposed faceted and stacked bar charts that include every building from the matrix and spanning the time period corresponding to the selected column. For *demand* data, the energy worker is directed to faceted and stacked line charts. At this point, the energy worker can perform the Drill Down and Roll Up tasks in alternation. We demonstrate this multiple-view workflow in the supplemental video.

Finally, we also envisioned drilling down further to individual buildings. If the energy worker selects a cell or row of the matrix, she will navigate to a single bar or line chart for the corresponding building and time period.

9 RESULTS

We are pleased to report that our collaborators adopted a number of our visualization designs into a new version of Energy Manager, shown in Figure 8, which will soon be deployed to their large client base. Their client base is also expected to grow dramatically as a result of their recent partnership with a large utility company: tens of thousands of energy workers will now be able to analyze the energy performance of their building portfolios with the redesigned Energy Manager.

As in our sandbox environment, options for filtering and aggregating buildings according to shared categorical tags are now prominently and persistently shown in the menu at the top of the interface. The interface also provides the ability to select units of interest and compare observed energy values against trusted historical values, as an alternative to comparing observed values to less trusted predicted values,

overcoming a limitation of the original Energy Manager.

Coordinated matrix and boxplots: a variant of our design described in Section 8 has been incorporated into the redesigned Energy Manager; the matrix is labeled as a *Site Overview* to avoid the *heat map* misnomer, and corresponds well with our characterization of the Overview (T1) task. The number of buildings shown depends on the window size, and more buildings appear as the energy worker scrolls. Our collaborators did consider the alternative calendar-based encoding, but ruled it out based on a requirement that arose late in the design process: that the redesigned Energy Manager be accessible on a tablet device. Partitioning the matrix cells into calendars may result in calendar dates too small to be selectable without zooming, which may incur a high implementation cost. Meanwhile, the boxplots, now referred to as *Value Range* visualizations, update when the energy worker brushes the matrix by hovering over a cell in the corresponding matrix row, similar to the behaviour of the prototype we described in Section 8. Unlike our earlier prototype, a single boxplot is shown instead of showing one boxplot for the entire range and another boxplot for the brushed time period; when no time period in the matrix is brushed, the boxplot for the entire time series is shown.

Interactive workflows realized: Another critical improvement over the original Energy Manager is the ability for an energy worker to drill down from a row, column, or cell of the matrix to stacked, faceted, superimposed, or individual line and bar charts, as shown in Figure 8. The selected energy unit is retained across these transitions, so faceted line charts and stacked area charts are used for *demand*, while faceted bar charts and stacked bar charts are used for *consumption*. In faceted views, individual facets can be manually resorted via drag and drop. Stacked and faceted visualizations are currently shown separately; our collaborators are considering juxtaposing stacked and faceted visualizations, such as in the design described in Section 8, which would allow for an uninterrupted alternation between the Drill Down (T2) and Roll Up (T3) tasks in the same display.

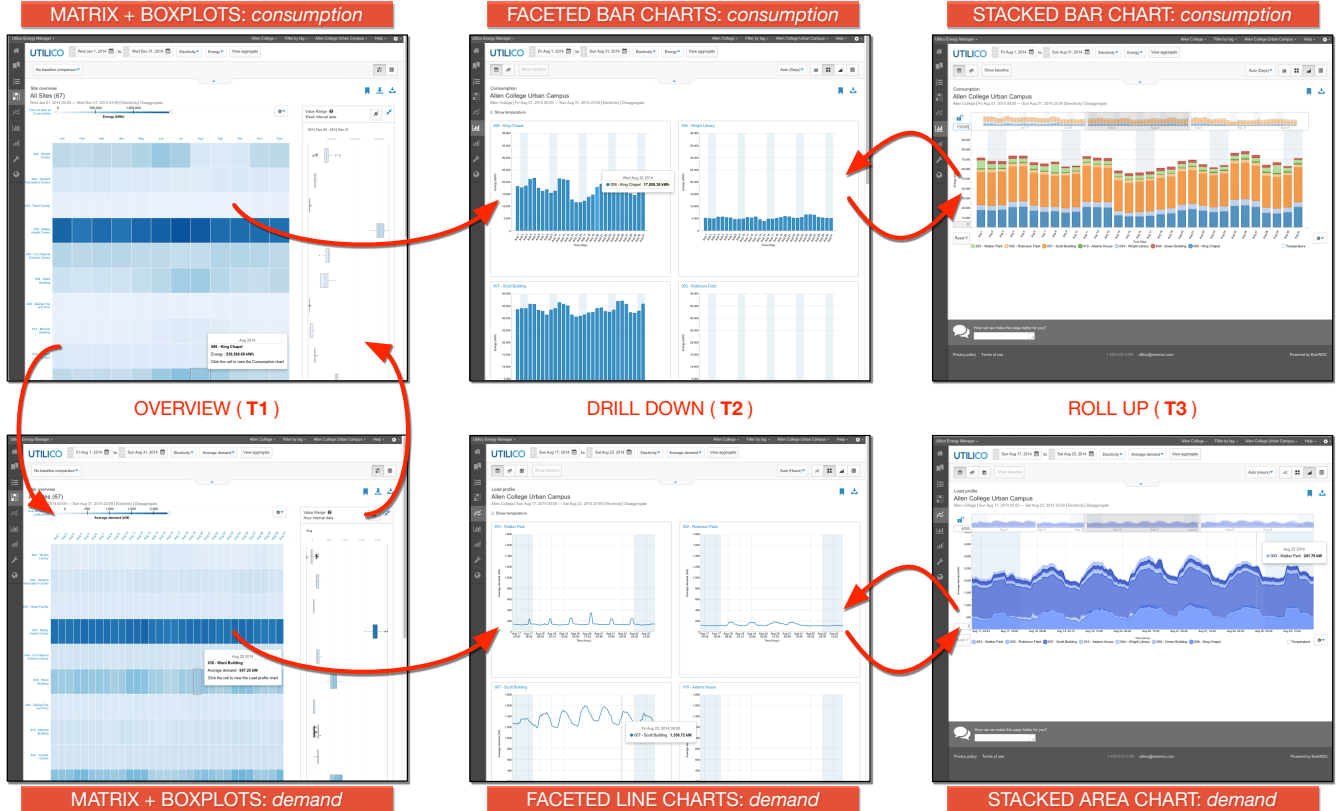


Fig. 8. The redesigned *Energy Manager* that incorporates many aspects of our prototype designs. On the left, the *Site Overview* (a time series matrix) is juxtaposed with coordinated *Value Range* (boxplot) views. An energy worker can easily switch between units such as *energy consumption* or *energy demand* and filter or aggregate the set of buildings to those that share a common categorical tag; by selecting a column of the matrix, she can drill down to faceted or stacked visualizations of *consumption* (top middle, top right) or *demand* (bottom middle, bottom right).

10 DISCUSSION

We now step back from specific aspects of visualization design for time-oriented data to discuss higher-level guidelines, to reflect upon our methodology, and to suggest future work.

10.1 Guidelines: Familiarity and Trust

In addition to the specific guidelines regarding matches and mismatches between idioms and abstractions described in Section 7, we also propose more general and succinct guidelines relating to the themes of *familiarity* and *trust*.

Familiarity: As with professionals in many other domains, energy workers are accustomed to working predominantly with familiar visual encodings, namely bars and lines. When we introduced them to unfamiliar visualization designs, we learned several things:

Complexity can reduce unfamiliarity: Though counter-intuitive, we learned that adding complexity increased interpretability; the juxtaposition of a matrix and a boxplot, two unfamiliar encodings, together with coordinated interaction and highlighting, received more positive feedback than either of these encodings in isolation. Similarly, energy workers indicated that the more complex diverging colour scales for derived *savings* values in the matrix were easier to interpret than the simpler unidirectional colour scales for observed energy values.

Take care with names: Names for unfamiliar visual idioms can be misleading: “*heat map*” [13, 46] and “*boxplot*” [45] were particular sources of confusion that we encountered. In introducing these unfamiliar visualizations, we primed the energy workers to think in a certain way, prompting expectations about the values being encoded. In hindsight, we could have explicitly solicited visualization names from energy workers based on their own descriptions [29].

Trust: When visualization is used as part of the hypothesis generation and verification process, trust is imperative, especially for derived and aggregated values. Previous work has investigated the trustworthiness of visualizations for text-based data [7], and we now discuss the topic of trust motivated by our design of visualizations of time-oriented data.

Auxiliary charts to combat information loss: When the number of concurrent time series grows large, it is difficult and overwhelming to visualize individual values from each time series; instead, a common approach is to visualize derived aggregate values [27]. This loss of detail is apparent in the original Energy Manager’s portfolio dashboard (Figure 1a) as well as in the cells of our time series matrix (Figure 2). Whenever there is a loss of detail, there is a loss of trust: one of the power user energy workers remarked that these derived aggregate values hide information such as extreme values. In juxtaposing the time series matrix with auxiliary boxplots that update whenever a matrix cell containing an aggregate value is brushed, we are not only restoring lost information: we are also restoring trust.

Promote agency over derived values: In our sandbox environment and in the redesigned Energy Manager, we provided explicit and obvious interactive controls for filtering, aggregation, normalization, and unit selection, controls that were missing in the original Energy Manager. With these controls, we provide energy workers agency over the creation of derived values and these values become more trusted. Similarly, the redesigned Energy Manager includes the option to compare observed energy performance to selected historical values, as an alternative to comparing against predicted values generated by a “black box” statistical model; until there is some visual indication as to how the underlying model algorithms generate these values [31], there will be little trust, and it is therefore preferable to provide the option to compare against trusted historical values.

10.2 Methodological Reflection

Though our overall methodological approach is similar to many other visualization design studies [26, 39], there are some specific aspects of our methodology that are unique to projects executed in company settings [38]: we negotiated access to clients and to their portfolio data at the very beginning of our collaboration, and we encourage researchers considering similar collaborations to do the same. In addition, we also

engaged primarily with remote energy workers, and our methodological decisions described in Section 2 reflect this logistical difficulty [4]. We now take the opportunity to reflect on three other aspects of our methodology:

Work domain analysis: *Worth it, and don’t be daunted.* Conducting a rigorous and systematic work domain analysis can be time consuming and logistically challenging. However, it is helpful to realize that authorities on work domain analysis [43] established their methodologies in the design of high-risk, high-cost systems, such as nuclear power plant control rooms. Work domain analysis and requirements analysis methodologies for many visualization design projects can be more flexible [28] and creative [15]. A thorough work domain analysis need not take a year to complete, and subsequent phases of abstraction and iterative design can be carried out while continuing to develop an understanding of the work domain.

Workflow prototyping: In addition to using our sandbox environment to identify appropriate visual encodings for individual tasks, we also used it as a tool to generate possible *workflows* that support a sequence of tasks. Some visualization research projects stop before this step, but we argue for its importance. We thought that confronting energy workers with a combinatorial explosion of possibilities from a large set of visual encodings and view parametrization options would fall short of a full solution to the problem at hand.

Grounding design decisions: *Document everything, strive to be consistent and systematic.* One collaborator remarked that our approach often confirmed some earlier suspicions rather than introduced major surprises, where the novelty lay in a clear path to design decision-making that was missing before: “*we performed an analysis of [Energy Manager’s] flaws in a systematic way, put a name on them, and then tested with users*”. The exhaustive collecting and analyzing of qualitative data before and during design allowed us to justify the design decisions described in Sections 7 and 8. Presenting our consolidated findings and design justifications in concise and consistent annotated slide decks was highly appreciated by our collaborators. Given this presentation of evidence, our collaborators adopted our designs with confidence, much in the same way that the results of a controlled quantitative experiment can convince stakeholders [38].

10.3 Future Work

This paper focused on the visualization design and evaluation process and how our designs were adopted into our collaborators’ production timeline. In the future, we would like to assess the adoption of the redesigned Energy Manager following its wide-scale deployment, tracking usage over an extended period of time and speaking to more energy workers via interviews and focus groups.

11 CONCLUSION

We conducted a design study in the energy domain, one in which we collaborated with an energy analysis software company and their clients to develop visualizations for analyzing the energy performance of large building portfolios. We described generalizable visualization design choices framed in terms of **matches** and **mismatches** between abstractions and visual encoding idioms that are transferable beyond the energy analysis domain. We also contributed more general guidelines pertaining to the themes of **familiarity** and **trust**, along with **methodological guidance** for visualization design studies. As a result of our research and design process, our visualization designs were adopted by our collaborator into their development of a redesigned commercial energy analysis application that will be deployed to tens of thousands of energy workers.

ACKNOWLEDGMENTS

We received financial support from NSERC and Mitacs. We thank our collaborators at EnerNOC, including client services liaison J. Christopher, as well as C. Crane and the Energy Manager development team. We also thank the energy workers that we interviewed. Thanks to M. Borkin, A. Crişan, J. Dawson, J. Fulda, E. Hoque, S.-H. Kim, N. Mahyar, and J. McGrenere for their feedback on the paper.

REFERENCES

- [1] W. Aigner, S. Miksch, H. Schumann, and C. Tominski. *Visualization of Time-Oriented Data*. Springer, 2011.
- [2] D. Albers, M. Correll, and M. Gleicher. Task-driven evaluation of aggregation in time series visualization. In *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI)*, pages 551–560, 2014.
- [3] M. Bostock, V. Ogievetsky, and J. Heer. D³: Data-driven documents. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis)*, 17(12):2301–2309, 2011.
- [4] M. Brehmer, S. Carpendale, B. Lee, and M. Tory. Pre-design empiricism for information visualization: Scenarios, methods, and challenges. In *Proc. ACM Workshop on Beyond Time And Errors: Novel Evaluation Methods For Visualization (BELIV)*, pages 147–151, 2014.
- [5] M. Brehmer and T. Munzner. A multi-level typology of abstract visualization tasks. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis)*, 19(12):2376–2385, 2013.
- [6] L. Byron and M. Wattenberg. Stacked graphs – geometry & aesthetics. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis)*, 14(6):1245–1252, 2008.
- [7] J. Chuang, D. Ramage, C. D. Manning, and J. Heer. Interpretation and trust: Designing model-driven visualizations for text analysis. In *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI)*, pages 443–452, 2012.
- [8] M. Correll, D. Albers, M. Gleicher, and S. Franconeri. Comparing averages in time series data. In *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI)*, pages 1095–1104, 2012.
- [9] K. Ellegård and J. Palm. Visualizing energy consumption activities as a tool for making everyday life more sustainable. *Applied Energy*, 88(5):1920–1926, 2011.
- [10] Energent EMIS. <http://energent.com/emis/>.
- [11] T. Erickson, M. Li, Y. Kim, A. Deshpande, S. Sahu, T. Chao, P. Sukaviriya, and M. Naphade. The Dubuque electricity portal: Evaluation of a city-scale residential electricity consumption feedback system. In *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI)*, pages 1203–1212, 2013.
- [12] S. Few. Distribution displays, conventional and potential, 2014. Technical report. <http://goo.gl/zy8uAo>.
- [13] K. Field. When is a heat map not a heat map, Feb. 17 2015. Blog post. <http://goo.gl/WnOPXo>.
- [14] J. Fuchs, F. Fischer, F. Mansmann, E. Bertini, and P. Isenberg. Evaluation of alternative glyph designs for time series data in a small multiple setting. In *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI)*, pages 3237–3246, 2013.
- [15] S. Goodwin, J. Dykes, S. Jones, I. Dillingham, G. Dove, A. Duffy, A. Kachkaev, A. Slingsby, and J. Wood. Creative user-centered visualization design for energy analysts and modelers. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis)*, 19(12):2516–2525, 2013.
- [16] S. Gratzl, A. Lex, N. Gehlenborg, H. Pfister, and M. Streit. LineUp: Visual analysis of multi-attribute rankings. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis)*, 19(12):2277–2286, 2013.
- [17] M. Hao, U. Dayal, D. Keim, and T. Schreck. Multi-resolution techniques for visual exploration of large time-series data. In *Proc. Eurographics / IEEE VGTC Symp. Visualization (EuroVis)*, pages 27–34, 2007.
- [18] HEAT (Heat Energy Assessment Technologies) Map. <http://saveheat.co/heat-scores.php>.
- [19] J. Heer, N. Kong, and M. Agrawala. Sizing the horizon: The effects of chart size and layering on the graphical perception of time series visualizations. In *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI)*, pages 1303–1312, 2009.
- [20] I. Hur, S.-H. Kim, A. Samak, and J. S. Yi. A comparative study of three sorting techniques in performing cognitive tasks on a tabular representation. *Intl. J. Human-Computer Interaction (IJHCI)*, 29(6):379–390, 2013.
- [21] W. Javed, B. McDonnell, and N. Elmqvist. Graphical perception of multiple time series. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis)*, 16(6):927–934, 2010.
- [22] H. Lam, T. Munzner, and R. Kincaid. Overview use in multiple visual information resolution interfaces. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis)*, 13(6):1278–1285, 2007.
- [23] T. Lammarsch, W. Aigner, A. Bertone, J. Gärtner, E. Mayr, S. Miksch, and M. Smuc. Hierarchical temporal patterns and interactive aggregated views for pixel-based visualizations. In *Proc. IEEE Intl. Conf. Information Visualization*, pages 44–50, 2009.
- [24] D. Lloyd and J. Dykes. Human-centered approaches in geovisualization design: investigating multiple methods through a long-term case study. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis)*, 17(12):2498–2507, 2011.
- [25] McGill Energy Project (MEP) Energy Map. <http://mcgillenergyproject.com/map.php>.
- [26] S. McKenna, D. Mazur, J. Agutter, and M. Meyer. Design activity framework for visualization design. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis)*, 20(12):2191–2200, 2014.
- [27] P. McLachlan, T. Munzner, E. Koutsofios, and S. North. LiveRAC: Interactive visual exploration of system management time-series data. In *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI)*, pages 1483–1492, 2008.
- [28] L. McNamara, K. Cole, and S. Stevens-Adams. “Where Do I Start?” Practical Methods for Design Studies in InfoVis and Visual Analysis, 2014. IEEE VIS Tutorial.
- [29] R. Metoyer, B. Lee, N. H. Riche, and M. Czerwinski. Understanding the verbal language and structure of end-user descriptions of data visualizations. In *Proc. ACM SIGCHI Conf. Human Factors in Computing Systems (CHI)*, pages 1659–1662, 2012.
- [30] M. Meyer, M. Sedlmair, P. S. Quinan, and T. Munzner. The nested blocks and guidelines model. *Information Visualization*, 2013 (in press). <http://doi.org/10.1177/1473871613510429>.
- [31] T. Mühlbacher, H. Piringer, S. Gratzl, M. Sedlmair, and M. Streit. Opening the black box: Strategies for increased user involvement in existing algorithm implementations. *IEEE Trans. Visualization and Computer Graphics (Proc. VAST)*, 20(12):1643–1652, 2014.
- [32] T. Munzner. A nested model for visualization design and validation. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis)*, 15(6):921–928, 2009.
- [33] T. Munzner. *Visualization Analysis and Design*. A K Peters Visualization Series, CRC Press, 2014.
- [34] Northwrite Energy Expert. <http://northwrite.com/energyexpert.asp>.
- [35] J. Rodgers and L. Bartram. Exploring ambient and artistic visualization for residential energy use feedback. *IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis)*, 17(12):2489–2497, 2011.
- [36] RStudio. Shiny: A web application framework for R, 2014. <http://shiny.rstudio.com/>.
- [37] Schneider Electric PowerLogic Ion EEM. <http://goo.gl/Nh7mJv>.
- [38] M. Sedlmair, P. Isenberg, D. Baur, and A. Butz. Information visualization evaluation in large companies: Challenges, experiences and recommendations. *Information Visualization*, 10(3):248–266, 2011.
- [39] M. Sedlmair, M. Meyer, and T. Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis)*, 18(12):2431–2440, 2012.
- [40] M. H. Shimabukuro, E. F. Flores, M. C. F. de Oliveira, and H. Levkowitz. Coordinated views to assist exploration of spatio-temporal data: A case study. In *Proc. IEEE Intl. Conf. Coordinated and Multiple Views in Exploratory Visualization (CMV)*, pages 107–117, 2004.
- [41] E. Tufte. *Envisioning Information*. Graphics Press, 1990.
- [42] J. J. van Wijk and E. R. van Selow. Cluster and calendar based visualization of time series data. In *Proc. IEEE Symp. Information Visualization (InfoVis)*, pages 4–9, 1999.
- [43] K. J. Vicente. *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work*. CRC Press, 1999.
- [44] M. Wattenberg. Baby names, visualization, and social data analysis. In *Proc. IEEE Symp. Information Visualization (InfoVis)*, 2005.
- [45] H. Wickham and L. Stryjewski. 40 years of boxplots, 2011. Technical report. <http://vita.had.co.nz/papers/boxplots.pdf>.
- [46] L. Wilkinson and M. Friendly. The history of the cluster heat map. *The American Statistician*, 63:179–184, 2009.