Grand Rapids Temperature Loggers 2029-2023

The Grand Rapids Soil Survey office established soil temperature probes from 2019 to 2023 at 4 locations representing high and low elevations and northern and southern latitudes within the Manistee National Forest with the goal of confirming a different soil temperature regime (mesic vs. frigid).

I initially attempted several global linear and random forest models to hindcast soil temperature regimes for two different 30-year periods. The global models optionally used case weights to compensate for the higher number of records in the NOAA air temperature data versus soil temperature station. The results compared to actual temperatures were very good. However, extrapolating the results to the earlier timeframes was inconsistent with hindcast estimates based on a more careful approach from individual station differences. It appeared that in the model only one soil temperature site with consistently within the frigid temperature regime (less than 8°C) for the cooler 1961-1990 reference temperature period. However, the manual calculations using the temperatures relative to Cadillac and Big Rapids, the nearest NOAA stations with long term air temperature records, suggests that both northern sites were frigid. However, the high elevation site to the south when compared to the differences from Big Rapids was mesic, unlike what we had anticipated when we established the new MLRA boundaries.

For the manually calculated hindcasts below, the 30-year average air temperature of the nearest NOAA station was compared to the average annual air temperature based on averaging by months for only the days when both soil temperature and air temperature stations were active. The time of overlap is variable depending on station, and the annual temperature averages the monthly temperatures equally irrespective of the number of days per month.

|  |  |  |
| --- | --- | --- |
| time period | station | temperature (°C) |
| 1961-1990 | Cadillac Air | 5.8 |
| cadillac X GRR1 | GRR1 @50 cm | 7.8 |
| cadillac X GRR1 | Cadillac | 7.3 |
| adjusted to 1990 | GRR1 @50 cm | 6.3 |

|  |  |  |
| --- | --- | --- |
| time period | station | temperature (°C) |
| 1961-1990 | Cadillac Air | 5.8 |
| cadillac X GRR2 | GRR2 @50 cm | 9.1 |
| cadillac X GRR2 | Cadillac | 7.1 |
| adjusted to 1990 | GRR2 @50 cm | 7.8 |

|  |  |  |
| --- | --- | --- |
| time period | station | temperature (°C) |
| 1961-1990 | Big Rapids Air | 7.1 |
| bigrapids X GRR5 | GRR5 @50 cm | 9.0 |
| bigrapids X GRR5 | Big Rapids Air | 7.8 |
| adjusted to 1990 | GRR5 @50 cm | 8.3 |

|  |  |  |
| --- | --- | --- |
| time period | station | temperature (°C) |
| 1961-1990 | Big Rapids Air | 7.1 |
| bigrapids X GRR6 | GRR6 @50 cm | 9.3 |
| bigrapids X GRR6 | Big Rapids Air | 8.0 |
| adjusted to 1990 | GRR6 @50 cm | 8.3 |

Random forest model to fill in missing data:

Covariables:

decdate = decimal date (year and fractional time of year).

nt = the mean regional NOAA air temperature average for the time period

ns = soil temperature of the mean regional NOAA air temperature based on linear model of stations that include both air temperature and 50 cm soil temperature, using 45 day running average air temperature and 5 and 22 day running averages of positive-only air temperature as inputs (optimized from several options in number of days to average)

forest = forested status of temperature probe. The soil temperature stations were all considered open except the NRCS stations which were under forest cover.

Soil50 = Whether temperature probe was at 50 cm depth (any other soil temperatures taken at 10 cm). Air = Whether soil versus air temperature.

rf <- ranger(t ~ nt+ns+forest+

air+soil50+sin0+

lat+lon+elev+decdate

, data=alldata.50, num.trees=200, sample.fraction = 0.5)

alldata.50 <- alldata.50 |> mutate(t.rf = predictions(predict(rf, data=alldata.50)))

mean((alldata.50$t - alldata.50$t.rf)^2)^0.5

mean error = 1.171905 m

