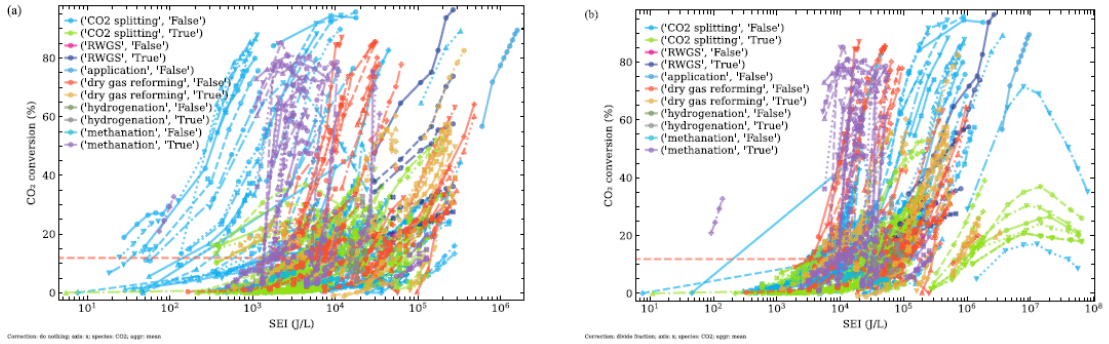

Fig. 4. Comparison of example graphs as they can be directly exported from the CO2-PDB app graphical user interface (left as-is, save for moving the legend) showcasing the CO_2 conversion (in %) versus the SEI (in J L^{-1}) with data grouped by the "main goal" and use of catalyst (either "True" or "False"). Both figures in principle show the same data, but in (b) the SEI has been normalised by dividing by the CO_2 fraction in the gas flow.

important metadata and minimal annotation on applied normalisation, which is crucial for comparison. Alongside these buttons, there are two tabs for further control over the plot and data. The "Plot setup" tab contains the widgets that control the appearance of the database plot. Using these widgets, it is possible to control the scale of the axes, the colour map of the plot and toggling automatically redrawing the plot after each filter change. Special attention should be given to the "Group data by" widget, which allows the selection of one or more categorical columns by which the data are grouped and colour mapped in the plot. By allowing the user to create custom groupings, insights into how certain groups and categories of experiments compare against each other can be made, see Fig. 4(a) as an example.

The "Normalisation" tab provides control over several different normalisation functions that can be applied to either the *X* or *Y* axis of the graph. Further input to some normalisation functions is controlled by the "Normalise species" and "Normalise aggregate func." widgets, which respectively control the species to normalise (e.g., normalising for CO_2 or H_2O partial pressure) and how to aggregate array-like metadata where applicable (e.g., taking the minimum, maximum or mean of the temperature if a range is defined). The same aggregation will be applied to all numeric columns that are needed for a calculation. The effect of applying a normalisation to the data can be seen in Fig. 4(b), which showcases the same filters and grouping as Fig. 4(a), but now with the data normalised for the CO_2 fraction in the initial gas mixture. In Section 4.2, a discussion of the benefits of this normalisation is presented.

It is important to note that normalisations only return data for which the necessary metadata used in calculation is present. In other instances, NaN values are returned. Consequently, applying a normalisation can result in a decrease in the amount of data shown.

3.3.2. Advanced data inspection

The main interface of the database tool provides convenient and powerful methods to filter, group, and visualise the data within the database, making it well suited to data exploration and comparison between publications. It however cannot replace all possible filtering and visualisation for a user's needs, and neither does it aim to. Furthermore, understanding nuances and outliers in the data requires a detailed comparison of metadata such as pressure, temperature and volume. For this reason, there is a secondary tab called "advanced database inspection" in the main viewport, which offers a powerful pivot table interface (referenced as "Data Table" going forth) to the user for data inspection. It gives full access to the data in the database but is limited to some extent in that it cannot represent array-like metadata, such as temperature ranges, as numeric array objects, but rather as string-like objects. Applied categorical filters in the main view are synchronised to the Data Table to support an explore-then-inspect workflow. However, please note that these filters do not work in the reverse direction. Furthermore, fine-grained control over data selection is enabled by extensive customisation of selected columns, filters, groupings, and aggregations across the database using drag-and-drop of column names. Using the toolbar on the bottom of the Data Table widget, it is possible to export the selected data or all data to several common formats such as, `*.csv` and `*.json`, thus enabling the user to utilise the data in their own workflow after initial inspection, and control what data they need. As a final feature, there is the option on the top left of the widget to change the view from a tabular "Datagrid" to other visualisations, such as, bar charts and scatter plots, which can be easily adapted using the column list and aggregations. This complements the feature set of the main interface, allowing further exploration, and represents the main way in which a copy of a subset of the database can be obtained. Examples of data extracted with the aid of the CO2-PDB app are shown and discussed in Section 4.1.

3.3.3. A note on calculations, assumptions and features

It is important to note that the database is not just a catalogue of data, but also a tool that allows quick comparisons of (normalised) data. However, there are some core assumptions and approximations that are made to compare as much data as possible across a wide range of conditions.

For example, SEI is seen as a macroscopic process parameter that inputs power into a given amount of flow rate regardless of the plasma conditions. Within the database, SEI values directly extracted from publications appear alongside values calculated from reported powers and flow rates, and are labelled accordingly. Contrary to the direct calculation of SEI, the adjustment with respect to standard conditions is barely mentioned in the publications used to build the database. Therefore, we can only assume that temperature and pressure inside the plasma or reactor are not considered when SEI or energy efficiencies are calculated in the majority of publications. Here, in the calculation of τ_{res}, the standard conditions are taken into account when adjusting for the discharge conditions. It is crucial to know how much time a molecule spends under plasma conditions. Similarly for the energy efficiency calculations, the correction of the SEI used here had to ponder temperature and pressure (values extracted from the database). Ultimately, the number of particles into which energy has been deposited is considered in the back end CO2-PDB app.