

HK Observatory Weather Forecast Accuracy

Analysis of HKO's 1 day and 7 day forecasts accuracy in predicting temperature range and rainfall with R.

COMP2501 Project
Name: Kam Ada Yi Man
UID: 3035662493

Table of Contents

1. [Rationale and background research](#)
2. [Research Question](#)
3. [Background Information](#)
4. [Methodology](#)
5. Results
 - a. [1 day forecast temperature range](#)
 - b. [7 days forecast temperature range](#)
 - c. [1 vs 7 days forecast rainfall](#)
6. [Discussion](#)
7. [Limitations and Future Works](#)
8. [References & Source Code](#)



1. Rationale and Background Research

We check the weather forecasts daily to

- a) Prepare how to dress
- b) Plan for outdoor activities
- c) Prepare to bring extra gears...



BUT just how accurate are weather forecasts? Are **some elements more predictable** than others?

Sometimes we **criticise HKO's forecasts** for their gross **inaccuracies**, are these criticisms **justified**?

Past research has been done on weather forecast accuracies in **other countries** around the world.

In Hong Kong, HKO has **published statistics** on how accurate they claim their predictions to be.

This research will apply methodologies to assess the accuracy of HKO's weather predictions.

2. Research Question

How accurate is the HKO's 1-day and 7-day forecasts of HK's temperature range and rainfall between 2021 to 2023?

3. Background Information

1. HKO Disclaimer

- "... average accuracy for the next one to three days, four to seven days, and eight to nine days are about **90%, 85%** and 80% respectively. The errors of PSR forecast in the next one to six days are within around **10%** and those in seven to nine days are within around **20%**."
- PSR forecast on each day indicates the **probability of the accumulated rainfall** generally over Hong Kong **reaching 10 mm**

2. Measuring Accuracy

- Multiple ways of defining and measuring accuracy

a. HKO

- i. Formula: $\pm 1^{\circ}\text{C}$

b. Mean Bias Error

- i. Formula: $\text{sum}(\text{forecast} - \text{recorded})$
- ii. Implications: Overall direction of error, detects **systematic bias** in data

c. Root Mean Squared Error

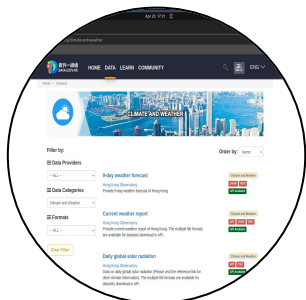
- i. Formula: $\text{sqrt}(\text{average}(\text{forecast} - \text{recorded})^2)$
- ii. Implications: Penalizes **greater absolute errors**

d. Brier's Score

- i. Formula: $(\text{predicted probability} - \text{outcome})^2$
- ii. Implications: Measure accuracy of **probabilistic forecasts**

4. Methodology

Data Source



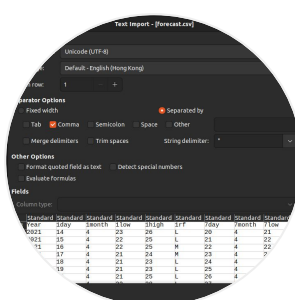
- Publicly accessible data provided by the government:
- 9-Day weather forecast
- Daily Total Rainfall
- Daily Maximum, mean, minimum Temperatures

Data Extraction



- VIA API
- Available daily data: 14-04-2021 to 28-02-2023
- HKO data for recorded data

Data Wrangling



- “Trace” rainfall = 0.1
- Date columns
- Removed empty values (~ 14 rows)
- Remove duplicates
- Join data by forecast and recorded date

Data Summary

Forecast:
673 obs. of 11 variables:
\$ Year : int 2021 ...
\$ X1day : int 14 15 ...
\$ X1month: int 4 4 4 ...
\$ X1low : int 23 22 ...
\$ X1high : int 26 25 2 ...
\$ X1rf : chr "L" "L" ...
\$ X7day : int 20 21 ...
\$ X7month: int 4 4 4 ...
\$ X7low : int 21 22 ...
\$ X7high : int 26 26 ...
\$ X7rf : chr "L" ...

Recorded:
673 obs. of 6 variables:
\$ Year : int 2021 ...
\$ Month : int 4 4 4 ...
\$ Day : int 13 14 ...
\$ HKO...MAX: num ...
\$ HKO...MIN: num ...
\$ HKO...RF : chr "0" ...

5. Results

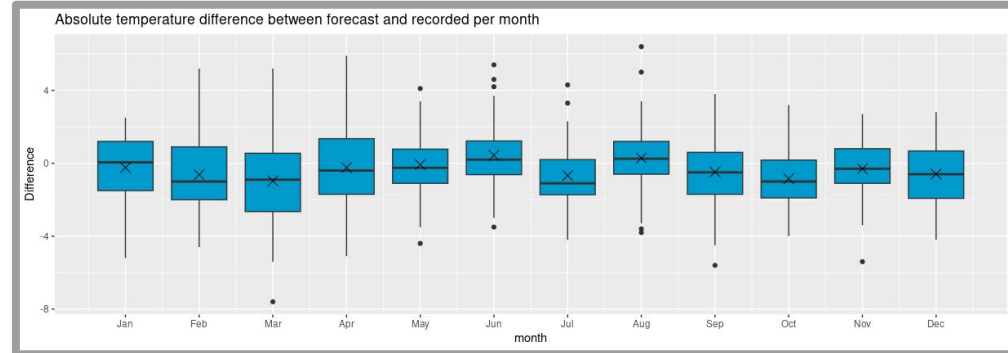
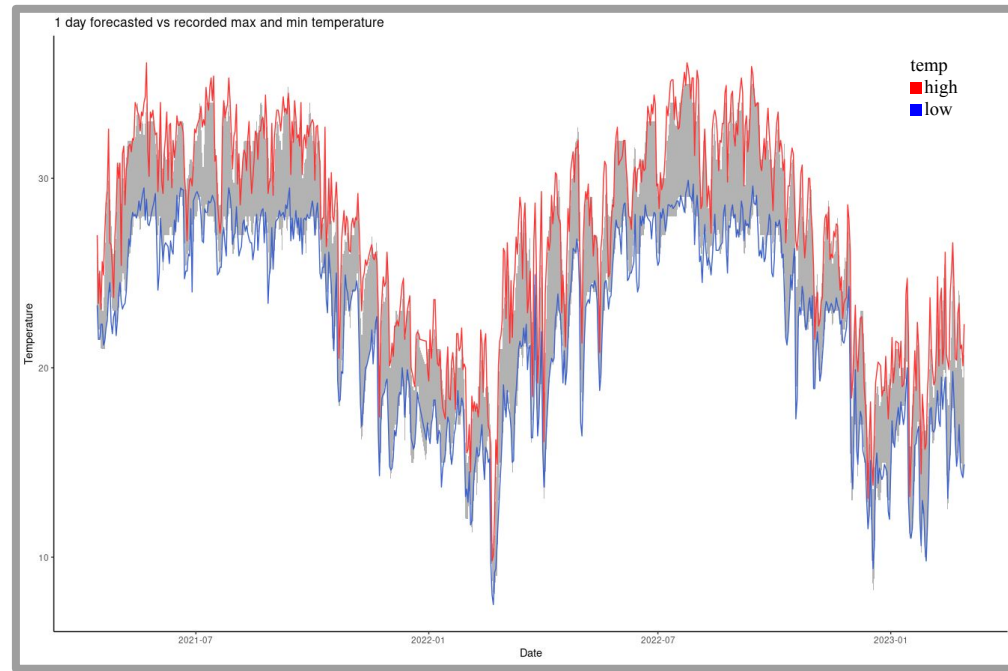
1-day forecast temperature range

	Upper Range	Lower Range
Mean Bias Error (MBE) <i>sum ((forecast - recorded) / observations)</i>	-0.236	-0.095
Root-mean-square Error (RMSE) <i>sqrt ((forecast - recorded)² / observations)</i>	1.293	0.941

[Top Left] Table summary of **statistical accuracies** of 1 day forecasted temperatures in upper and lower range.

[Top right] **Line chart** showing the actual **max temperature in red**, **min in blue**. **Forecasted range is shaded in grey**. Ideally easy to identify inaccuracies based on colors e.g. shaded grey area out of bounds, but graph is slightly claustrophobic because of vast data points and **fluctuations in daily temperatures**.

[Bottom right] **Box plot** of absolute **temperature difference** between forecast and recorded in **each month**. **Outliers** represented as dots, 'x' marks **mean** temperature difference.



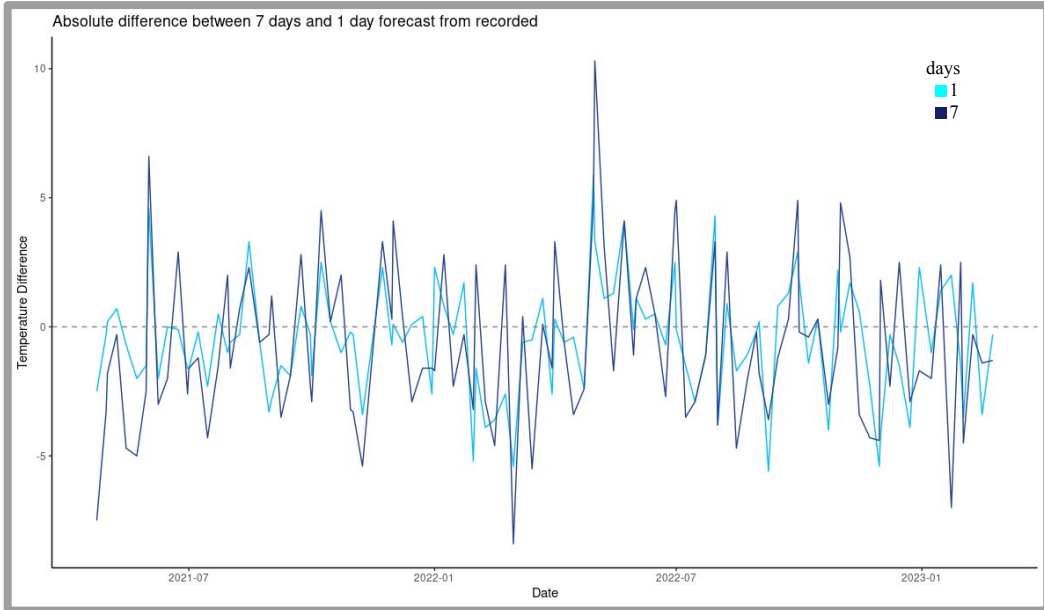
5. Results

7-days forecast temperature range

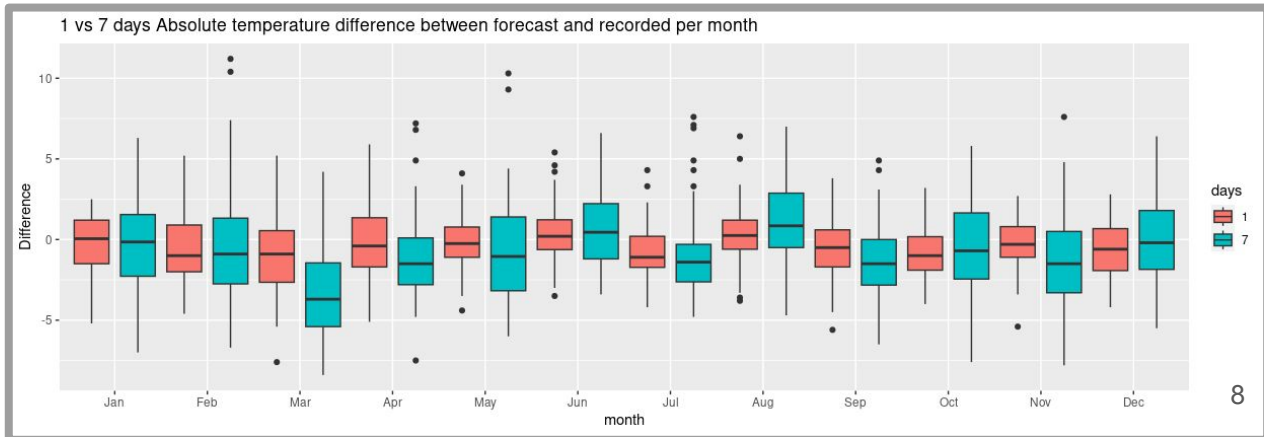
[Top Left] Table summary of **statistical accuracies**.

[Top Right] **Line chart** comparison of 1 day vs 7 day absolute difference in temperature. **Visual representation** of difference, closer to the 0 means more accurate, can easily **spot outlying spikes** for dark blue on some days.

[Bottom Right] **Boxplot** summaries of difference to compare **distribution of results**, **average** of results and **outliers** between each month and the days of forecasts.



	Upper Range	Lower Range
MBE	-0.236 → -0.472	-0.095 → -0.065
RMSE	1.293 → 1.936	0.941 → 1.508



5. Results

1 vs 7 day forecast rainfall

Forecast	Implied %	Actual % (1 → 7 days)	Rainfall % (1 → 7 days)
High	>= 70%	73.4 → 16.7%	96.9% → 75.0%
Medium High	55-69%	30.0 → 32.1%	96.7% → 71.4%
Medium	45-54%	32.4 → 20.0%	100% → 76.0%
Medium Low	30-44%	30.0 → 22.8%	79.5% → 69.5%
Low	< 30%	2.56 → 9.17%	49.9% → 39.2%

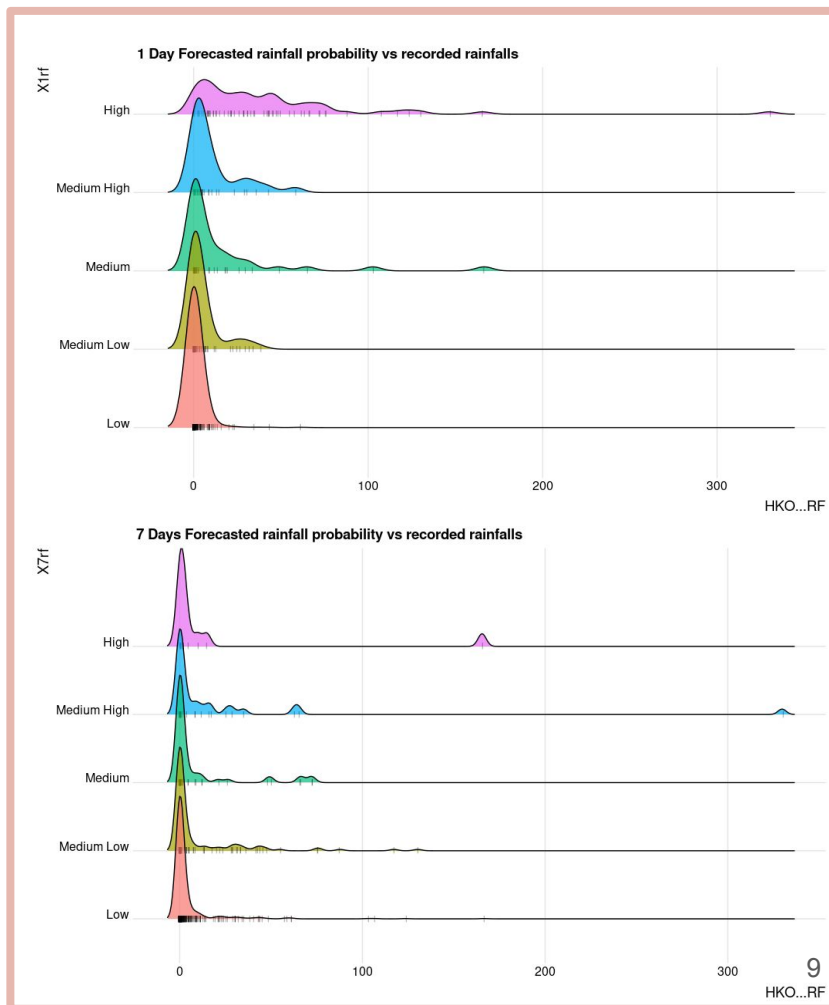
[Left] Forecast: Classification used by HKO forecast in PSR.

Implied %: Given by HKO.

Actual %: For 1 day and 7 day forecasts, how many days (in %) were there actually significant rain? (**Direct comparison**)

Rainfall %: For 1 day and 7 day forecasts, how many days (in %) were there any rainfall? (**Looser definition** of HKO's definition to compare accuracy with less focus on precision)

[Right] **Ridgeline plot** to show **distribution** of actual rainfall volume for each forecast. Clear **visual comparison** between 1 vs 7 day forecasts and each classification



5. Results

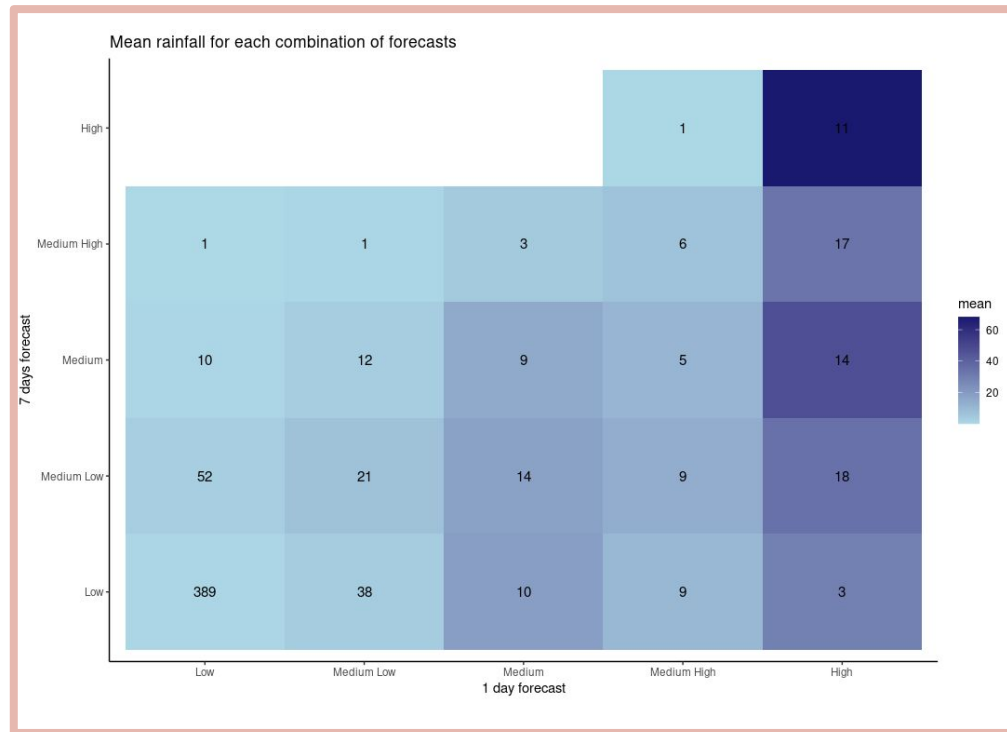
1 vs 7 days forecast rainfall

	1 day	7 days
Brier Score	0.353	0.438

[Left] Table of Brier Score. Score uses the lower bound of each classification in the formula. **Lower score indicates higher accuracy.**

- How to interpret
- If PSR is High, it means forecasted PSR is 0.7%. If actual day rainfall $\geq 10\text{mm}$, outcome = TRUE (1), else outcome = FALSE.
- So if none of the days have significant rain, Brier score is 0.7%

[Right] **Tile plot** to show **correlation** between the change (or unchanged) forecasts from 7 to 1 day. **Color** represents actual rainfall, **number labels** are number of days with the respective predictions



6. Discussion

1. 1 day vs 7 days forecast accuracy
 - a. Better statistical scores for 1 day than 7 days forecasts
 - b. Outlier
 - i. 7 day LOWER TEMPERATURE RANGE slightly less biased
 - ii. For 7 day PSR forecasts, “Medium High” has the highest probability of significant rain, “High” has the 2nd to lowest
 - c. 1 day PSR forecast of “High”, “Medium High”, “Medium” has > 95% of rain, drops to around 75% for 7 day
2. Max & Min temperature forecasts
 - a. Consistent negative bias (recorded > forecast)
 - b. Seasonal Trends
 - i. April, May, Jul, Aug has the most outliers (Summer months)
 - ii. No outliers for Dec, Jan, Oct (Colder months)
 - iii. Typhoon? More unpredictable weather?
3. PSR forecasts tile map
 - a. “Medium” 1 day forecast more likely to rain than “Medium High”?
 - b. More days of escalated forecasts from 7 to 1 days than lowered forecast

7. Limitations and Future Works

Choices of statistical measures

- Brier Score
 - Rainfall probability given as a range
 - The analysis only took lower bound
- Missing statistical significance analysis
 - Results not robust because no analysis is performed on statistical inferences
 - e.g. Pearson's value

Lack real-world implications

- No benchmark to compare against
 - Only an objective figure of accuracy
- Error vs Discernment vs Uncertainty
 - Difference with actual value, segmentation across days, uncertainty in variance
 - Combine all factors for more meaningful model

Compare across districts and other forecasts

- Measure discrepancies between urban vs rural?
- Compare accuracies between weather stations?
- Find maximum inaccuracies?
- Objectively benchmark accuracy
- Evaluate shortcomings of HKO's model statistically
- More practical implications on which model to use to access information

Machine Learning Models

- Predict outcome based on forecast?
- Look into trends within forecasts and how they may or may not real life?
- Identify correlations between different forecasts variables, e.g. extended days of above average heat may foreshadow typhoon?

8. References & Source Code

Babb, D. (2019). *Assessing Forecast Accuracy* | *Learning Weather at Penn State Meteorology*. Learningweather.psu.edu. <https://learningweather.psu.edu/node/77>

DATA.GOV.HK. (n.d.). | *DATA.GOV.HK*. Data.gov.hk. Retrieved April 30, 2023, from <https://data.gov.hk/en-datasets/category/climate-and-weather>

Hong Kong Observatory. (n.d.). *9-day Hong Kong weather forecast provided by the Hong Kong Observatory*. Wwww.hko.gov.hk. Retrieved April 30, 2023, from <https://www.hko.gov.hk/en/wxinfo/currwx/fnd.htm>

Jason. (2016, June 28). *How to measure the accuracy of forecasts*. @ASmartBear. <https://blog.asmartbear.com/forecast.html>

Source Code

<https://github.com/kamada01/HKO-Weather-Forecast-Accuracy>