

晴天钟(7TIMER!)第一期模型预测准确度研究

7TIMER 项目负责人/叶泉志

The accuracy study of 7TIMER!'s first model

By Quanzhi Ye, the director of 7TIMER!

• 前言

晴天钟已经进行了一个多月的试运行，其准确度是大家所关心的。为了让大家对目前晴天钟模型的准确度有一个了解，也方便我开展模型修改工作来完善晴天钟，从7月25日到8月4日，我在海口、广州和南京的合作者进行了实测，我汇总了实测和预测数据并进行了分析，下面就是分析的结果。

云量问题是晴天钟的核心，因此这次研究仅针对云量。在完成这篇论文之后，我将对晴天钟云量模型进行修改，并在完成之后再发布一篇关于云量预测说明的论文。

• Introduction

After more than a month's test run, you might be interest at the accuracy of 7TIMER!. So for your interests, and for my intend to improve the model, from July 25 to August 4, some users from Haikou, Guangzhou and Nanjing made observations for me, to compare the observation and the forecast, so we could know exactly about the accuracy.

The cloud cover is the most important part of 7TIMER!, so I only did study on it. After finish this study, I'm going to modify the cloud cover model, and I'll publish another paper on it after I finish modifying.

• 观测数据

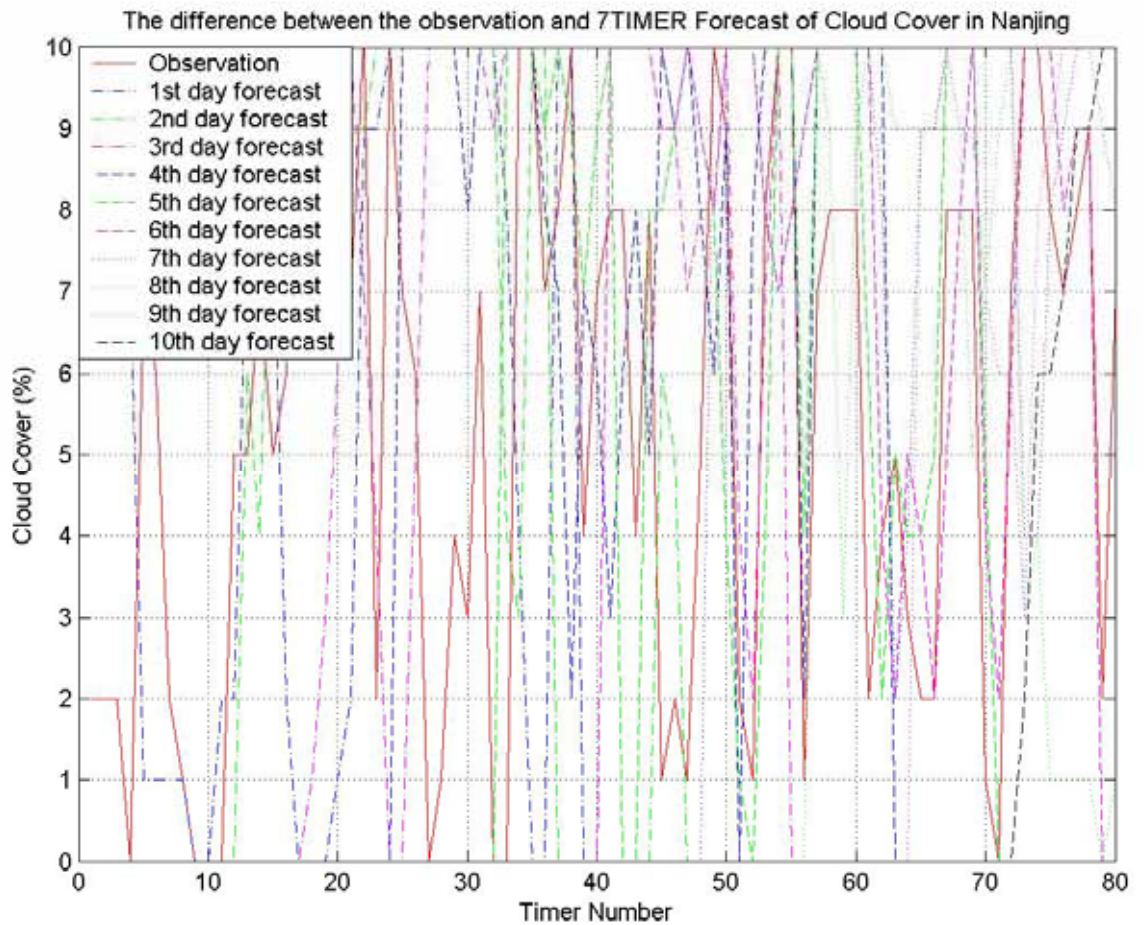
数据采集在三个点进行：南京（虞俊）、广州（叶泉志）和海口（AAXX 数据），观测时间为7月25日23时到8月4日20时，每3小时观测一次云量，对空缺以及质量低处用 AAXX 数据予以填补。同时，7TIMER 将 25 日-3 日的上述地点预测存档，以供分析比较之用。

预测存档文件共 11 个，每个文件从当日 23 时开始进行 120 小时的预测（因此没有 4 日的预测）。

以下是三个地点的观测/预测曲线绘图，其中的红线是观测值，不同的线代表不同时间的预测。（从上到下为海口、广州、南京）

• Observations

The observations are made in three different sites: Nanjing (by J.Yu), Guangzhou



• 分析结果

上面的图象可能比较复杂，以至于不好看懂。那没关系，下面是根据上面的时间所综合的分析结果。

1. 平均偏差率

将所有观测站的所有同一时长预测数据偏差平均所求得的值，这是为了看看在不同时长预测数据相对于真实值的趋向。

上面的解释可能让人一头雾水，请让我举个例子：

预测时间/预测差	25日14时发布	26日14时发布	27日14时发布	平均偏差率
9小时后	+3	-3	+3	+1
12小时后	+5	-2	-3	0
15小时后	+0	+7	+5	+4

预测差：云量分为11级，0~10；预测差=实际云量-预测云量。

根据上面的例子可以看出，在样本中，15小时后的预测值总体偏高，而12小时后的预测值总体接近实际值。下面是平均偏差率曲线。

*时间号(Timer Number)：等于实际预测时间除以3。比如预测60小时后，时间号=60/3=20。

• **Analyze result**

The images above might be difficult to read, so you can read these results, they might be easier for you to understand.

1. Average difference

The average difference between the real rate and the forecast rate by forecast time.

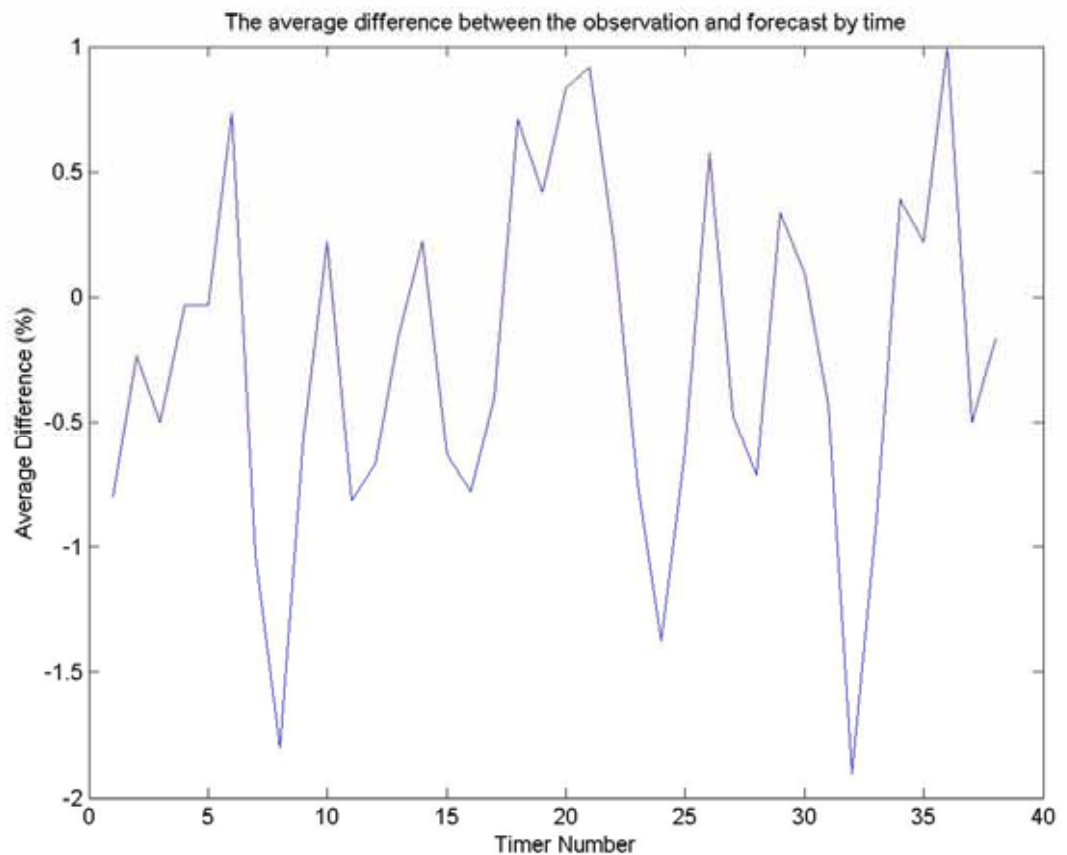
And this explanation might not be easy to understand, let me give an example.

Forecast time/the difference between obs and real rate	Released on 14h, Jul 25	Released on 14h, Jul 26	Released on 14h, Jul 27	The average difference
+9h	+3	-3	+3	+1
+12h	+5	-2	-3	0
+15h	+0	+7	+5	+4

The difference between observations and the real one=Real cloud cover (0~10) – forecast (0~10)

And you can see, in +15h the forecast average is higher than the real one.

*Time Number: The forecast time / 3. e.g. Forecast time is +60h, Time Number = $60/3 = 20$.



2.皮尔森偏差系数

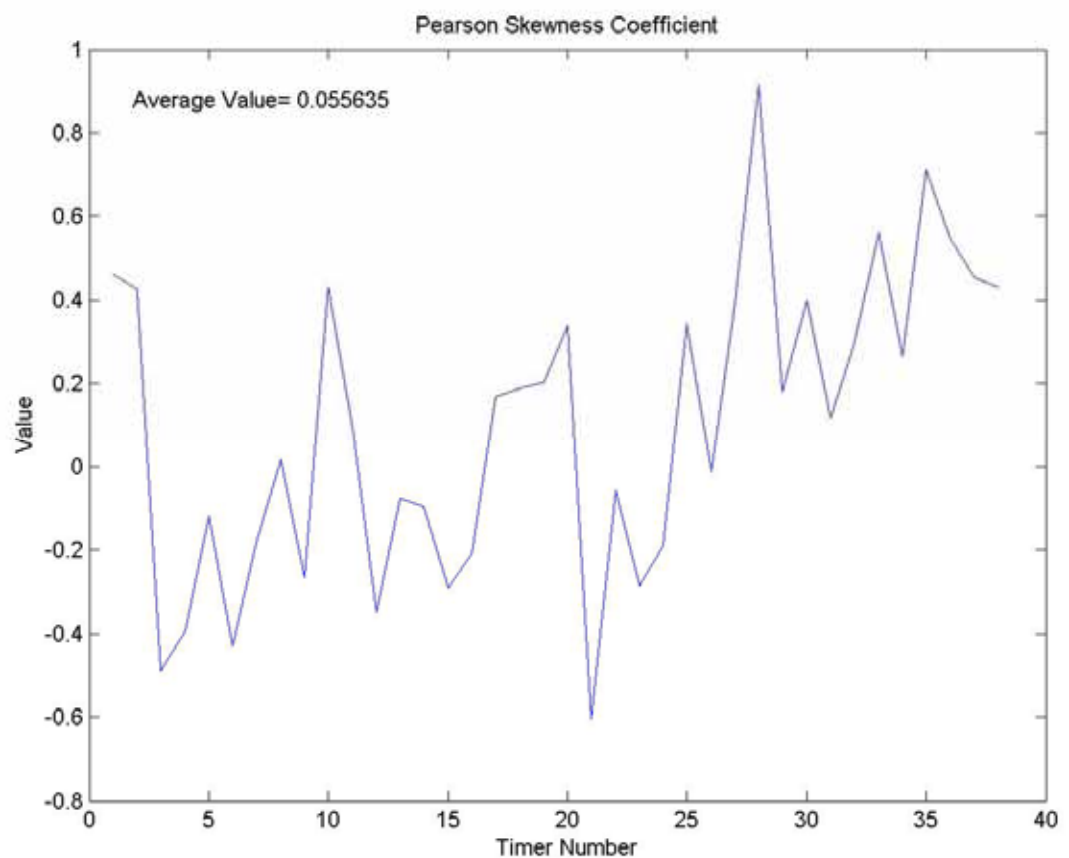
上面的方法并不是那么科学的,因此皮尔森偏差系数应该能够给我们一个更为准确的值。

皮尔森偏差系数用相对的办法求得预测值和实际值的偏差,从结果来看偏差系数为+0.05,稍微偏高,因此要将云量模型的输出改低一些。

2.Pearson Coefficient

The method above is not so scientific so, I use another one to calculate the average difference.

The Pearson Coefficient gives the results +0.05, a little bit higher, so we should discard the cloud cover output a bit...



3.预报准确率

预报准确率由如下公式计算:

$$A = [10 - \text{average}(C)] * 10$$

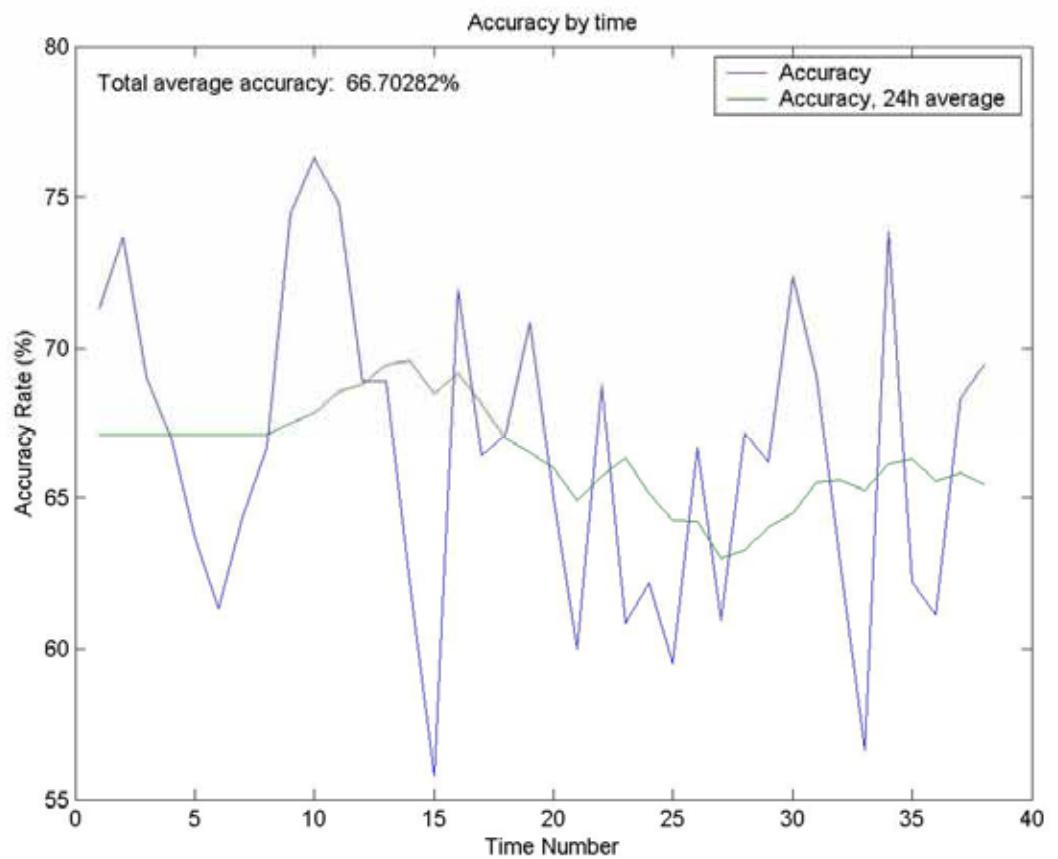
其中 A 为准确率, C 为指定时间内的云量。计算所得的最终准确率为 66.7%。

3.Accuracy

I use this formula to calculate the accuracy:

$$A = [10 - \text{average}(C)] * 10$$

A is accuracy, while C is the cloud cover in period. The result is 66.7%.



• 结尾语

虽然 66.7% 不能算是很高，但考虑到数据源的精度为 1 度，我想准确度还是大家都可以接受的。

由于云量预测由几条函数组成，其中有若干个参数是可以调整的。完成本文之后，我将对这些参数进行调整，以保证模型发挥到最佳性能。此外，我还打算对其他几个量的准确度进行测试和调整。

• References

The GFS Model description, NOAA, <http://www.emc.ncep.noaa.gov/gmb/moorthi/gam.html>

August 7, 2005