

We thank the reviewers for their helpful feedback which has led to many improvements in the paper. The point by point description of changes are below: the reviewers' comments are in red and our response is in black.

In addition, we have expanded the paper by including a subsection titled '2.4 Reasons to use calendar-based graphics' to explain why to use calendar plots, and a short section titled '3 Making calendar facets' which describes a new function `facet_calendar()`, that extends **ggplot2** intervals and complements the existing function `frame_calendar()`. It allows this algorithm to integrate more deeply with the **ggplot2** architecture.

## Reviewer 1 comments

Figure 2 and Figure 10 show much the same information. You can see the Australian Open period in both. There is more detail in Figure 10 for the State Library. The gaps in the Birrarung Marr data are easier to see in Figure 2.

We have explained the purpose of showing both displays in the revised paper, by including a new subsection '2.4 Reasons to use calendar-based graphics' to explain why to use calendar plots, and what information is typically revealed by calendar plots but not by other plots. The calendar display doesn't replace traditional displays, but serves to complement them to further tease out interesting structures in the data.

Figure 11 is not described well. What does the V-pattern mean? How can you see the strong autocorrelation at weekends? There are only dots.

We have changed the explanation of Fig 11 to make it clearer. In the revised paper this is now Fig 10. The V pattern is due to morning and evening commuter pattern on weekdays. We realised that the daily lag relationships would be read more easily if local scale was used, that is, each day is scaled relative to itself. This has been done, and the different autocorrelation pattern for week days and weekends is now clearer.

The colours are not well chosen in Figure 15. It is hard to see the averages for households 1 and 3.

Colours have been changed to make it clearer.

There are too many large plots in the paper. Figure 6 is not needed and Figure 12 is not needed.

We are not sure if you mean Fig 6 here. Fig 6 is the diagram illustrating the yearly layout, and we think it is important to keep this illustration in the paper. On Fig 12, we think you mean Fig 13 is not necessary along with Fig 12. We have kept Fig 13 with the Chinese labels, to illustrate variety in language labels as well as the examination of variability using boxplots.

How do you know that the New Year's Eve outliers in Figure 12 are from the same sensor? That is not in the plot.

Yes, you are correct. You can't see that it is one sensor from this plot. The explanation of Fig 12 (now 11) focuses on the variability of counts across days, instead of outliers.

Do Figures 16-19 show much? Drawing all four series as rows in facets one above the other would show the holidays better. Comparisons of the households would be easier. The six-weeks months confuse looking at holidays. What are the one-day peaks in the middle of holidays?

Yes, we agree, they take up too much space. But the message of comparing multiple households is good to illustrate the benefits of the calendar plots, so we have reduced the plots to focus only on Jan-June 2018 and show all four households in one plot.

There is a shiny app that presents some of the data. It is not well described. The plot at lower left is cumulative precipitation for each month, but it does not say so. Why is cumulative precipitation by month interesting? Months are of different lengths. You can select single days in the time series. Can you select more than one day?

We have added a text box underneath the plots in the shiny app. To describe the app in this paper would be too distracting for the reader. Cumulating monthly precipitation provides a more interesting look at rainfall, and it approximates how the NY Times displays weather. Only one day can be selected.

## Reviewer 2 comments

p.10, l.1-15 It wasn't clear to me why the subtraction was performed in the equation (2) because I couldn't find any description for the global coordinate system of the canvas. The value of  $y$  could be negative. Please add the description how to deal with the coordinate system, especially the  $y$ -axis.

This is already explained directly after the equation. It is because plots are measured from top left of the display area, which is the common computer graphics standard.

## Reviewer 3 comments

Thank you for carefully reading of our paper.

p1, l32: 'information, and events.' - remove the comma (and all of the other similar small details)

Fixed in multiple places. The language in many places has been modified according to these comments.

p2, l9: 'grammar of graphics' - provide reference immediately here and not only later in the article and explain what the grammar of graphics is; also 'piped into' may not be understood by outsiders

Done.

Fig.1 (& related figs): I suppose you use a '3-class Dark2' color scheme from Rcolorbrewer? This does not work well when printed in gray scale: The colors become almost indistinguishable. ...

We have changed the colour scheme from ColorBrewer palette 'Dark2' to '4-class PuOr' in Figure 1 and related figures. This works better on the grey background, but still can be read on a white background in the other plots.

Fig.4: Colors orange & purple are hard to distinguish in grayscale. Moreover, you should use different colors here as these 2 colors are already related to the 3 stations.

In Fig 4, we have changed the colour to red and blue picked from a ColorBrewer '4-class RdYlBu' colour scheme. Fig 1-3 colours used for the 3 stations use a different palette.

Fig.2: Instead of 'Jan 2016' ... 'Jan 2017', be specific and list 'Jan 1, 2016' ... 'Jan 1, 2017'

Done.

Fig.2: Can you match the minor grid lines with the start of a month, rather than the middle of a month? This will make it much easier to identify the approximate dates for some of the other spikes. No need to label these minor grid lines.

Done.

Fig.2: You mention 'small multiples' only in a figure caption. This concept is a central part of your calendar graphics and should be summarized in more details (including references) in the main text. Also mention Unwin & Valero-Mora's 'Ensemble Graphics', JCGS, 27(1), as a major concept that applies to both sets of related figures in your article.

We have explained the concept of faceting and also referred to small multiples, trellis charts, and ensemble graphics in a new paragraph in the Introduction section.

Fig.3: Add tick marks at 0 and 24. See whether tick mark labels fit. If not, OK to omit those for 0 and 24.

Done.

p6, l8: You mention ggTimeSeries & ggcal. For completeness, also cite Jones (2016) Calendar Heatmaps, <https://rpubs.com/haj3/calheatmap>, and possibly Wong's TimeProjection R package, <https://cran.r-project.org/web/packages/TimeProjection/index.html>

Done.

p6, l49-...: Starting a section called 'Data transformation' with an example/figure seems to be strange. Can you first start with the formal steps and then place this example/figure after the formal part, i.e., around p.8, l38.

Done.

Fig.4: Colors orange & purple are hard to distinguish in grayscale. Moreover, you should use different colors here as these 2 colors are already related to the 3 stations.

Done.

Fig.4: Prior to reading the text on p8, I was really confused and even assumed there was a major bug in your R code. Having 2 days in early May and 1 day in early Oct, then 3 days missing, and then the remainder of the month is really confusing for someone who only looks at the figure without reading the text in detail. At least mention this layout anomaly in the figure caption.

Thanks. We have pointed out the layout anomalies for May and October in the Figure 4 (now 6) caption.

p8, l19: 'wrap the last few days up to the top row of the block': This answers my comment for Fig.4 now, but this layout still remains misleading. We expect to see some similar temporal pattern in nearby graphs. But there could be considerable differences over a 30-day period, e.g., in your example from Section 3: What if these 1 or 2 days are after some summer vacation with lots of air conditioning use, but the vacation already started in the previous month and continued through the middle of the current month. Suddenly, there will be a few huge spikes that interrupt the overall low-energy pattern. I could think of 2 possible solutions: (i) Add a 6th week for each month; or (ii) Add the extra days to the start of the next month. This is sometimes called a 'Calendar Heatmap Tetris Chart', see for example <https://stackoverflow.com/questions/27000131/calendar-heat-map-tetris-chart> [both of these features could become additional user options for your R function]

It is intentional to conform a calendar layout of 5 rows and 7 columns for the purpose of compactness. Only one or two months in a year will span over 6 weeks, which would result in more unused space for a year-long calendar plot. We don't shift the extra days to the next month because it is difficult for readers to distinguish which month contains these additional days without proper labelling. We agree that it is disruptive since time is usually perceived as linear, but it is also a common calendar convention.

That said, the algorithm that lays out the calendar is flexible to incorporate these two options. We have included them as the options for future work. Thanks for the suggestions.

p8, l44: 'Between each month requires some small amount of white space, denoted by b.' - strange sentence; rephrase

Done.

p11, l44: 'star plots': briefly explain what these are and cite a basic reference for these.

Done.

Fig.7: Where is 'noon' and 'midnight' here - on top/bottom or on the 0/180 degree position on the right/left side? And what is the direction? I suppose clockwise, but this needs to be mentioned somewhere.

The star plots have been removed from the paper, since it doesn't add more information to other existing plots.

p13, l39: 'Figure 13 shows the same plot as Figure 12 labelled': You can't jump forward to a figure before all intermediate figures have been introduced. The previous fig was Fig.8. Figs 9-11 have not been mentioned yet. Either rearrange your figures so that 12-13 become 9-10, or use Chinese characters for one of the previously introduced figs (1-8).

Fixed. 'Figure 13 shows the same plot as Figure 12 labelled' is deleted.

p13, possible subsection 2.2.6: You seem to have a 'sunday = FALSE' default. If TRUE, does this make Sunday the start of the week (as used in the US & Canada)? If so, mention this as an option. No need for a figure.

Done. We deprecated argument `sunday` in favor of `week_start` that is more general. We added a new subsection on 'Start of the week' to briefly explain the option. We have reorganized the order of the subsections to be consistent with the function argument's ordering.

p13, possible subsection 2.2.7: How can you enter specific holidays? These are different from country to country, e.g., based on national holidays such as Thanksgiving, Labor Day, independence days, religious holidays, etc.

We used `tsibble::holiday_aus()` for generating Australian public holidays, prior to creating calendar displays. Whether to use holidays or not is a data-specific problem, and would depend on local analysts incorporating a holiday calculator into their country specific analyses. This is beyond the scope of this paper.

Fig.8: 'impossible to compare the size of peaks between days.' - Correct; since this is based on ggplot, could you use colors from a sequential color scheme that maps the counts, e.g., from faint yellow (0) to dark red (for the overall max)? This might be another useful argument for a future extension of your function.

We recommend against this idea. The better approach is to re-make the plot with global scaling to focus on counts. No need to make one plot show everything possible, but better to make multiple plots.

p15, l12: 'idea of faceting': explain what 'faceting' means in general (and cite 1 or 2 main references)

We added the explanation of 'faceting' in the Introduction section.

p15, l16: 'In particular, it can be immediately learned that when Birrarung Marr was busy and packed, for example Australian Open in the last two weeks of January.' -> 'In particular, it can be immediately learned that Birrarung Marr was busy and packed, for example during the Australian Open, a major international tennis tournament, in the last two weeks of January.'

Done.

p15, l45: 'the day before Christmas, go shopping on the Boxing day, and stay out for the fireworks on New Year's Eve.': Some of these names may not be known to non-Christian / non-British readers. List the date in parentheses, e.g., December 24, December 26, etc.

Done.

Fig.12: 'loess smooth': cite main reference

Done.

p21, l36: 'these four households are the data of colleagues of the authors.' - strange sentence; rephrase

Done.

p22, l17-35: inconsistent use of times, e.g., 'before 6', 'around 18', '3pm', etc. - adjust to times used in the figures and be consistent. Also adjust caption of Fig.15.

Done.

p22, l36: 'Figure 16, 17, 18 and 19' -> 'Figures 16-19'

Done.

'individually for each household': This seems to suggest that a scale from 0 to max(household\_i) [but not to max(all households)] has been used. But what is this max in each of the 4 figures? The reader has no way to easily infer this. Also mention that these 4 figs are not true small multiples as the vertical scales for the 4 households differ from figure to figure.

We merge these four figures into one plot (Figure 15) using global scales.

Fig.15: Use labels 6, 12 & 18 and extra tick marks at 0 & 24, matching Fig 3.

Done.

p25, l3: What about April? If added here, also add to caption of Fig.19.

Done.

p25, l14: 'Anzac Day in Australia, or Thanksgiving Day in the USA,': Never mentioned before in this article. Refer back to the special events/holidays from earlier in the article.

This is removed.

p25, after l35: List my suggestions for software additions from above that you do not want to implement now as options for future work.

Done.

Fig.16 caption: 'Calendar display for household 1, indicates higher weekend usage, and in the summer months, November-February. It seems that they took a vacation in June.' -> 'The calendar display for household 1 indicates higher weekend usage and higher usage in the summer months (November-February). It seems that this household took a vacation in June.'

Done.

Fig.17 caption: 'Calendar display for household 2, reveals their tendency to use air conditioning and heating continuously. Not many vacation were taken.' -> 'The calendar display for household 2 reveals their tendency to use air conditioning and heating continuously. Some vacation days were taken in late December and in the second week of June.'

Done.

Fig.17: I am concerned: Is this a 'fair' comparison when you drop the months August-October (& part of November) 2017 for household 2, while these 3+ months are part of the graphics for households 1, 3 & 4? Wouldn't it be better to restrict all households to Dec 2017 - July 2018? This would drop the winter/spring months from the aggregated displays in Figs.14-15. But household 2 still would have most variation, but now, this is based on matching time periods.

Thanks. We have restricted the data to Jan-June 2018 and combined four separate plots to a single plot using faceting. They all share common scales so that comparison can be made.

Fig.18 caption: 'Calendar display for household 3. Their energy use reveals higher energy use in the winter months, with multiple peaks daily on both week days and weekends. There are some high peaks in summer, perhaps indicating occasional air conditioner use. There have been several long vacations in the past year.' -> 'The calendar display for household 3 reveals higher energy use in the winter months (May-August), with multiple peaks daily on both week days and weekends. There are some high peaks in summer (November-February), perhaps indicating occasional air conditioner use. There have been several long vacations ... [match the main text].'

Done.

Fig.19 caption: 'Calendar display for household 4, shows energy use mostly in the evenings and on weekends. Three short trips were taken in October, December, and June.' -> 'The calendar display for household 4 shows energy use mostly in the evenings and on weekends. Three short trips were taken in October, December, and June.' [also April? - match with p25, l13]

Done.

p.30: City of Melbourne (2017) [& other URLs]: Does JCGS require the 'last accessed' information for web-based references? If so, add where needed.

Done.