SPATIAL INTERPOLATION OF GROWING DEGREE DAYS IN SOUTH-WEST NOVA SCOTIA, CANADA

Dana Melamed
Centre of Geographic Sciences, Advanced Diploma in Geographic Sciences
Contents

Introduction .................................................................................................................................................. 1
Research Questions ...................................................................................................................................... 2
Methodology ................................................................................................................................................. 2
Results and Discussion .................................................................................................................................. 3
Conclusion ..................................................................................................................................................... 6
References .................................................................................................................................................... 7
Introduction

This project builds upon previous efforts by COGS and the Applied Geomatics Research Group (AGRG) to produce spatially interpolated growing degree-day (GDD) maps for South-West Nova Scotia (SWNS). GDD maps are key tools in agricultural research with geographic information systems (GIS). Problems such as site exploration, soil suitability, crop selection, and choice of pest management strategies, are examples of analyses that are best addressed with GDD rasters to represent temperature. The capacity for air temperature to influence plants is often described as the accumulation of heat over time, known as the thermal sum (Scarepare, Scarpare Filho, Rodrigues, Reichardt, & Angelocci, 2011). This concept was first introduced in the 1735, by Reaumur. Growing degree-days (GDDs) have been widely accepted for over 200 years to numerically represent the thermal sum (Equation [2]) (McIntyre, Kliewer, & Lider, 1987).

\[
GDD = \sum_{i=1}^{n}(T^{\circ}_{\text{daily average}} - T^{\circ}_{\text{base}}),
\]

The GDDs to be modelled were derived from the surface air-temperature recorded at 74 weather stations across SWNS over five growing seasons (i.e., April to November for 2012 to 2016). Previous work by COGS/AGRG (Ambrose, 2013) has already determined that generalized additive models (GAMs) (Hastie & Tibshirani, 1990) are an effective method for interpolating the data.

Now, a closer look is being taken into identifying the input layers that most highly correlate with GDD in the GAMs. A combination of Python, R, and ArcMap tools were used for generating maps of the cumulative GDDs over the five growing season. The input landscape layers used include: elevation; slope; aspect; distance to coastline; latitude and longitude; potential solar radiation; actual solar radiation (as derived from mapping cloud conditions with GOES imagery). Results focus on the layers selected for use in the GAMs, the interpolated GDD maps that were derived for each year, and the differences that were observed from one year to the next.

In Nova Scotia, there are approximately 3900 farms covering about one million acres (Statistics Canada, 2011). Nova Scotia is one of four major grape producing provinces in Canada; alongside Ontario, British Colombia, and Quebec (Hope-Ross). South-West Nova Scotia contains three of the four major wine growing regions in Nova Scotia: the Annapolis Valley, the Gaspereau Valley and the South Shore (Bell, 2013). Since temperature is also a major factor in acidic compound production in grapes (Kliewer, 1973), the production of GDD maps are of considerable use to Nova Scotian viticulturists.
Research Questions

1. At which resolution do the input variable rasters significantly model the GDDs?
2. At which time-frame do the input variable rasters significantly model the GDD?
3. Can potential solar radiation rasters be used instead of actual solar radiation data in the model?

Methodology

The project took place in three phases (Figure 2). In the first phase, GDD$_{10}$ rasters were modelled from those accumulated each month in the growing season (A-N), and those accumulated over the entire growing season, in 2012. The outputs were compared to examine the influence of time-frame in the model.

In the second phase, the A-N time-frame in 2012 was modelled again, this time using potential solar radiation (generated by ESRI Area Solar Radiation tool) and again using lacking solar radiation (negative control model).

In both phases, each model was generated from rasters at five different resolutions: 20m, 100m, 250m, 500m, 700m.

All input variables were processed from starting data and/or generated in a Python Script Tool. The values of each input raster were extracted to the weather station’s attribute table. The table was converted to a .CSV and brought into the R (The R Foundation for Statistical Computing, 2013). In R, GAMs were generated, and metrics for comparing the models were collected.

The significance of the correlation for each variable (p-value), adjusted $R^2$ and RMSEs were used to determine relative accuracies of the models. Difference rasters and root mean square errors (RMSEs) from external validation stations, were also calculated to compare the models.

For the third phase, using the information collected from phases one and two, GDD$_{10}$ rasters for the A-N time period for 2014, 2015, and 2016 were generated.
Results and Discussion

Latitude, longitude and elevation consistently make significant contributions to the models, regardless of time-frame (Figures 3 & 4). The proximity to coast does well for A-N, but for the monthly models is insignificant when averaged. The long standard error bars in Figure 3 tell us that the significance changes between months. Aspect and slope only contribute significantly in few select models.

The solar radiation rasters, whether they were remotely sensed (from the Geostationary Operation Environmental Satellite System) or generated with ESRI, did not correlate significantly in the model at the A-N time-frame (Figure 6).

![Figure 3 Mean significance of covariates in monthly models (n=8)](image1)

![Figure 4 Significance of covariates in A-N models](image2)

![Figure 5 Significance of solar radiation for all models (GOES: monthly & A-N, ESRI: A-N)](image3)
The models generated from data at the monthly time-frame, when averaged, had greater adjusted $R^2$ squared value than the Apr. – Nov. time-frame (Figure 6). The values of the Apr. – Nov. time-frame, although lower fall with one standard deviation of the monthly time-frame.

![Adjusted $R^2$ for monthly (average, n=8) vs. A-N modelled with goes solar radiation](image)

Figure 6 Adjusted $R^2$ for monthly (average, n=8) vs. A-N modelled with goes solar radiation

The RMSEs for the monthly time-frame was lower than the Apr. – Nov (Figure 7). There appears to be positive correlation between RMSE and resolution with Pearson coefficients of 0.3155 and 0.8786 for the monthly and Apr. – Nov. time-frame, respectively. This suggests that high resolution models may be suitable for modelling GDD. The RMSEs agree that the monthly time-frame has more accuracy that the Apr. – Nov. time-frame.

![RMSE for Monthly vs. Apr. - Nov. Models made with GOES Solar Radiation](image)

Figure 7 RMSE scores for monthly (square) and Apr.- Nov. (diamond)modelled with GOES solar radiation

\[
y = 0.007x + 103.51 \\
R^2 = 0.8786 \ (GOESan)
\]

\[
y = 0.0053x + 102.14 \\
R^2 = 0.3155 \ (GOESmon)
\]
Figure 3 The five-year average of GDD\textsubscript{10} in SWNS 2012 - 2016

Figure 8 RMSE scores for monthly (square) and Apr.-Nov. (diamond) modelled with GOES solar radiation

\begin{equation}
y = 0.0157x + 102.41 \\
R^2 = 0.9228 \text{ (ESRI)}
\end{equation}

\begin{equation}
y = 0.007x + 103.51 \\
R^2 = 0.8786 \text{ (GOESan)}
\end{equation}

\begin{equation}
y = -0.001x + 104.09 \\
R^2 = 0.0357 \text{ (NoSolAn)}
\end{equation}
The GOES and ESRI-generated solar radiation models at the Apr. – Nov. time-frame had higher RMSEs (Figure 8) than the negative control. At this time-frame, the solar radiation is not a suitable variable, for it detracts from the accuracy.

The final GDD rasters for 2012 – 2016 and the five-year average (2012-2016) were generated using Apr. – Nov. models, without solar radiation, at the 20m resolution. This time-frame was selected on its majorly lessened processing effort, compared to the monthly modelling, and was comparable in accuracy (104 vs. 99 RMSE).

Conclusion

The April – November time-frame may be too long to model GDD. The monthly models scored higher adjusted $R^2$ values and RMSEs. The comparison between GOES solar and ESRI-generated solar was unable to be made, as they were rarely significant at this time-frame (and not in comparable models the unique time they each were). Next steps include generating ESRI solar radiation data at smaller time-frames (potentially daily) and making a comparison with daily GOES solar radiation.
References


