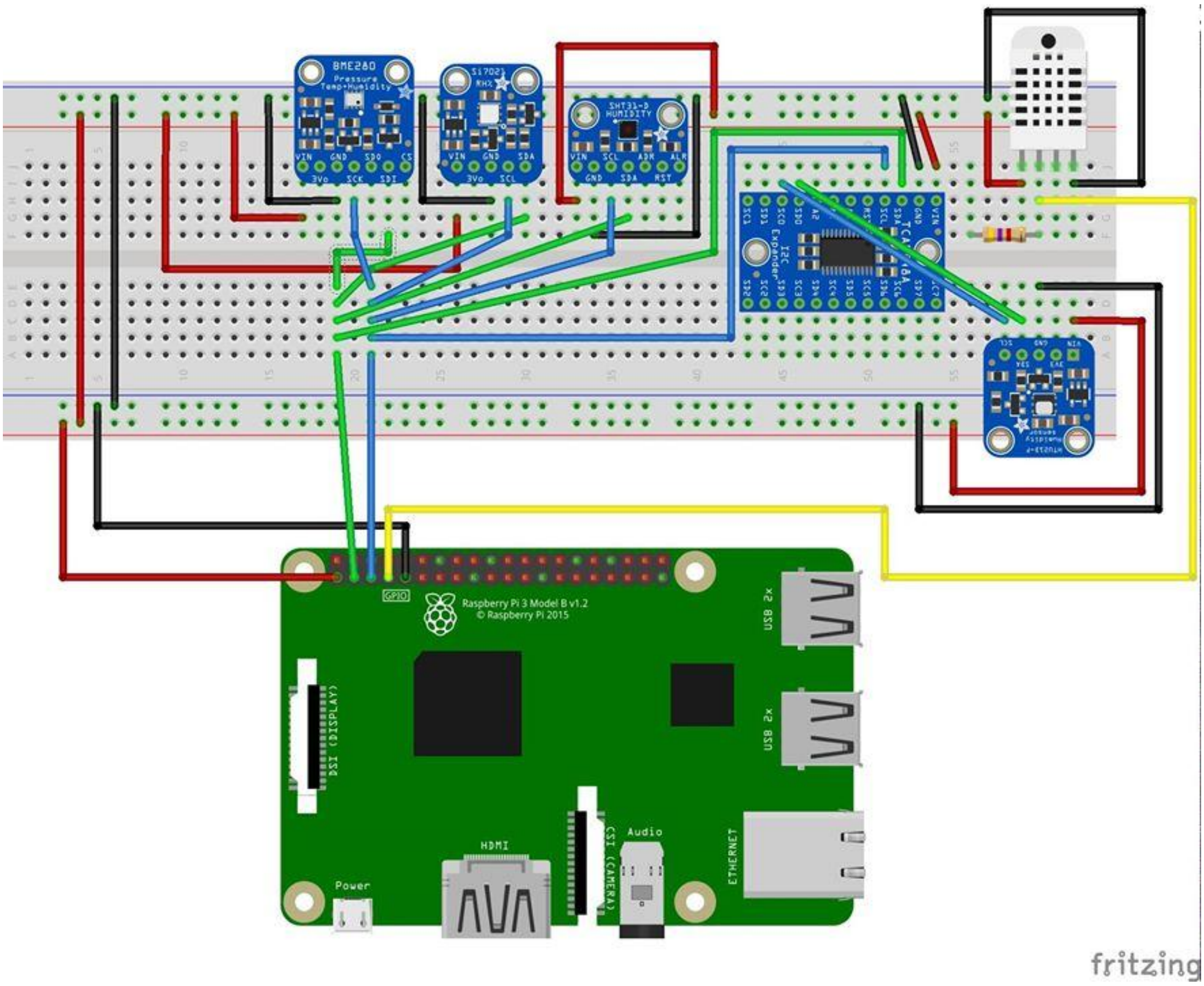
Comparing Temperature/Humidity Sensors

Adafruit has a number (i.e., a lot) of options for temperature and humidity sensors. Most of them use I2C to communicate (the exception is the DHT22, which uses a special one-wire data protocol). Some of these sensors are supposed to be more accurate than others, which may be true, and some are supposed to be faster at data acquisition than others, which is definitely true. I decided to compare sensors using a Raspberry Pi 3 for sensor control and data acquisition and a fairly powerful laptop running R (an extremely powerful statistical analysis programming environment - check it out) to do the analysis.

The sensors I used were the Adafruit HTU21D-F, the Adafruit SHT31-D, the Adafruit DHT22, the Adafruit BME280 breakout, and the new Adafruit Si7021 sensor breakout.

My test procedure was to wire all the sensors to a single Raspberry Pi 3. Since the Si7021 and the HTU21D-F both use 0x40 for their I2C address, and it can't be changed, I had to include the Adafruit TCA9548A 8-to-1 I2C multiplexer to resolve the address confilcts. Here is a scruffy Fritzing diagram showing how I wired it together onto a single I2C bus.



Wiring diagram for temp/humidity test environment

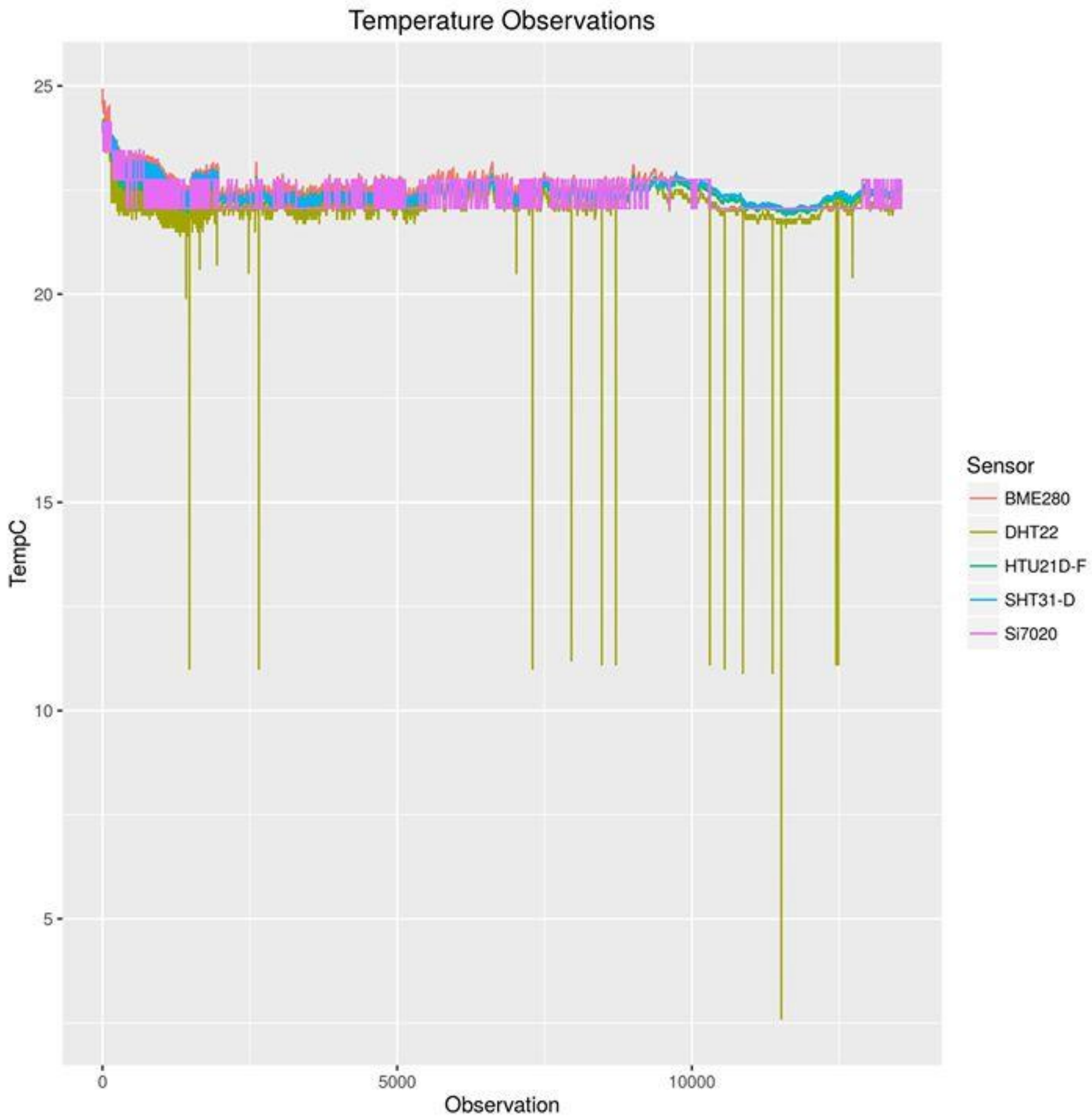
As you can see, all the I2C SCL lines are pulled together onto one row of the breadboard with blue wires, and all the SDA lines are pulled to another row with green wires. The DHT22 is the only non-I2C device in the circuit; it's connected to the Raspberry Pi GPIO pin 4 (the yellow wire). The data line has a 4.7K Ohm pull-up resistor, as needed.

I wrote a Python program to acquire data from each sensor every two seconds. This data acquisition rate was imposed by the DHT22 update rate. It’s fast enough, though, to highlight any random variance in the data over time.

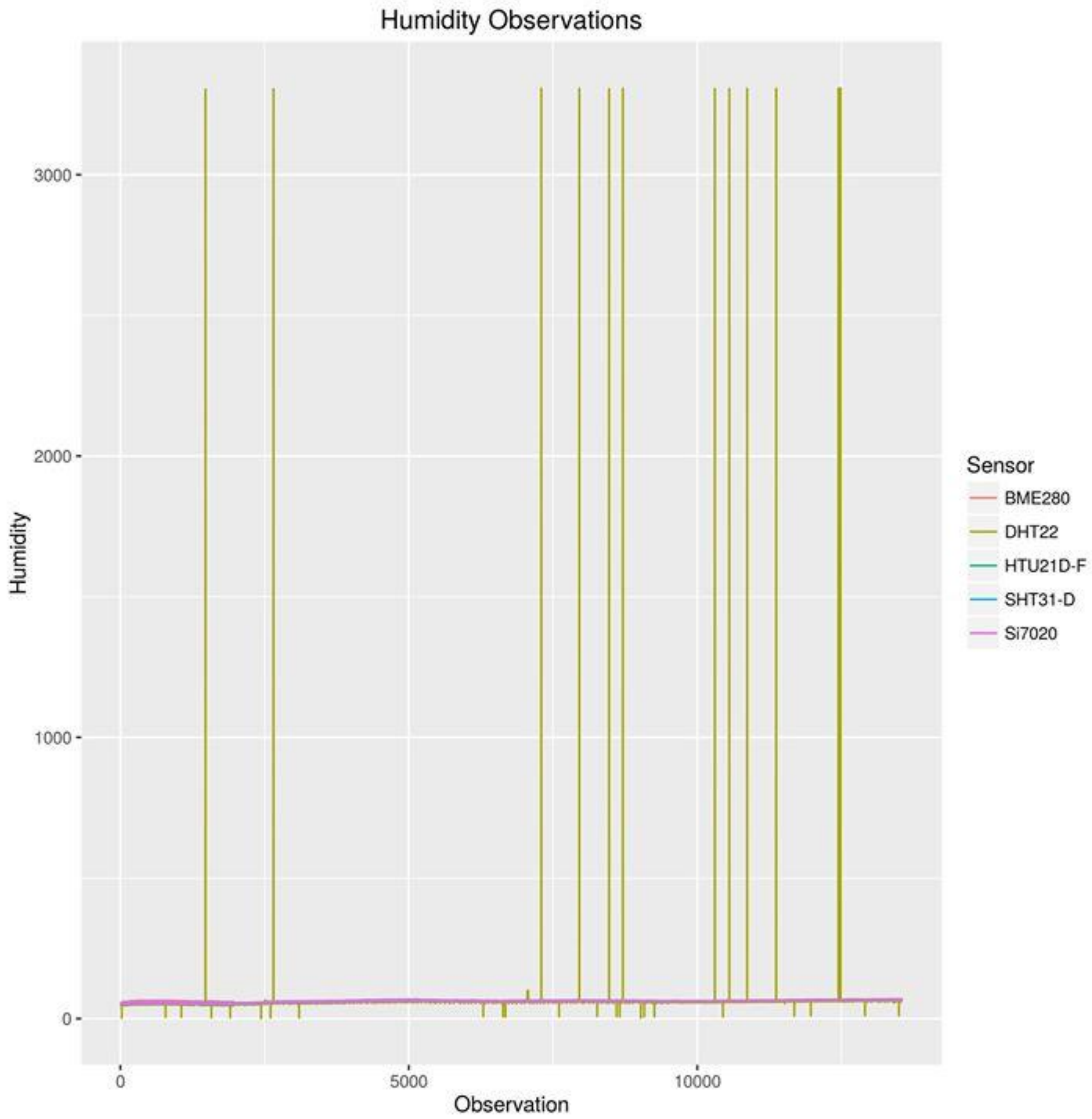
It also (as you will see) is detailed enough to show the resolution of the different sensors. I wrote the Python program myself, but I relied heavily on GitHub for the different sensor drivers. The authors and licenses are included in the program headers; many thanks to the original authors who did most of the heavy lifting in the program. The software is open source. I will upload the program and supporting libraries to GitHub and post a link. I then loaded the data into R, cleaned and munged it, and used the ggplot2 library to generate the graphs. I also generated summary statistics, which I discuss below. I will include the R program in the GitHub repository. Keep an eye on this space for the link.

In this note I am discussing both the temperature and humidity data analysis. If you have questions, or want other analytics run, just send me e-mail or message me on FB and I'll probably add the new analysis you'd like performed. 🙂:-)

On to the results! First we'll look at the raw temperature and humidity data.

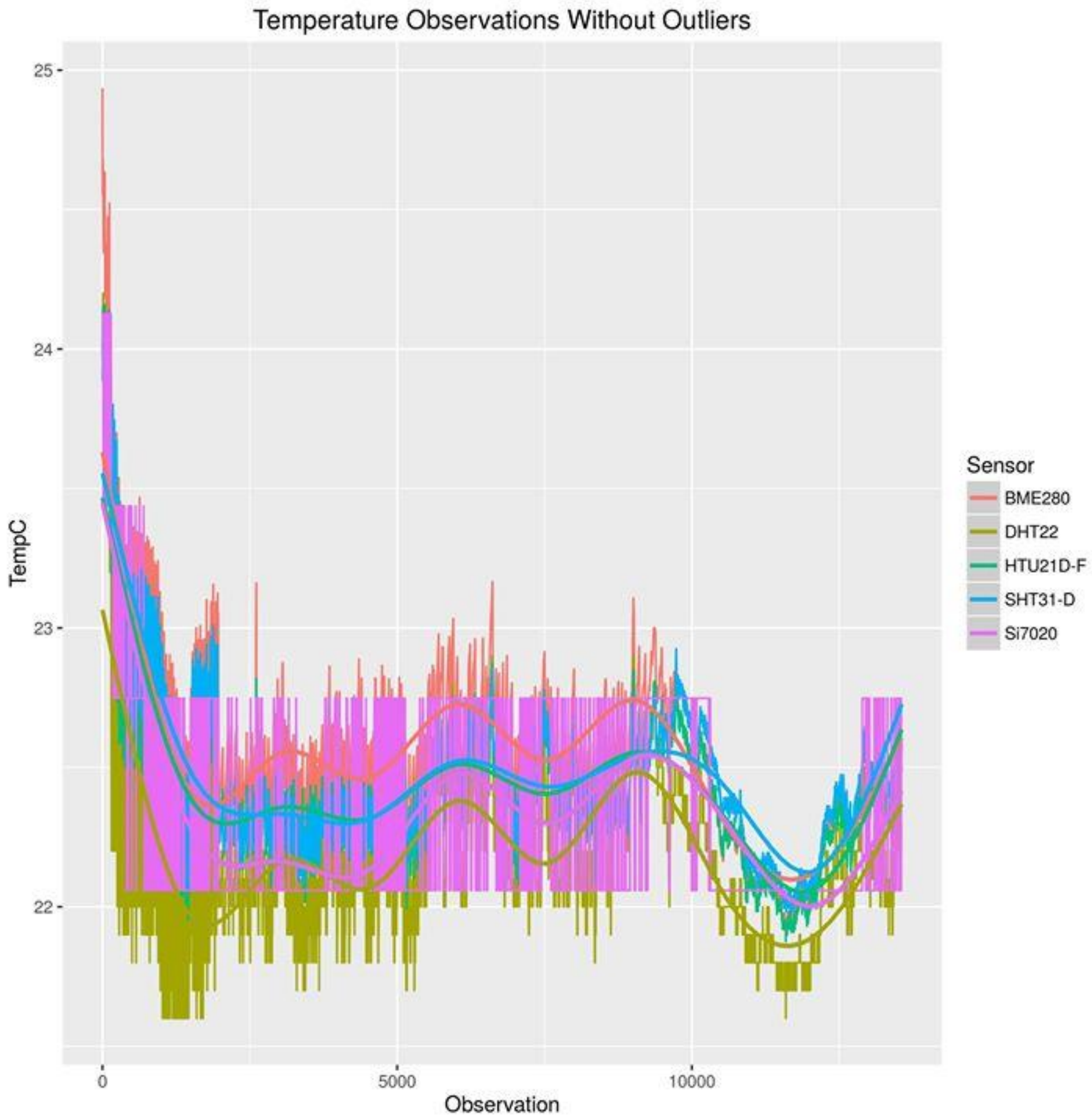


As you can see, the DHT22 occasionally generates extraneous outliers in the data set. We'll have to remove those.



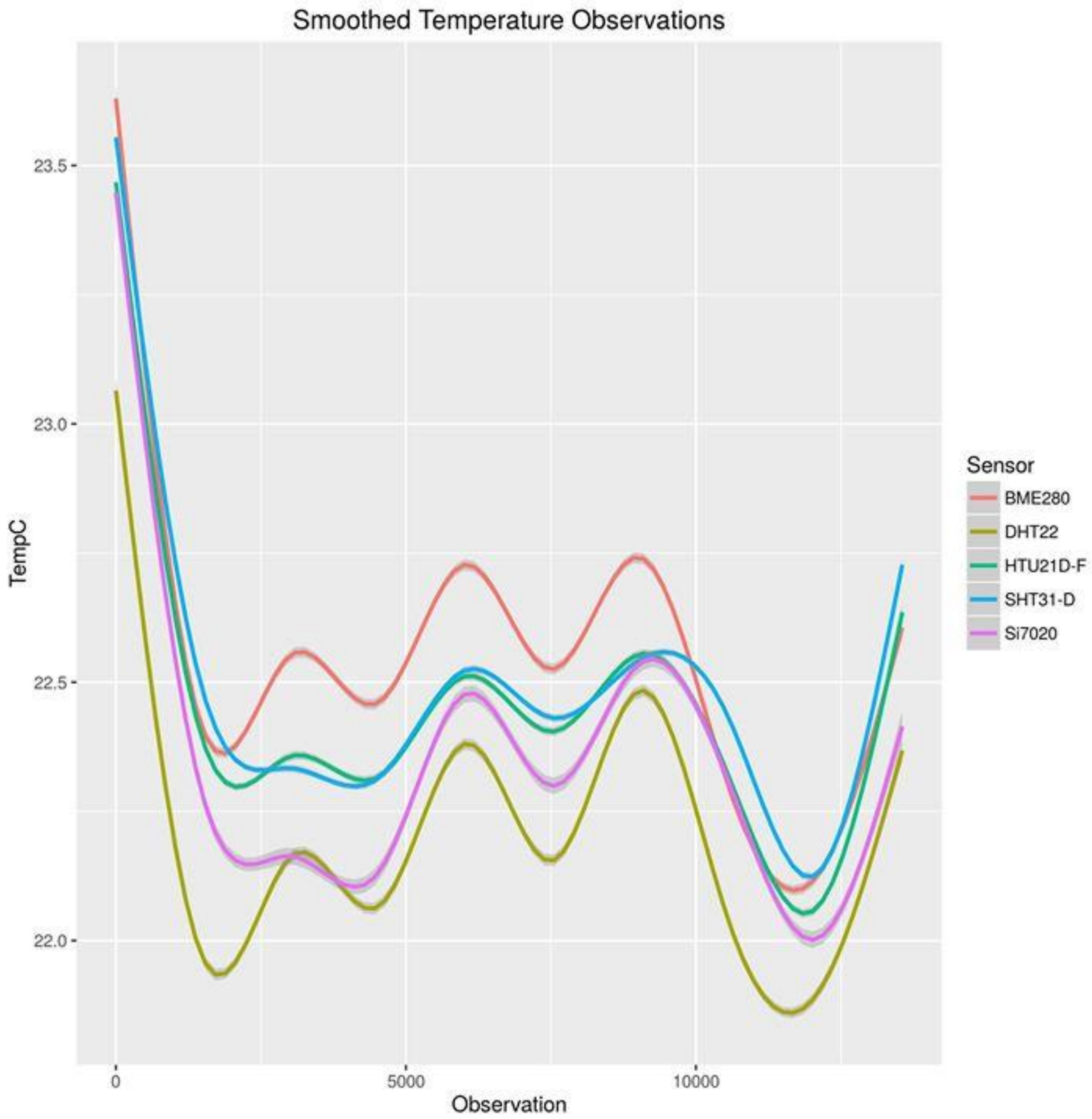
Raw Humidity Data

As you can see here, the DHT22 also generates extraneous outliers in the humidity data. After removing the outliers, we get more interesting and informative graphs. Incidentally, the number of observations in the data set is 13,559 (I left the data collection program running on the Raspberry Pi for quite a while since I was watching Elon Musk's speech about making humans a multi-planetary species, followed by replays of rocket launches, anyway).



Temperature data with outliers removed and smoothing curves added

These are the raw data plotted as a time series. It includes all five sensors. It's interesting to me that, first, the room was warm before I turned on the air conditioning and, second, that the air conditioning doesn't do a very good job of keeping the temperture in the room stable. The variability in the sensor readings, which is error fluctuation, makes this graph kind of hard to read. Here is the same graph with the data plots removed to show only the smoothed data:



Smoothed Temperature Data Time Series

Some observations on this figure.

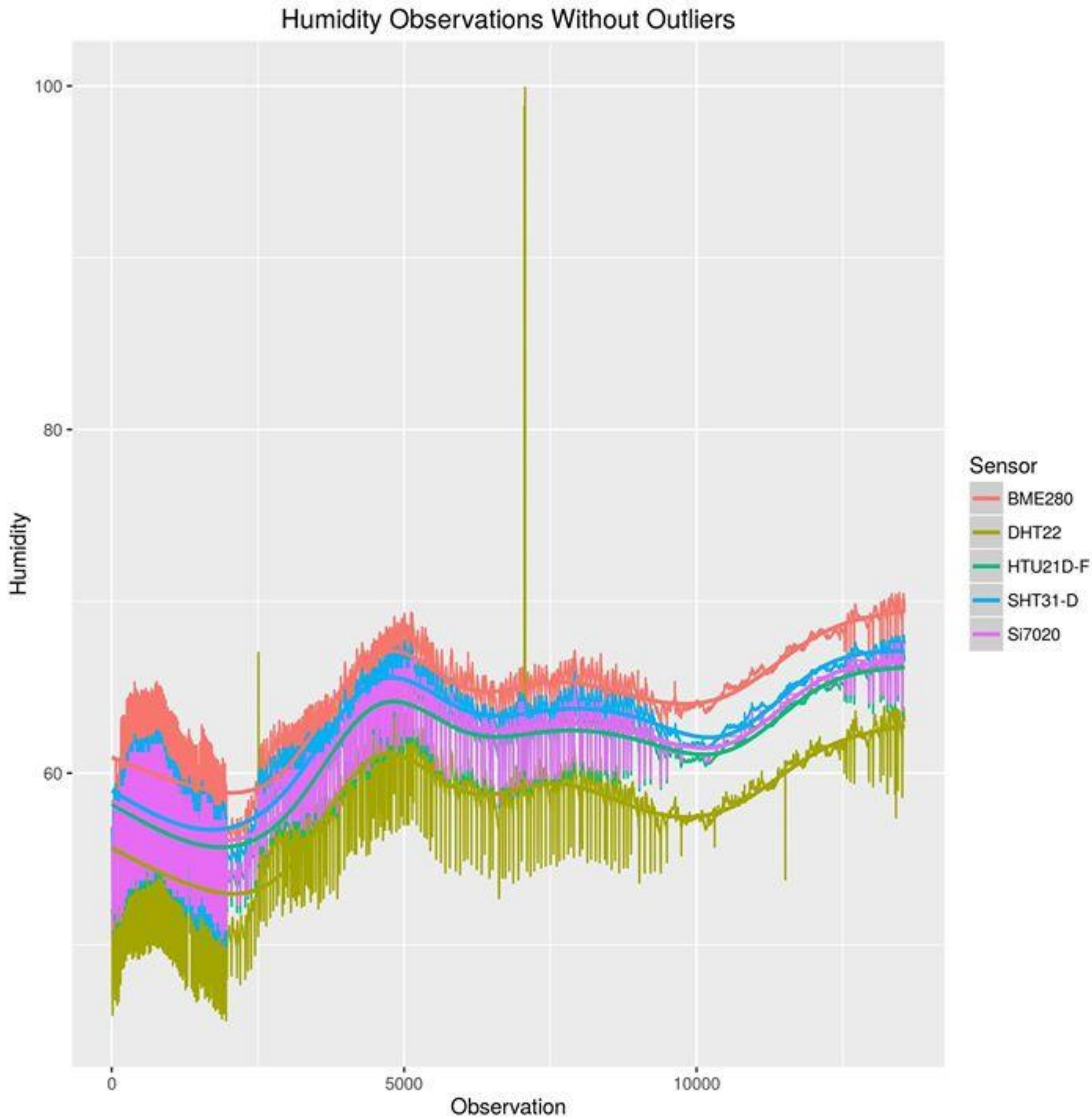
The differnet sensors roughly follow the same patterns but some are consistently high (the BME280) and some are consistently low (the DHT22).

Certain sensors are much more sensitive to shifts in temperature, while others appear to have a built-in smoothing algorithm. The most touchy sensors are the BME280 and the DHT22. I think I would also add the Si7020 (the purple line) to the list.

The room tempertaure definitely fluctuates over time. I don't know the actual room temperature because I don't have a sensitive lab-quality recording thermometer to give me a "real" temperature in the room. I think any of the sensors will give you a relatively correct answer but I would pick one close to the center of the graph, and one that is less sensitive to transient temperature changes, for applications like home automation (otherwise your heat will come on every time someone opens the door to come in or go out). The most stable sensors, from this perspective, are the HTU21D-F and the SHT31-D.

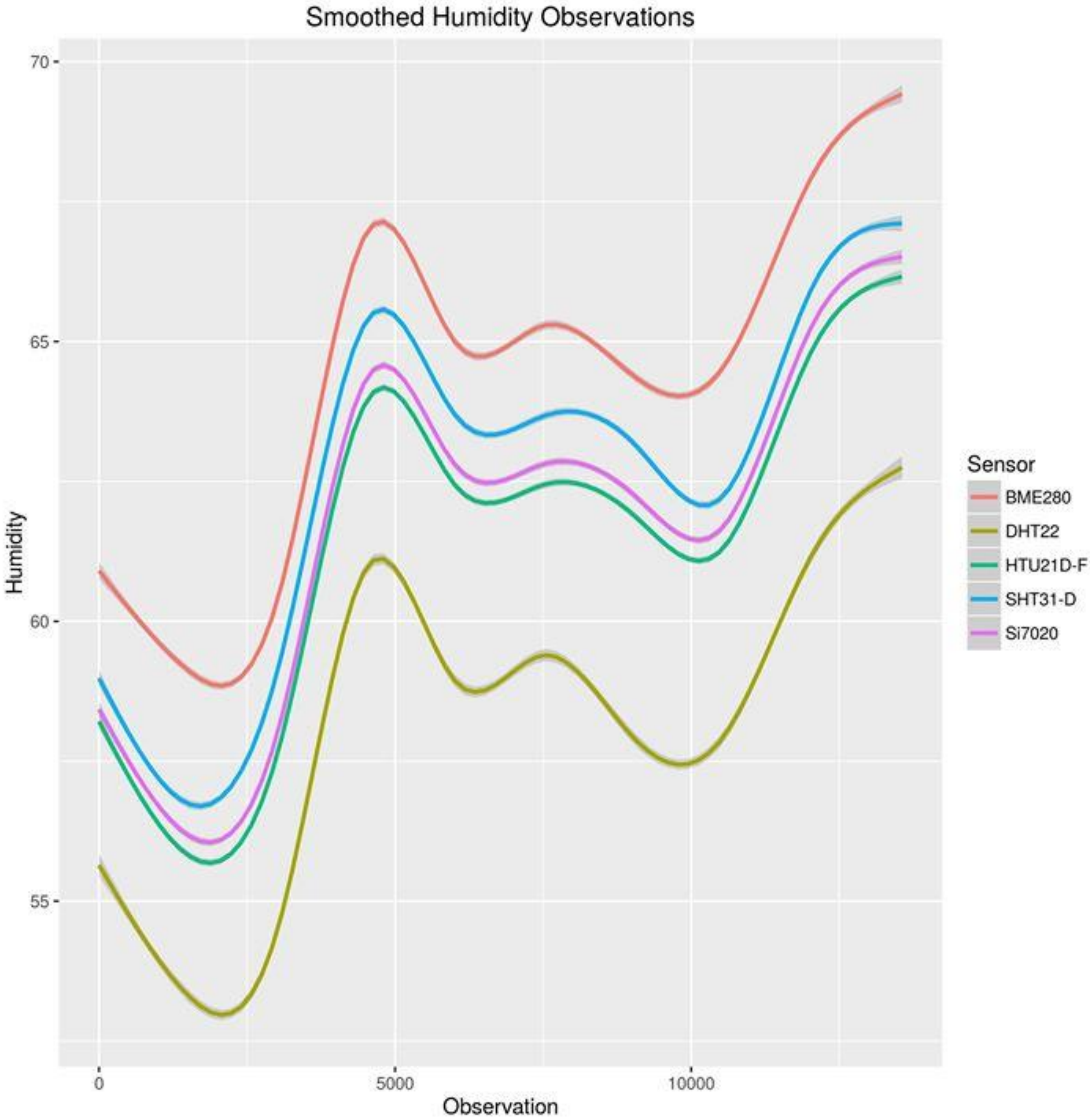
I'm a bit surprised at the results, since the Silicon Labs sensor (the 7021) is, according to the data sheets, the most accurate. If we assume this is the case, then the Si701 line (purple) is close to the "real" temperature and the other sensors have systemic error in them. This is not surprising. What's surprising is that the HTU21D-F and SHT31-D performed close to one another and are the "middle of the road" picks for temperature readings.

Now let's look at the humidity results using the same methods. First, here's the humidity time series plot with outliers (well, all but one) removed:



Humidity observation time series

Again, the fluctuations/variance in the sensor data makes this plot a little difficult to read. Here is the smoothed data time series plot:

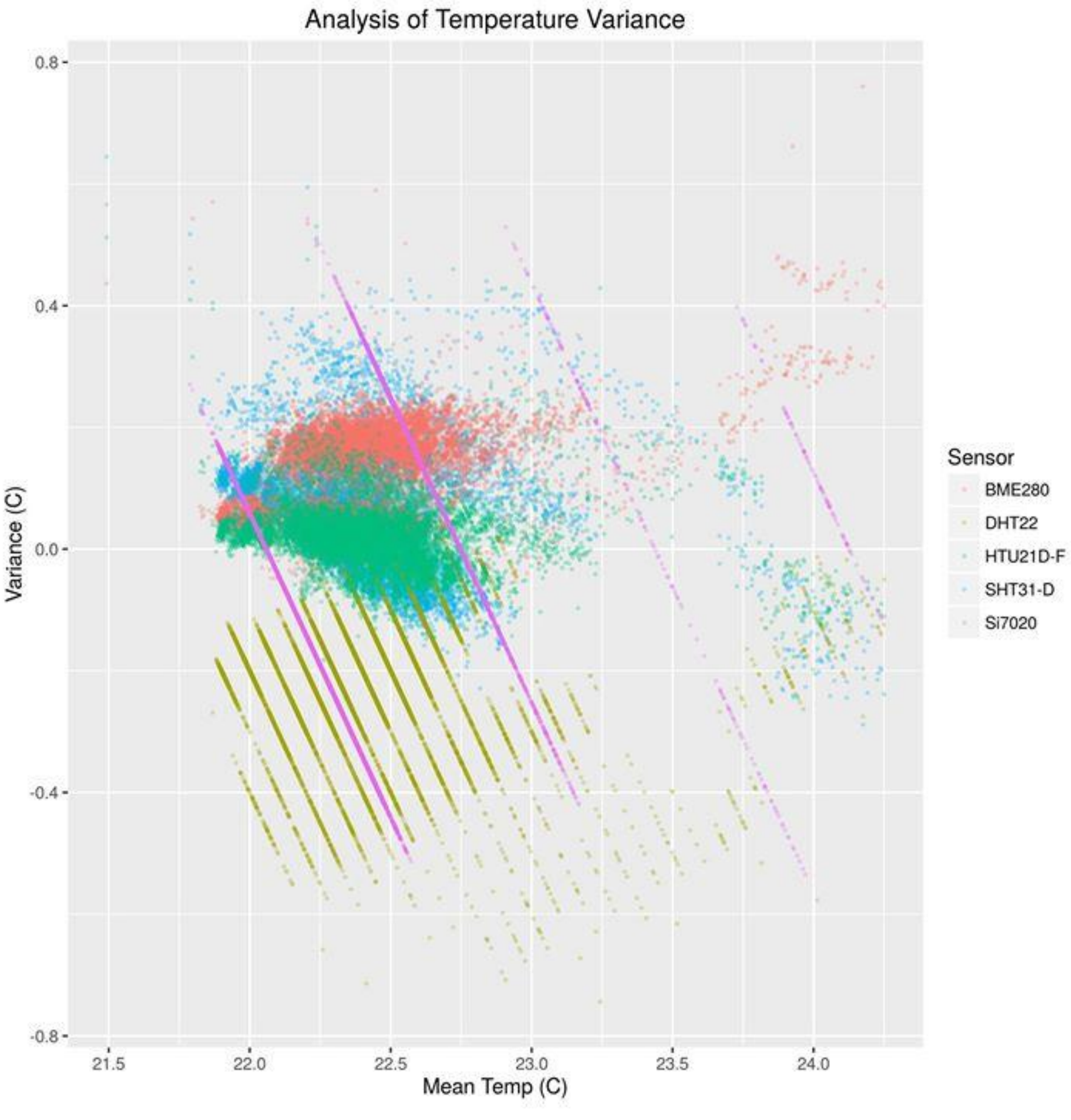


Humidity Smoothed Data Time Series

The relative performance of the sensors, from a humidity perspective, are all in lock-step. The DHT22 is, again, the one with the most suspicious data (consistently low and greater fluctuation than the others). If you need consistently accurate humidity results, I don't think you'll go wrong with any of the HTU21D-F, SHT31-D, or the Si7021. I suspect the BME280 is reading high for both temperature and humidity.

Now let's look at some more interesting visualizations.

These next two plots are extremely interesting.

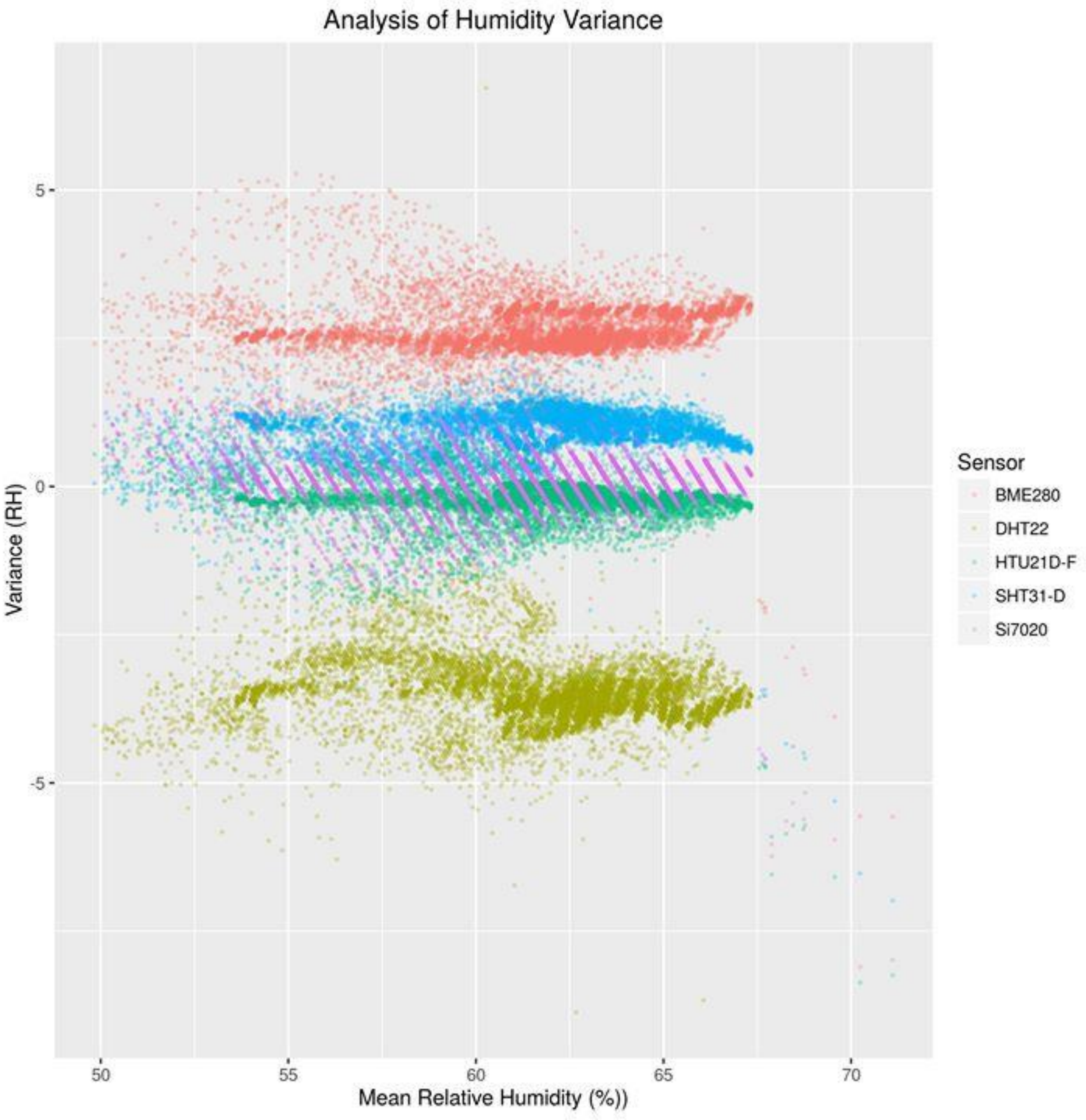


Analysis of Temperature Variance by Temperature

This is the plot of the temperature variance (which is the actual value minus the mean for each observation). The diagonal striping shows the "steps" in the error for each sensor. Ideally, you won't have any steps in the error because steps indicate that the values "jump" more. It's fascinating that the Si7021 (which is, according to the data sheet, an excellent sensor) consistently has readings that are "close to" previous readings, but when they have errors they're pretty wide errors. The next worse sensor is the (no surprise) DHT22.

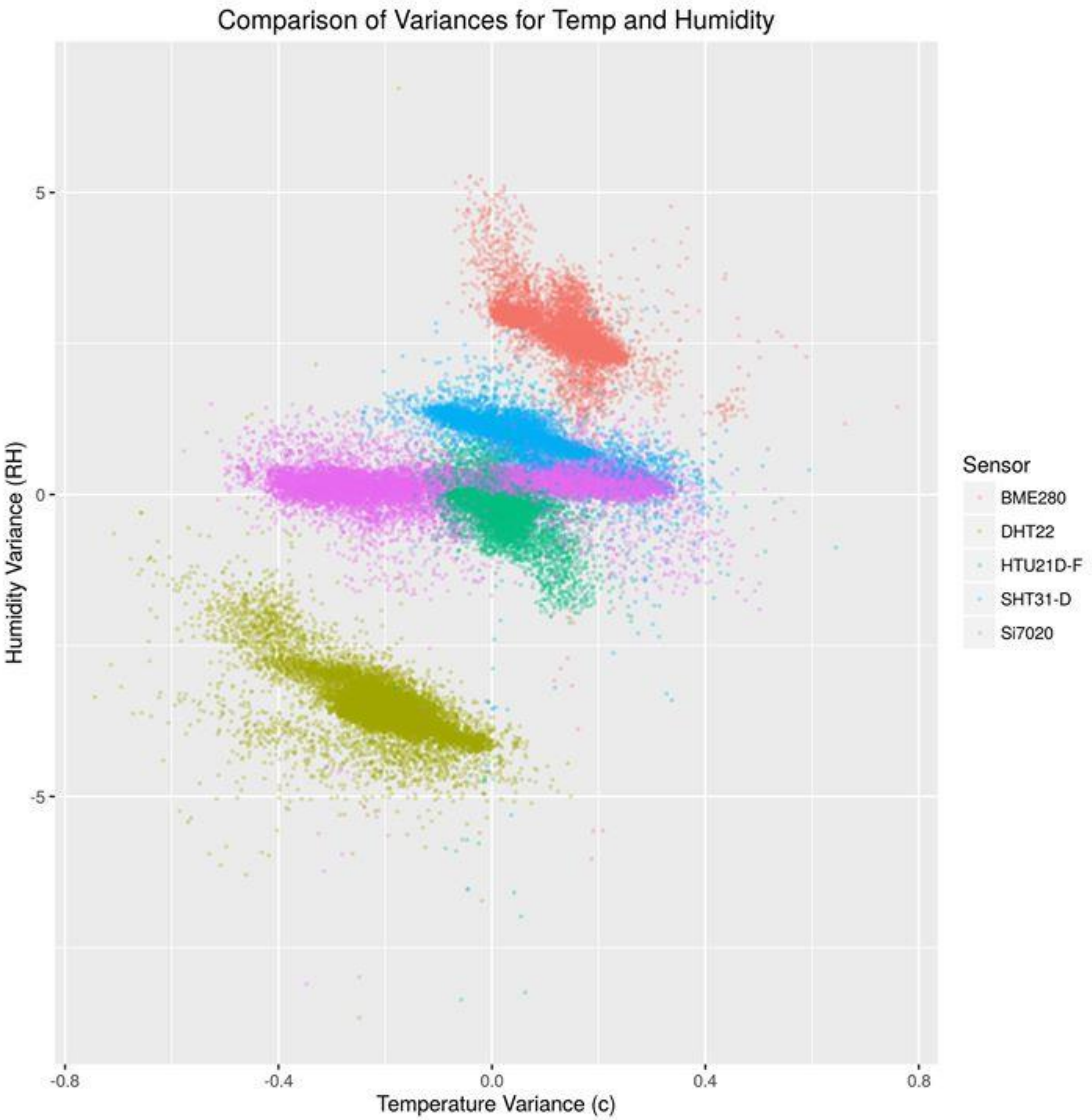
It's a bit hard to read the legend. The red points are the BME280 data, the olive points are the DHT22 data, the blue points are the HTU21D-F data, green points are the SHT31-D data, and, finally, the purple points are the Si7021 data.

This is only half the story. Let's look at the humidity plotted the same way.



Humidity Variance by Mean Humidity

Here we have the same plot mechanism. As before the most jumpy data is from the Si7021 sensor. The DHT22 is better at humidity than temperature readings. The other sensors seem to be pretty good. Again, though, this is only half the story. Let's combine both the temperature and humidity variance graphs:



Remember, red is the BME280, olive is the DHT22, blue is the HTU21D-F, green/teal is the SHT31-D, and purple is the Si7021. What we want to see is a tight grouping of data points. This will indicate a consistent variance in the temperature and humidity performance. In other words, the sensor won't be "way better" at one than the other. The Si7020 has relatively poor temperature variance performance compared to its humidity performance. That's a surprise to me since the 7021 is supposed to be the newest/greatest temp/humidity sensor.

Looking at this graph, and knowing what we've learned from the previous graphs, I would say my preferences for sensors are the HTU21D-F or the SHT31-D. The SHT31-D has a tighter group in this last graph, so if I had to advise people on which temperature/humidity sensor to buy, I would actually recommend the SHT31-D based on the statistics we've seen. This sensor is available in the Adafruit shop for $13.95, which isn't outrageous by any stretch of the imagination.