McGill Univeristy

ATOC/PHYS 404

Analyzing Cloud Movement in Relation to Wind Velocity

Project: Final report

Guillem Alvarez Bernal/261184788

December 5, 2023

1 Introduction

The relationship between atmospheric phenomena is a key focus in meteorological studies, particularly in understanding cloud dynamics and wind behavior. Cloud movement in the sky is influenced by various factors, with wind velocity playing a crucial role in shaping their distribution and path. This project aims to clarify the correlation between cloud movement and wind velocity by implementing a Centroid-Based object tracking system.

While it might seem intuitive that cloud velocity directly mirrors wind velocity, experimental evidence has revealed potential differences of up to 5 m/s between the two [1]. This finding highlights the intricate nature of cloud dynamics, suggesting that factors beyond wind influence the movement and behavior of clouds in the atmosphere.

Establishing a correlation between rapidly changing and widely dispersed variables like cloud coverage and wind velocity presents a challenge. To simplify this issue, this study narrows its focus to explore the spatial distribution of low, mid, and high-level clouds. Distinguishing between these cloud layers facilitates a clearer connection to their respective wind patterns, allowing for the establishment of distinct correlations under three different conditions.

To streamline cloud detection and tracking, an implementation of a Centroid-Based object tracking system is employed. This system utilizes the *OpenCV* library to extract cloud centroids and a custom-built *CentroidTrackerWithVelocity* class, effectively monitoring cloud movements. Enhancements in cloud detection involve the utilization of *SciPy's* ndimage function to create a downsampled cloud coverage map. By employing logical conditions and using *numpy* functionalities, such as threshold values and minimum region size, isolated regions with high cloud coverage are delineated in a separate mask, simplifying the detection process.

The conclusive outcomes of this study emphasize the effectiveness of the Centroid-Based object tracking system. Importantly, the obtained results align closely with empirical observations documented in the research conducted by A. F. Hasler, W. Shenk, and W. Skillman in the Journal of Applied Meteorology and Climatology (1976). We think that these findings could underscore the potential geographical characteristics of the Earth's surface.

2 Methodology

2.1 Data Extraction

Our study utilizes data sourced from the Copernicus Climate Change Service repository, specifically extracted from ERA5, an atmospheric reanalysis of the global climate.

The dataset comprises hourly recordings of cloud coverage and wind velocity spanning August 13th, 14th, and 15th, 2023, covering the entirety of the globe. While these dates were selected arbitrarily for demonstration purposes, precautions were taken to ensure the absence of extreme climate events, such as hurricanes or tornadoes, which could introduce biases into our data analysis. This dataset is represented by a 4-dimensional array, encompassing 1440 longitude points, 721 latitude points, 72 time points spanning three days, with 24 hours each, and delineating 3 distinct cloud distributions.

The primary parameters under study are the fractions of high, mid, and low cloud coverage. These parameters, scaling from zero to one, represent the proportion of a grid box covered by clouds occurring in different layers of the troposphere: higher, medium, and lower levels. This classification corresponds to pressures greater than 0.8 times, between 0.45 and 0.8 times, and less than 0.45 times the surface pressure, respectively [6]. This classification is determined by referencing the vertical distribution of wind velocities at 900 hPa for low coverage, 600 hPa for medium coverage, and 250 hPa for high coverage.

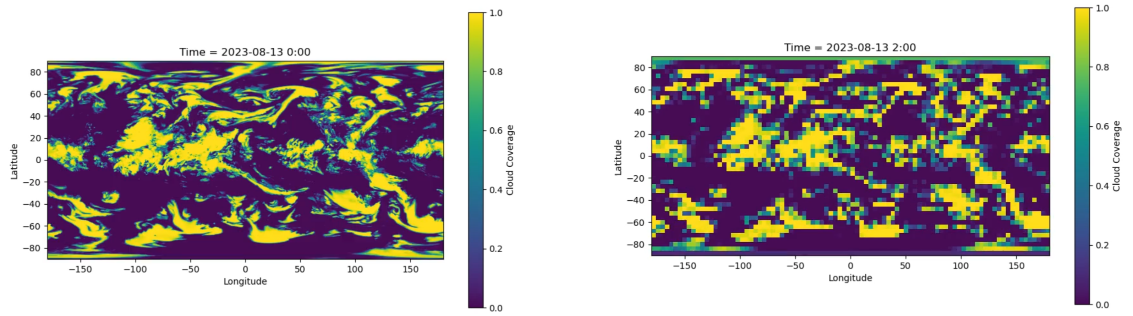
Furthermore, our project involves the extraction and analysis of the U-component and V-component of wind, sourced from the same repository and sharing the same temporal and spatial distribution as the cloud coverage. Both wind components are measured in meters per second m/s, signifying the eastward and northward movement of the wind, respectively. Due to the fact that the vertical wind component has limited influence on the horizontal movement of clouds, we excluded it to consolidate our focus on understanding the horizontal dynamics primarily involved in cloud migration.

2.2 Cloud detection

All the code referred is well presented and explained in the Jupyter Notebook attached to this report. First of all, the initial identification of clouds is a fundamental step in the analysis of cloud coverage data. Upon retrieving data from the NetCDF file in the form of heat maps, we employ the *downsample_heatmap* function. This function downscales the cloud coverage heatmaps using the *ndimage.zoom* method, increasing computational efficiency while preserving crucial cloud coverage details across specified spatial coordinates. The operational process of this function is visually depicted in Figure 1.

Simultaneously, the *highlight_isolated_regions* function operates on the downscaled heatmap, employing a multi-step process. Initially, it applies a threshold to discern areas with substantial cloud coverage. Next, it labels connected components within the thresholded heatmap, identifying small regions smaller than a specified minimum size that signify isolated high cloud coverage clusters. Finally, it constructs a mask to isolate and highlight these distinct regions, thereby facilitating further analysis into localized patterns of high cloud coverages. The resulting visualization is presented in Figure 2a.

Overall, this code allows us to do the downsampling for visualization and identification purpose. This is critical for managing the approximately 72 million initial data points without compromising the integrity of the primary cloud clusters.



(a) Original horizontal component of the wind for high range.

(b) Downsampled horizontal component of the wind for high range.

Figure 1: Applying downsampling dynamics to the initial frames of the high cloud coverage distribution.

2.3 Centroid Tracking System

2.3.1 Cloud velocity computation

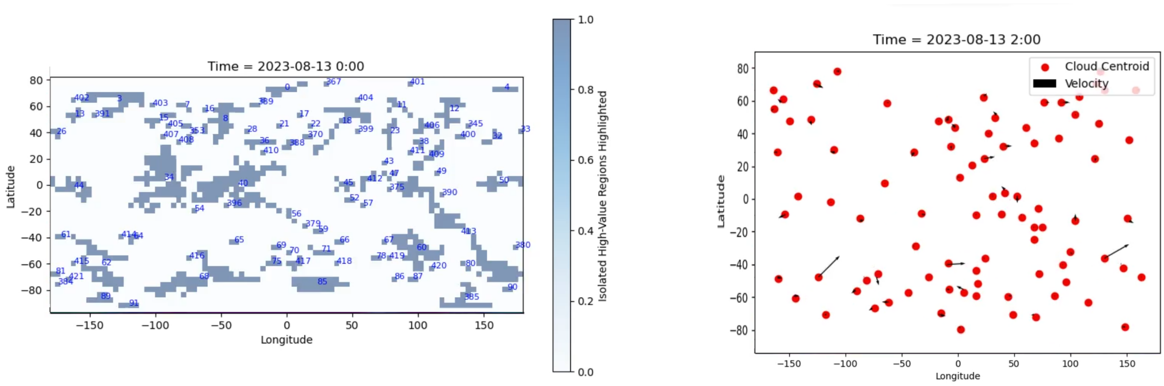
To compute cloud velocity, it is crucial to track the positional changes of clouds over time. Our approach involves labeling the same cloud consistently across different time intervals, enabling the calculation of positional differences, the centroids. A custom code has been developed to implement a Centroid-Based object tracking system. This system analyzes and tracks cloud movement across successive frames of downsampled heatmaps, utilizing OpenCV and classes designed for effective cloud centroid identification and velocity computation.

The *find_centroids* function employs OpenCV's *connectedComponentsWithStats* method to locate centroids of cloud clusters within the downsampled heatmaps. These centroids are visually depicted as red dots in Figure 2b. By identifying connected components and extracting centroids, this function establishes pivotal reference points crucial for following cloud tracking.

The CentroidTrackerWithVelocity class organize the object tracking process, managing cloud centroid positions and velocities across frames. It maintains a comprehensive record of cloud centroids using the *self.objects* dictionary and tracks their disappearance via the *self.disappeared* and *self.velocities* dictionaries. The *register* and *deregister* functions manage the addition and removal of cloud centroids, respectively.

The *update_with_velocity* method within this class dynamically updates existing centroids with their respective velocities, handles the appearance and disappearance of centroids between frames, and computes their velocities based on the elapsed time. The system executes centroid identification and tracking sequentially, providing the *tracked_objects* and velocity variables as outputs, enclosing updated centroid positions and their corresponding velocities.

Inside the *update_with_velocity* method, we employ a distance-based approach to determine correspondences between centroids in successive frames. It calculates the Euclidean distance between centroids from the current and previous frames, comparing new centroids with existing tracked centroids stored in the *self.objects* dictionary. Pairwise distances are computed, and the closest pairs of centroids are identified based on minimum distances. This distance calculation allows for the assignment of centroids in the current frame to the closest tracked centroids from the previous frame, ensuring continuity in the labels seen in Figure 2a across frames. The method updates existing centroids' positions and computes their velocities based on the elapsed time between frames, typically 3600 seconds in our case.



(a) Highlighted and labeled high cloud regions in a separate mask.

(b) Computed velocities for each centroid considering a time step of $dt = 3600$ s. We observed few higher unrealistic velocities.

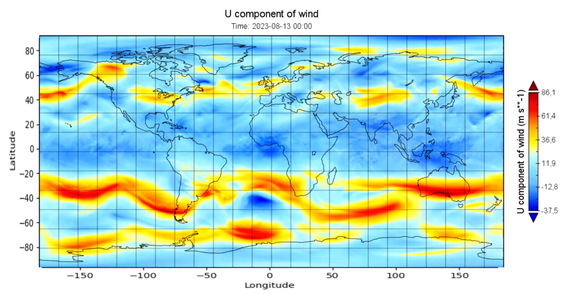
Figure 2: From labeling the clouds to generating the cloud velocity vector field map. We see in the first image the clouds correctly labeled and in the second image the centroids and their velocity.

2.3.2 Wind/cloud velocities comparison

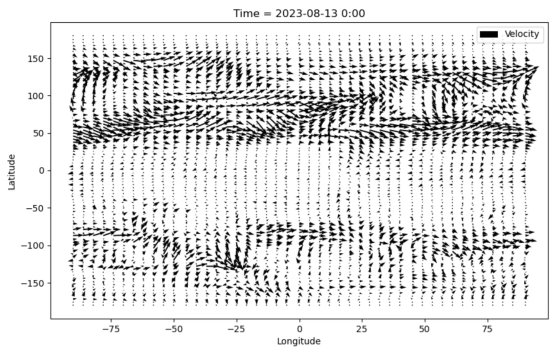
The primary objective is to evaluate the relationship between wind and cloud velocities using the Mean Squared Error (MSE), which is an statistical measure, represents the average of squared errors between the data and a function. A lower MSE the closer is computed to actual values, while a higher MSE signifies the opposite. This enables the assessment of higher or lower correlations with the aiming of identify potential correlations between cloud movement and wind patterns, specifically focusing on both land and ocean regions.

Initially, we extract atmospheric wind velocities from the netCDF file corresponding to the mentioned pressure levels. This data results in a heatmap, depicted in Figure 3a. Secondly, these velocities get downsampled via the *downsample_heatmap* function to align with the resolution of the cloud coverage data.

To differentiate between land and ocean areas, we use the *global_land_mask* library. By iterating through the tracked cloud centroids at each time step, our script constructs a vector field by associating velocities with respective grid positions based on centroid locations. It then computes the MSE between the calculated cloud velocities and actual wind velocities, visualized in Figure 3b, separately for land and ocean regions. Additionally, an overall MSE encompassing all velocities is computed.



(a) One frame of the U-Component of the wind for a pressure of 250 hPa.



(b) Vector field of the combination of the downsampled distribution of the U-component and V-component of the wind for a pressure of 250 hPa.

Figure 3: retrievement and data treatment of the wind velocities.

The MSE calculations involve iterating through grid positions, comparing vector field velocities with the true wind velocities, and aggregating the squared differences. This process yields to an average MSE values distinct for land, ocean, and overall regions, quantifying the disparity between derived cloud velocities from centroid tracking and authentic wind velocities.

Finally, to evaluate how different cloud velocity is from wind velocity, we compute the standard deviation of the differences to understand the dispersion of errors from the mean.

3 Results

The methodology described was applied to the three distinct altitude regimes under consideration. MSE values were obtained for each wind component, the horizontal and vertical, corresponding to high, medium, and low cloud coverage, as detailed in Table 1. Additionally, an examination was conducted to explore potential differences in the correlation between clouds over land and those over oceanic regions.

	Standard Deviation (m/s)	Land MSE (m^2/s^2)	Ocean MSE (m^2/s^2)	Total MSE (m^2/s^2)
High	25.19	(22.25, 8.12)	(17.30, 9.75)	(18.81, 9.27)
Medium	12.56	(5.23, 2.53)	(3.90, 2.43)	(4.32, 2.46)
Low	10.17	(3.17, 2.06)	(2.16, 1.70)	(2.50, 1.82)

Table 1: General results for the MSE and the standard deviation of the cloud velocities compared to wind velocities.

As well as we did a quantitative analysis, a qualitative assessment is provided through attached videos showing the dynamic movement of clouds and winds. These videos illustrate the intermediate steps involved in arriving at the final calculations. Particularly noteworthy is the observation in *'high clouds full'* and *'u component of the wind high range'* where the movement of clouds appears linked to horizontal wind patterns.

3.1 High cloud coverage

The data presented in Table 1 reveals distinct MSE values for high clouds over both land and ocean regions. For high clouds observed over land, the MSE showed a higher magnitude in the horizontal component compared to ocean regions, indicating a more significant deviation in cloud movement patterns over terrestrial areas. However, the MSE for high clouds over ocean regions demonstrated a relatively lower discrepancy between velocities.

This discrepancy in high cloud velocities between land and ocean areas cannot be solely attributed to the geographical features of the land surface, as high-altitude clouds are minimally influenced by such features [3]. Further analysis is necessary to discern the underlying factors contributing to these disparities in cloud and wind velocities across distinct geographical regions.

3.2 Medium cloud coverage

Moving to medium clouds, the MSE analysis reveals distinct horizontal and vertical components for both land and ocean regions. Over land areas, medium cloud MSE showed a higher magnitude in the horizontal and vertical components compared to oceanic regions, signifying a relatively larger deviation in cloud movement patterns over terrestrial areas.

Similar to the high cloud distribution, the observations suggest a less correlated relationship between medium cloud velocities and wind velocities over land areas compared to oceanic regions. These differences in medium cloud velocities between land and ocean regions may be influenced by earth topography, especially when examining the coverage between 800 hPa and 450 hPa.

The overall MSE and the lower standard deviation for medium clouds suggests that medium cloud velocities exhibit a relatively closer agreement with actual wind velocities compared to high clouds. This could potentially indicate a more predictable association between these elements. Further investigation into the influence of earth topography on these disparities between land and ocean regions could provide valuable insights.

3.3 Low cloud coverage

Examining low clouds, the MSE analysis presented distinct horizontal and vertical components for both land and oceanic regions. Over land areas, the MSE values for low clouds displayed a more considerable difference in both horizontal and vertical components compared to ocean regions, indicating a larger deviation in cloud movement patterns over terrestrial areas.

The overall MSE for low clouds suggests a relatively closer alignment between computed low cloud velocities and actual wind velocities, further supported by the lower MSE values observed over both land and ocean.

3.4 General Analysis

Table 1 provides an overview of our findings, revealing an escalating trend in MSE values from low to high clouds. We suggest that this discrepancy might be caused by geographical factors, specifically how topography influences the relationship between cloud movement and wind patterns.

To elaborate, our analysis considers lower and medium cloud distributions, approximately ranging between 2 km and 6 km in a standard atmosphere. We hypothesized that high mountains may act as barriers, potentially inducing chaotic behavior in the lower range that complicates establishing a clear correlation between wind and cloud velocities. Contrary to our expectations, we observe an inverse trend, where higher clouds exhibit less correlation with wind velocity. We propose that this might be attributed to the fact that the chaotic behaviors happens at smaller scales, not accounted for due to the elevated threshold required for cloud detection. Then landmasses such as mountains and plateaus could act as larger-scale buffers, slowing down both cloud and wind movements uniformly. This influence on both elements could contribute to their higher overall correlation [2].

Moreover, variations in MSE values between land and ocean for different cloud levels suggest a potential influence from temperature discrepancies. The consistently higher MSE over land clouds compared to those over oceans cloud imply that land's faster heating and cooling capabilities may induce greater vertical wind motion. Land's lower heat capacity results in quicker temperature fluctuations, leading to stronger vertical movements and increased large-scale turbulence, consequently affecting cloud dispersal and making their general movement less dependant on planar winds.

It's noteworthy that horizontal wind velocity displays more significant variations across all scenarios and consistently yields higher MSE values. This could reflect the impact of Earth's landform. As the transition between water and land primarily occurs horizontally [**?**], more obstacles and changes in this direction could disrupt the correlation. However, the most substantial difference between vertical and horizontal winds is observed at higher altitudes, where landform changes are negligible. In this context, jet streams, characterized by intense west-to-east wind flows [4], may contribute significantly to these distinctions.

4 Discussion

We successfully replicated the EUMETSAT visualization of cloud movement [7], allowing the extraction of velocity information for comparison with actual wind velocity data. Furthermore, the centroid tracking system has proven its efficacy by generating plausible results, serving as a valuable resource for analysis and future comparisons in subsequent investigations. These results have the potential to offer insights into Earth's surface characteristics.

Our observations have indicated a decrease in the correlation between cloud velocity and wind velocity with increasing altitude. We attempted to provide a topological explanation for this behavior. To validate this supposition, further research should focus on localized areas with well-defined landmarks for precise analysis.

Additionally, we noticed a general trend where the Mean Squared Error is higher over land compared to oceans, implying a weaker correlation on land surfaces than over water bodies. This disparity might be explained by considering the faster temperature fluctuations in land compared to oceans. These rapid surface temperature changes may induce vertical air currents, leading to chaotic cloud movements. Confirming this hypothesis would require a simultaneous temperature analysis.

We've taken measures to avoid biases in our data analysis. The distribution of data points between land and ocean agrees with the Earth's surface ratio of 30% land and 70% water [8]. Therefore, our system does not exhibit discrimination towards any particular cloud cluster type, mitigating bias-related concerns in isolating clouds.

Furthermore, we've attempted to explain the dissimilarities between horizontal and vertical correlations, attributing these distinctions to Earth's surface shape, particularly the general vertical landform.

To finalize, regarding the computed standard deviation of the velocity differences, our results are reasonably close to the actual values. On average, the vector difference between cloud motion and ambient wind at the cloud base averaged 1.2 m/s, while at higher cloud levels, this difference ranged from 3 to 5 m/s [1]. Despite our results being generally higher than these real values by up to two or four times, we've observed a consistent trend of increasing values with altitude.

However, it's crucial to note that the standard deviation consistently surpasses the real differences. This discrepancy may arise when the tracking system misidentifies a cloud or registers it in distant locations, leading to an overestimation of traveled distances. This error occurs only two or three times per frame, as evident in the centroid and velocity videos *'centroids and velocities'*. Although it does not significantly impact our overall understanding of cloud behavior, it contributes to an increase in the MSE across all analyses.

5 References

References

[1] A. F. Hasler, W. Shenk, & W. Skillman (1976) *Wind Estimates from Cloud Motions: Phase 1 of an in situ Aircraft Verification Experiment*, Journal of Applied Meteorology and Climatology, 15(1), 10-15

[2] Moore, D. S., McCabe, G. P., & Craig, B. A. (2020). *Introduction to the practice of statistics* (9th ed.). W. H. Freeman.

[3] Ackerman, S. A., & Knox, J. A. (2019). Meteorology: Understanding the Atmosphere (3rd ed.). Jones & Bartlett Learning.

[4] Ahrens, C. D. (2018). Meteorology today: An introduction to weather, climate, and the environment (12th ed.). Cengage Learning.

[5] J. Rover et al. (2020). Land change monitoring, assessment, and projection. United States Geological Survey. 2020-3024. doi: 10.3133/fs20203024

[6] European Centre for Medium-Range Weather Forecasts. (2022)ecCodes Home. ECMWF. https://confluence.ecmwf.int/display/ECC/ecCodes+Home

[7] Looking back over a year of weather. (2020). EUMETSAT. https://www.eumetsat.int/features/looking-back-over-year-weather

[8] Petersen, J., Sack, D., & Gabler, R. E. (2018). Fundamentals of Physical Geography (2nd ed.). Cengage Learning.