CS-732 Data Visualisation Report for A2

Arjun Subhedar (IMT2021069)
IIIT Bangalore
Arjun.Subhedar@iiitb.ac.in

***Abstract*—This report presents an analysis of three SciVis tasks and three InfoVis tasks. The three Scivis tasks- color-mapping, contour-mapping, and generating quiver plots, focus on understanding the ocean datasets. The first two of the aforementioned tasks are performed using the parameter, columnar water vapour [?], from the ocean dataset, for the time period of three months from August 2015 to October 2015. Whereas, the quiver plots are generated for wind-speeds over the sea surface [?], from August 2013 to October 2013.**

I. SciVis

A. Dataset Description

The dataset for the first two scientific visualization tasks, namely color mapping and contour mapping, focuses on columnar water vapor over the sea surface. This measure indicates the total gaseous water within a vertical atmospheric column, expressed in units of kg/m^2. Spanning globally, the dataset comprises scalar values across 720 latitudes (from 89.875N to 89.875S) and 1440 longitudes (from 0.125E to 0.125W). Land points are identified by a bad flag value of -999, resulting in 720 lines for each latitude.

For the third task, involving vector mapping through quiver plots and streamplots, distinct meridional and zonal wind speed datasets are utilized. Meridional winds, representing the x-component of the total wind speed, flow along latitudes, while zonal winds, the y-component, flow along longitudes. The speed unit in both datasets is m/s. Each dataset contains scalar values at 360 latitudes (from 89.75N to 89.75S) and 721 longitudes (starting from 0E, completing a full circle, and returning to 0E). Ignoring header information, this results in 360 lines for each latitude. Both datasets for these tasks cover nine dates at ten-day intervals in the year 2015: August 5, August 15, August 25, September 5, September 15, September 25, October 5, October 15, and October 25.

B. Data Pre-Processing

The datasets, initially in ASCII format, were text files with removed header information containing longitude positions. Some datasets for columnar water vapor had periodic missing values for longitudes, which were reconstructed. Another challenge was incomplete visibility of data in certain lines due to length, resulting in a variable number of data values instead of the expected 1440. To address this, the minimum of the maximum length for each line was considered, determining the number of workable longitudes. Consequently, the dataset was truncated to $720 \times M$ (where M is the minimum of maximum lengths for each line) data values, rather than the original 720×1440. There were no such problems with the zonal and meridional wind speed datasets.

C. Methods

We have used **matplotlib** for plotting the graphs. Specifically the functions we have used are - **pcolor** for the color maps, **contourf** or the contour plots, **quiver** for the arrow plots, and **streamplot** for the streamline plots. **numpy** is used to convert the data into a format as required by the functions to do the visualisations. **Basemap** has been used to plot the visualisations over the world map.

D. Color Map

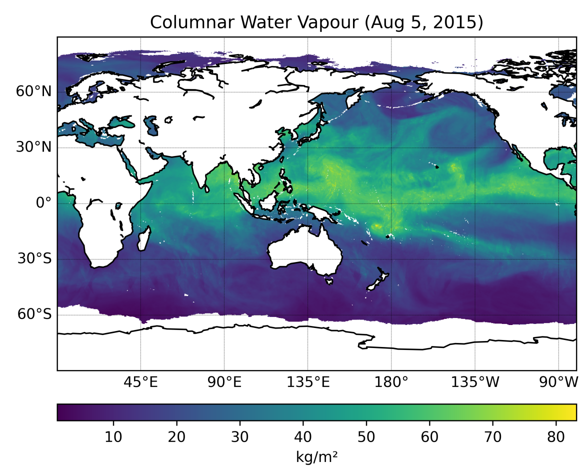
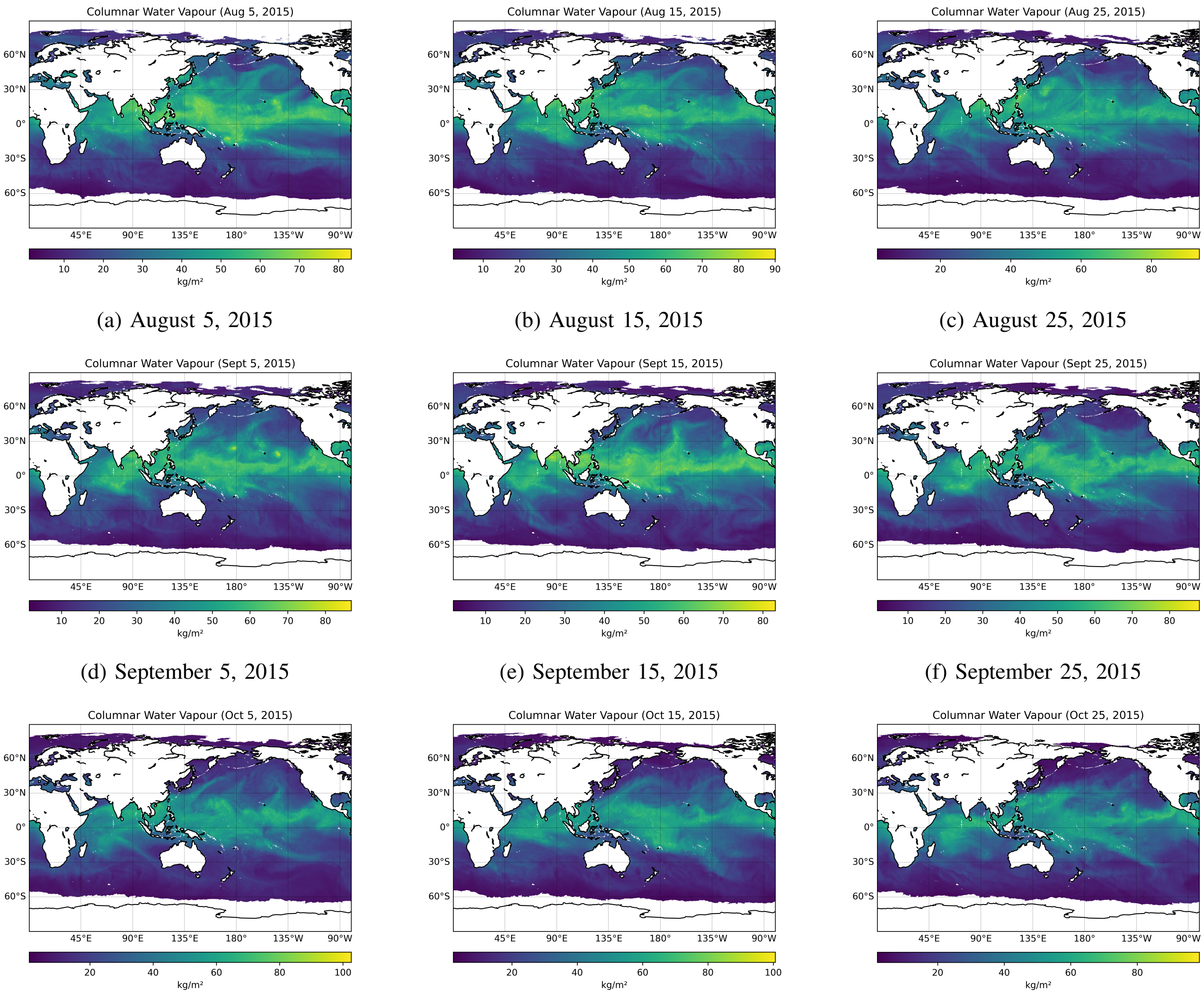


Fig. 1: Columnar Water Vapour over the Sea Surface

In Figure 1, we can see the colormap generated for the columnar water vapour data for August 5, 2015. The colormap used over here is a perceptually sequential colormap, *viridis*.

The colormaps for each of the nine dates that we have chosen, has been shown in Figure 2. All of these color maps are *viridis*. The maximum values of the columnar water vapour on each of these dates is summarised in table I.



(a) August 5, 2015
(b) August 15, 2015
(c) August 25, 2015
(d) September 5, 2015
(e) September 15, 2015
(f) September 25, 2015
(g) October 5, 2015
(h) October 15, 2015
(i) October 25, 2015

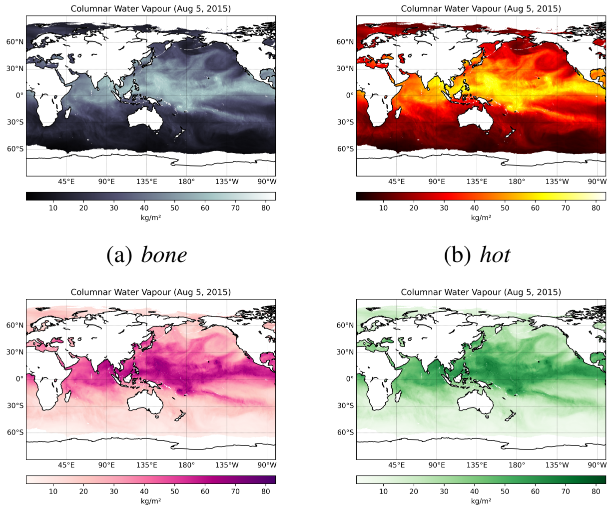
Fig. 2: Color maps of Columnar Water Vapour data across the nine sampled dates.

Date	Value
Aug 5, 2015	83.24195
Aug 15, 2015	90.05486
Aug 25, 2015	93.628
Sept 5, 2015	86.85767
Sept 15, 2015	83.13477
Sept 25, 2015	87.97
Oct 5, 2015	102.636
Oct 15, 2015	100.612
Oct 25, 2015	98.6653

TABLE I: Maximum values of the Columnar Water Vapour

We can notice a change in the scale under each of this colormaps indicating the change in the atmospheric water vapour conditions throughout the three months. The analysis reveals a crucial discovery: on October 5, 2015, the global maximum of columnar water vapor reached 102.636, spanning nine dates. Notably, all three October dates present notably higher values, signaling increased humidity and precipitation. This correlation is clear—elevated water vapor in the air corresponds to heightened humidity and a larger volume of precipitable water in the atmosphere. The findings emphasize October's distinct atmospheric conditions, shedding light on the interconnected dynamics of water vapor, humidity, and precipitation during this period.

To enhance the visual representation of these variations, a series of experiments were conducted with alternative perceptually sequential colormaps, including *bone*, *hot*, *RdPu*, and *Greens*, showcased in Figure 3. Notably, the *viridis* colormap emerges as the most fitting choice, aligning its color scheme with the characteristics of columnar water vapor associated with water content, humidity and precipitation patterns.



(a) *bone* (b) *hot*

(c) *RdPu* (d) *Greens*

Fig. 3: Different colormaps for the Columnar Water Vapour over the Sea surface

We have also visualised the same data with a logarithmic scale of the color map, which can be seen in Figure 4.

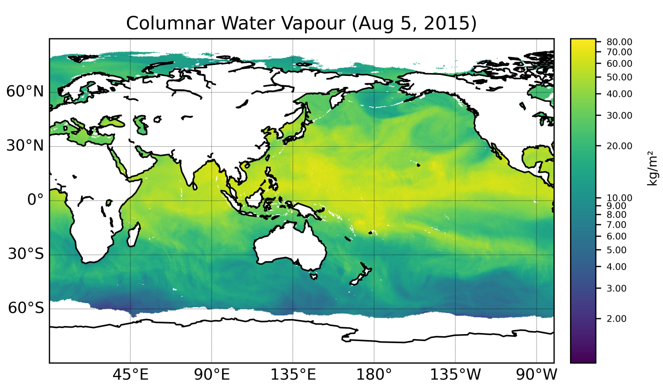
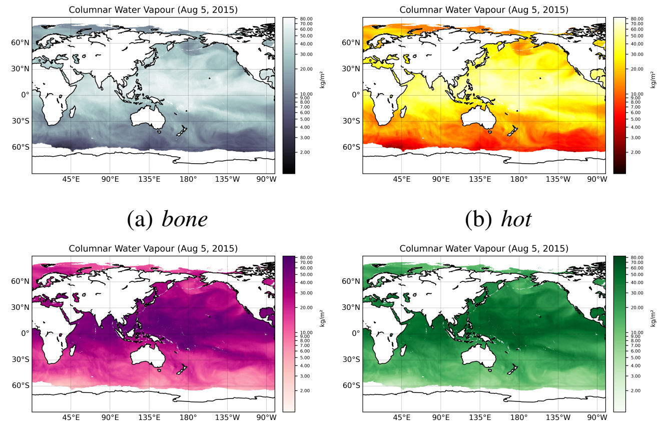


Fig. 4: Logarithmic Scale *viridis* colormap for the Columnar Water Vapour over the Sea Surface

Logarithmic colormaps are instrumental in representing data with a wide dynamic range, particularly when certain values span several orders of magnitude. They are designed to map data values on a logarithmic scale rather than a linear one. In this logarithmic scale exploration, various colormaps were experimented with (refer to Figure 5). Notably, the *Greens* and *RdPu* colormaps exhibit enhanced visual perceptibility in comparison to their linear scale counterparts in Figure 2, with darker colors becoming more distinguishable. Conversely, the effectiveness of the overall colormapping for the other two colormaps is reduced due to the introduction of more number of lighter shades.

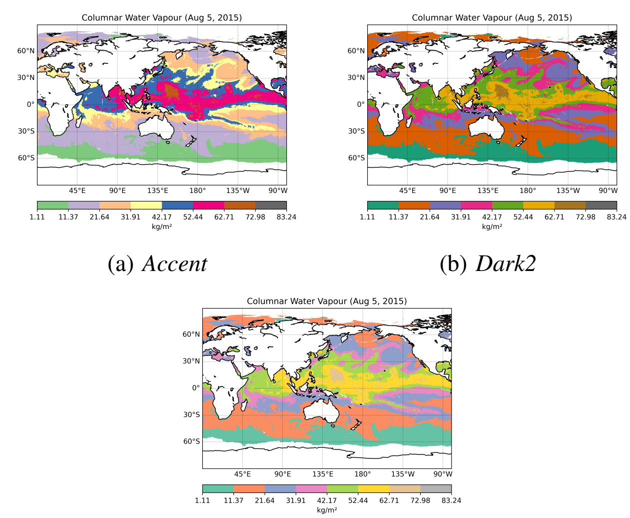


(a) *bone* (b) *hot*

(c) RdPu (d) *Greens*

Fig. 5: Different logarithmic scale colormaps for the the Columnar Water Vapour over the Sea Surface

As the data values are continuous rather than categorical, opting for continuous maps is more suitable for colormap visualizations. In Figure 6, discreet colormaps (Accent, Dark2, Set2) are generated for the August 5, 2015 dataset. However, discrete/qualitative colormaps lack perceptual order, making it challenging to discern which colors represent larger or smaller values without referencing the scale.

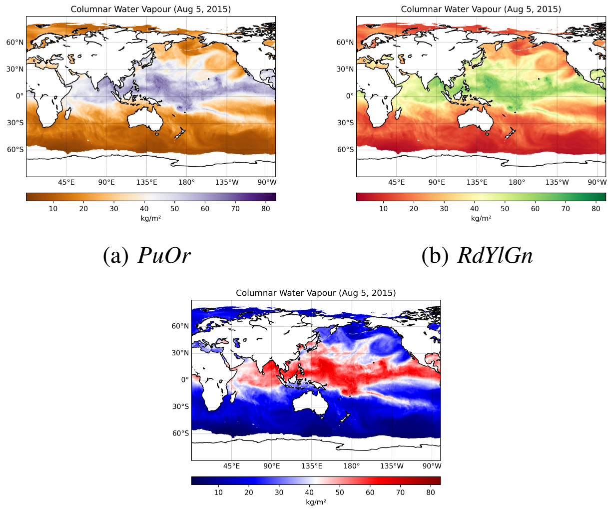


(a) *Accent* (b) *Dark2*

(c) *Set2*

Fig. 6: Different Discrete Colormaps for the the Columnar Water Vapour over the Sea Surface

During our investigation into effective visualization methods, we probed the application of diverging colormaps, as depicted in Figure 7. These colormaps serve as pivotal tools in data representation, especially when the goal is to underscore variations or deviations from a central reference point. In our context, the choice of a perceptually sequential colormap such as *viridis* was deliberate and well-suited to our dataset. The dataset in question showcases a discernible unidirectional trend in the columnar water vapor data, without a distinct threshold value around which the data values pivot.



(a) *PuOr* (b) *RdYlGn*

(c) *seismic*

Fig. 7: Different Diverging Colormaps for the the Columnar Water Vapour over the Sea Surface

E. Contour Maps

When creating contour plots, the selection between the Marching Squares algorithm and the Contour Fill algorithm becomes crucial. This choice is exemplified in the context of generating contour maps for columnar water vapor data, where the Contour Fill algorithm is preferred. The Marching Squares algorithm, while computationally efficient and straightforward, exhibits limitations when applied to large datasets, often producing jagged or pixelated outcomes that may not accurately convey the underlying smoothness of patterns. In contrast, the Contour Fill algorithm employs a more sophisticated approach, resulting in visually appealing and smoother contours. This superiority is particularly pronounced when dealing with extensive datasets or those characterized by intricate patterns, as the Contour Fill algorithm excels in portraying subtle changes in the data, such as variations in weather patterns. Despite potential increased computational demands, the Contour Fill algorithm's ability to enhance the clarity of contour representations makes it the method of choice for discerning and accurately visualizing nuanced features within the columnar water vapor data.

We have plotted the contour maps for different levels of contouring, as can be seen in Figure 8.

It can be observed that as the levels of contouring, i.e, the number of contour in a map increases, the contour mapping tends to become smoother and at very high levels of contouring, it might become sequentially continuous map too (see Figure 9)

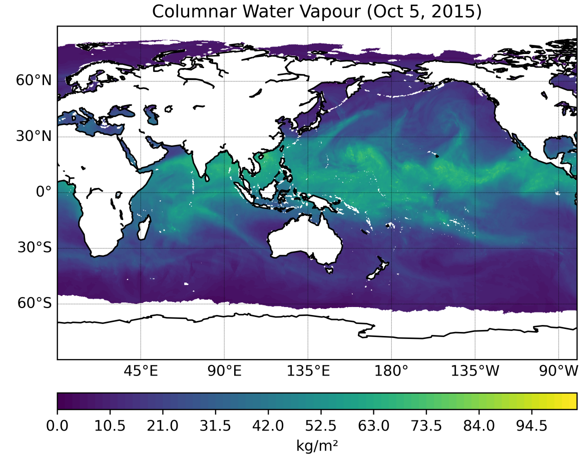
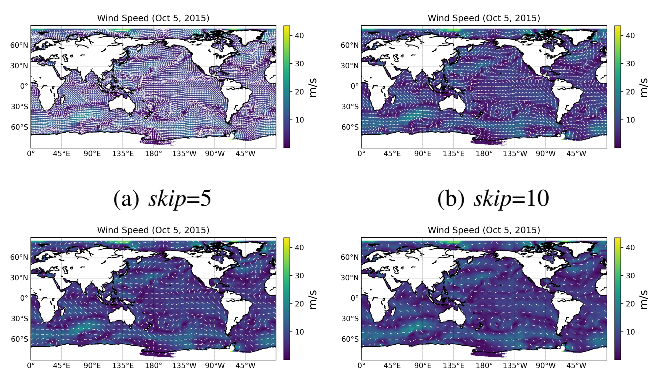


Fig. 9: Contour Map of Level 100 for Columnar Water Vapour Data. This nearly resembles the same as sequential continuous colormap (see Figure 2 (g))

F. Quiver and Streamline Plots

For the vector data plotting of the wind data, we have generated arrow plots and streamplots. Experimentation has been done with different grid sampling strategies as can be seen in Figure 10. In this figure the size of arrows is normalized (all arrows of the same size). Beneath the arrow plot, there is a colormap of the total wind speed. This has been done for all nine dates, an instance of which is given below.

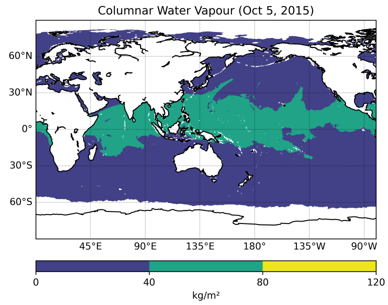


(a) *skip*=5 (b) *skip*=10

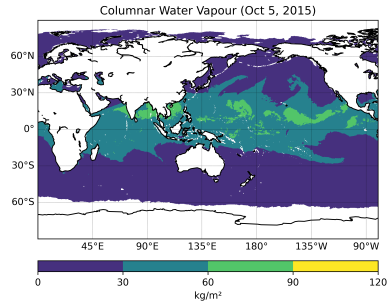
(c) *skip*=15 (d) *skip*=20

Fig. 10: Different grid sampling experiments. In this above figures, the data has been plotted by skipping some values, because considering all values causes visual clutter. Skip values are given below each figure. For a given skip value, it means that we skip the arrow data value by that value.

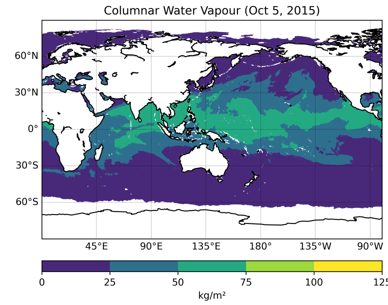
For the same grid sampling experiments, we have generated arrow plots where the size of arrows is directly proportional to the magnitude of the total wind speed at a point. This is shown in Figure 11. We can see that, this technique is not very useful, as at points, where the wind speed is relatively lower, the arrow size of the vector reduces to a point nearly a



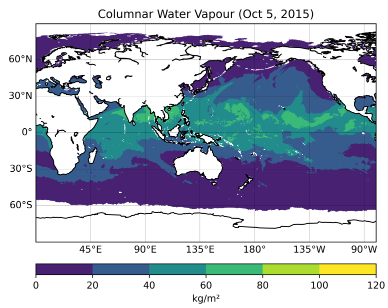
(a) Level 2



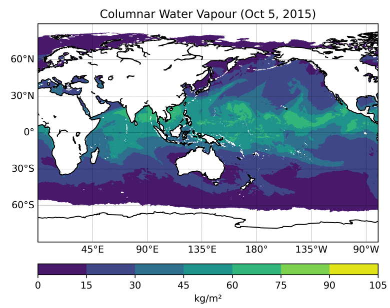
(b) Level 3



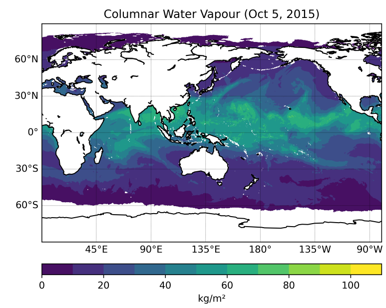
(c) Level 4



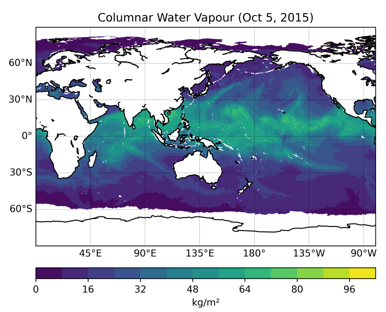
(d) Level 5



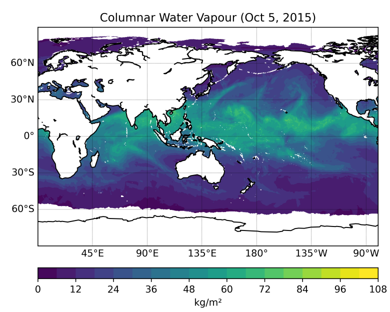
(e) Level 7



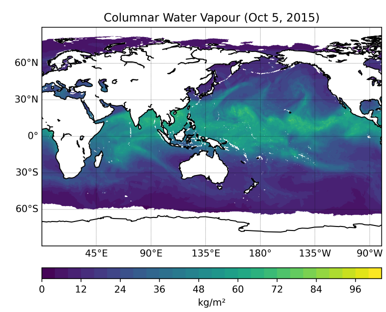
(f) Level 10



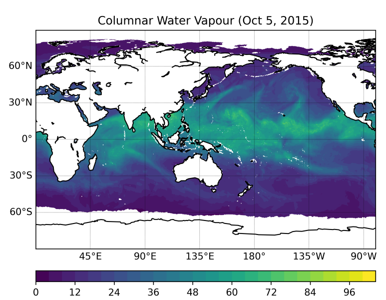
(g) Level 13



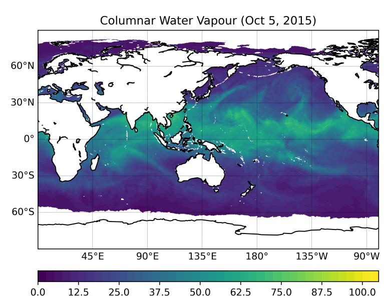
(h) Level 18



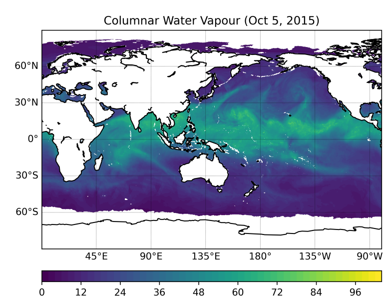
(i) Level 25



(j) Level 32



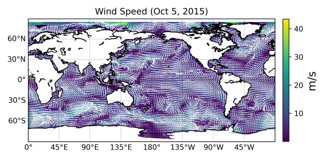
(k) Level 40



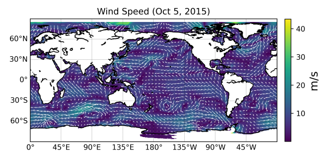
(l) Level 50

Fig. 8: Contour maps of columnar water vapour with different levels of contouring.

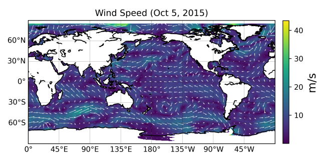
point, making it hard to discern which direction is the arrow pointing to.



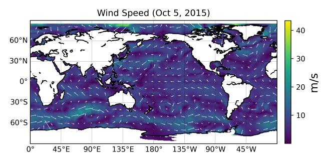
(a) *skip* = 5



(b) *skip* = 10



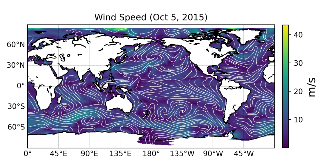
(c) *skip* = 15



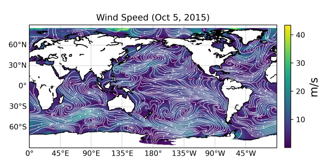
(d) *skip* = 20

Fig. 11: Different grid sampling experiments. Here, the lengths of the arrows are proportional to the magnitude of the total wind speed at their respective points.

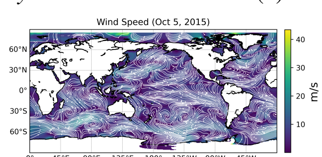
For the same wind speed data, we have plotted the streamline plot (see Figure 12). The density of lines in the maps is something we have explored. We can see as the density increases, the lines get closer, and beyond density 9, the lines get so close that it leads to visual clutter.



(a) *density* = 3



(b) *density* = 6



(c) *density* = 9

Fig. 12: Streamline plots for Wind Speed data with different line densities.

As we compare the quiver and streamline plot, we note down the following points.

- The streamline plot shows the flow patters of the winds, whereas the quiver plot shows the correct representation of wind speed vector (magnitude and direction) at a point.
- Quiver plots may become hard to interpret as clutter occurs when the density of arrows increases. On the other hand, streamline plots do not give any information about the magnitude of the vectors at a point.