Water Level Variability and Trends

Analysis of groundwater levels in south-central Ontario

Oak Ridges Moraine Groundwater Program

26 August, 2023

Table of Contents

[Introduction 2](#_Toc143939991)

[The ORMGP 2](#_Toc143939992)

[Longer-term levels 3](#_Toc143939993)

[Methodology 3](#_Toc143939994)

[Pattern + Noise 3](#_Toc143939995)

[Generalized Additive Model 4](#_Toc143939996)

[Processing steps 5](#_Toc143939997)

[Maps 6](#_Toc143939998)

[Water Level Range 6](#_Toc143939999)

[Watertable 6](#_Toc143940000)

[Deeper system levels 7](#_Toc143940001)

[Distributions 8](#_Toc143940002)

[Long-term Trend 9](#_Toc143940003)

[Shallow (<20m) wells 10](#_Toc143940004)

[Deep (>20m) wells 11](#_Toc143940005)

[Distributions 12](#_Toc143940006)

[References 12](#_Toc143940007)

[Extras 12](#_Toc143940008)

[Presentations 12](#_Toc143940009)

[Complete GAM formulation 13](#_Toc143940010)

[Example in ggplot 13](#_Toc143940011)

[see also 16](#_Toc143940012)

# Introduction

Groundwater levels are fundamental to the practice of hydrogeology. They are necessary for flow system analysis and understanding from determining hydraulic gradients to estimating material properties to calibrating regional groundwater flow models.

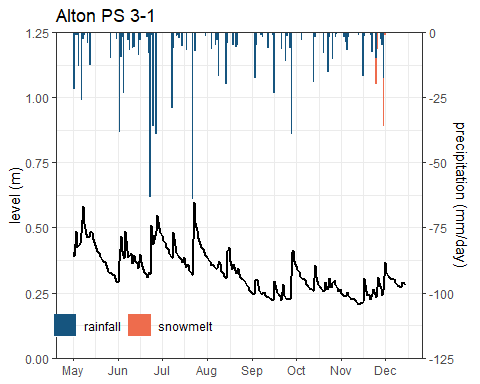
From a hydrological perspective, groundwater levels and their seasonal fluctuation affect runoff generation in southern Ontario. Longer-term (5-10yr) fluctuations should also impact on land use development; think of constructing and maintaining buried infrastructure (e.g., sewers, water mains), or applying certain Low Impact Design (LID) strategies, etc.

For a number of good reasons, groundwater resource management tends to boil down to a set of steady-state (static) targets. This, immediately implies that the steady-state condition represents some long-term average state, without determining, say: *How much should I expect the water table to vary from its current position?*

### The ORMGP

The [Oak Ridges Moraine Groundwater Program](https://www.oakridgeswater.ca/) is a municipal agency charged with synthesizing and disseminating water resources data, geologic interpretations, and numerical models over a 30,000 km² area situated in south-central Ontario. One of our interpretive mapping products gaining increased use are estimations of the depth to the water table and deeper potentiometric surfaces. These maps are constructed from thousands of data points collected over many years and at various times of the year considered to represent an “average” state. An effective use of these maps must consider i) longer-term deviation the measured and average conditions, ii) short-term seasonal patterns, and iii) water table sensitivity to storm and/or snow melt events.

From our database, 690 good-quality monitoring wells exist within the study area, collecting regular (e.g., daily, hourly or less) groundwater level data over decades. At some locations, when evaluated alongside climate data (e.g., precipitation and snow melt estimates) the water level response to events or stresses can be analysed.



*A 4.2m deep well located 120m away from Shaw’s Creek, Alton, ON.*

### Longer-term levels

Long-term seasonal variability in shallow groundwater levels have been quantified using a 12-point Generalized Additive Model (GAM) that provides confidence intervals to a fitted seasonal trend.

From 273 shallow wells (depths < 20m), the southern Ontario water table fluctuates ±0.7 (s.d. 0.6) m. Deeper (n=416) wells show greater fluctuation of groundwater levels at ±1.4 (s.d. 1.6) m.

# Methodology

## Pattern + Noise

The intention here is to determine the expected range of variability one should expect when (say) reading our interpolated water table or potentiometric surface maps. Our [***water table map***](https://owrc.github.io/watertable/) should be thought of as a *long-term average* with uncertainty associated with *when* the measurements were made. It is a surface interpolated from 100,000s measurements made over the past century, and constrained by mapped watercourses, water bodies, etc.

Realistically, the depth to water table and potentiometric surfaces can vary naturally by a metre or more, and even more if affected by local groundwater pumping centres. Long-term trends in water levels and aquifer storage further impacts expected water levels, as this only changes the stage from which the expected variability is to occur.

For now, lets skip the event-base groundwater response and break the expected variability into 2 components:

Further more, lets assert:

* seasonal variability is assumed to have a cyclic nature; and
* inter-annual variability (or long-term trend) can be used to de-trend the time-series.

## Generalized Additive Model

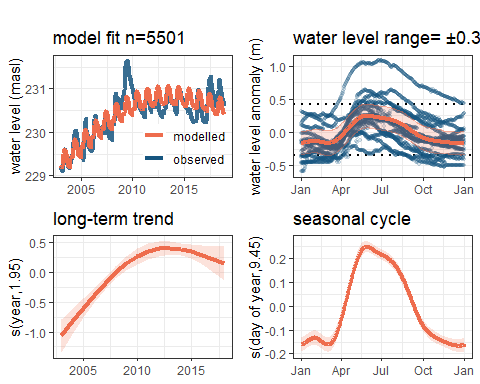
The Generalized Additive Model (GAM) is a flexible means of fitting a model to time series data. It is a generalization of the linear models we have all fit to a scatter plot, yet retains the ability to efficiently compute confidence/prediction intervals using an intuitive control (Wood, 2017).

GAMs are especially useful as hydrographs contain a good deal of noisy auto-correlated data (that is, a sequence of measurements made in a short time frame tend to be correlated). For instance, it’s apparent that there is a seasonal pattern to most hydrological data in southern Ontario. Here we design the GAM with 12 smoothing spline “knots”, one for every calendar month. These monthly knots are further specified as being a cyclic regression spline by assuring that the surface remains continuous to the second derivative at a years end (Wood, 2017). The GAM will fit a smoothing spline that captures the underlying seasonal pattern.

The same goes for the long-term trend, which is determined using knots specified to every year. The GAM is then specified as:

where is the expected/modelled water level, the “smooth function” is the long-term trend and is the seasonal variability of the GAM, and is the AR(1) autocorrelation function. The GAM is initially set to having 365 + degrees of freedom, is the number of years.

What’s important is that the above equation is linear, and so the GAM can be decomposed into long-term and seasonal patterns fitted to the the data:



The bottom 2 plots illustrate the first 2 terms in the *RHS* of the above equation, with a 90% prediction intervals shaded. Top 2 plots show the GAM vs. observed both over the period of record (left) and after being de-trended (right).

The *y*-axis label of the bottom 2 plots report numbers called the “effective degrees of freedom,” which relates to the minimum number of constraints needed to reproduce this GAM, simplifying the regression from the initial constraints given (i.e., .)

### Processing steps

Of the 689 locations with >34 water level **logger** measurements are queried from our database. Of these, data were screened for having a reported well screen depth and having at least 4 years of data only 690 wells remained. In addition, for every location time-series:

1. water level data were re-sampled to daily averages
2. outliers were automatically removed using the ±1.5 IQR (inter-quartile range) method
3. values are factored according to their day-of-year and year of the reported level
4. data are de-trended by subtracting from the data
5. 90% prediction intervals surrounding are taken as the inter-annual (long-term) variability that is added to the smooth function
6. annual maximum and minimum extent of the prediction intervals defines *the anticipated range in groundwater levels one should expect at a given location in the near future*.

This gives the upper-right plot in the above figure.

* [Download GAM plots prduced for all 690 locations here](https://www.dropbox.com/scl/fi/uhl611zqhmqcb0ea925s5/gwGAM-summary.pdf?rlkey=71shw1gnl0tckjblwudw3atiy&dl=1).
* [Tabular form here](https://www.dropbox.com/scl/fi/973q2ts8qm7yo7a6dbjfc/gwGAM-summary.R.csv?rlkey=7uw31hllk1yixo7gh0dxksmzo&dl=1).

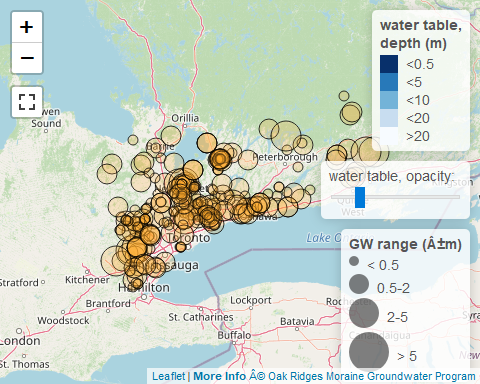
# Maps

## Water Level Range

hover over any circle to reveal station properties, click for more details. Full-screen available in the top-left corner

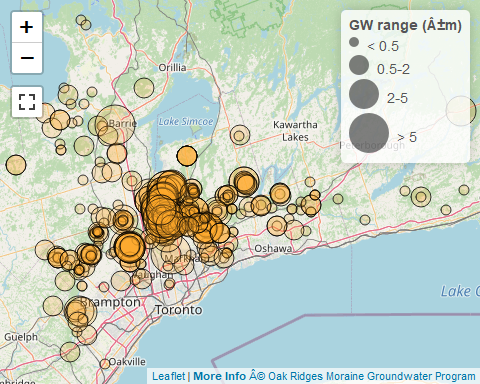
### Watertable

Showing only shallow (<20m) wells.

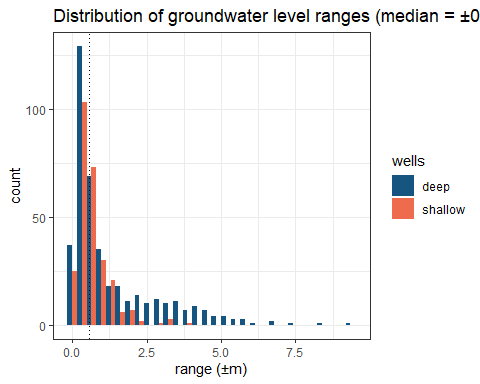


### Deeper system levels

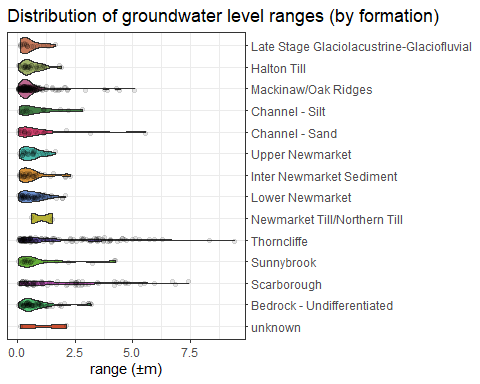
Showing only shallow (<20m) wells.



### Distributions



#### By formation



## Long-term Trend

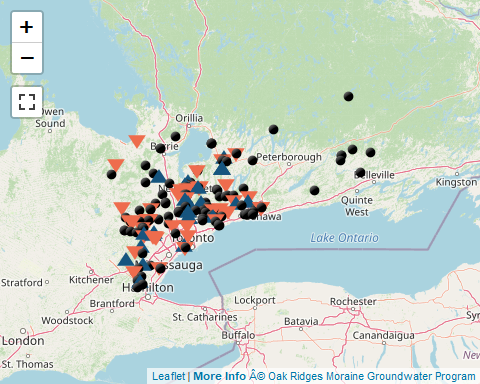
Groundwater level monitoring over the past decade have been put to the Mann-Kendall test for trend. The maps below identify wells that have exhibited increasing vs. decreasing change over the past 10 years.

Mann-Kendall test for trend.

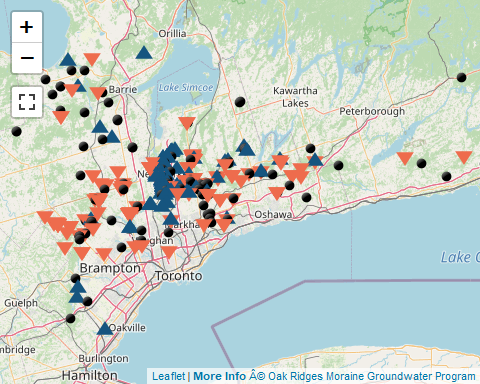
* Blue Triangles: increasing trend (p<0.05)
* Orange Triangles: decreasing trend (p<0.05)
* Circles: no trend

hovering on any triangle to reveal station properties, click for more details. Full-screen available in the top-left corner

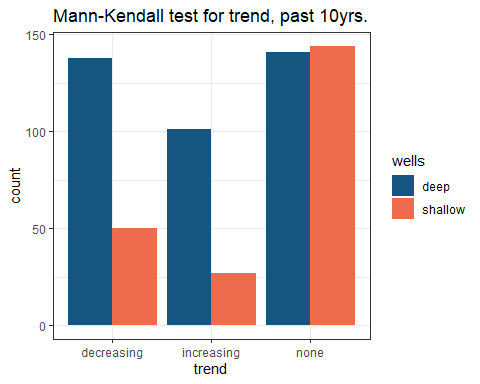
### Shallow (<20m) wells



### Deep (>20m) wells



### Distributions



# References

Wood, S.N., 2017. Generalized Additive Models: An introduction with R, 2nd ed. CRC Press. 476pp.

# Extras

## Presentations

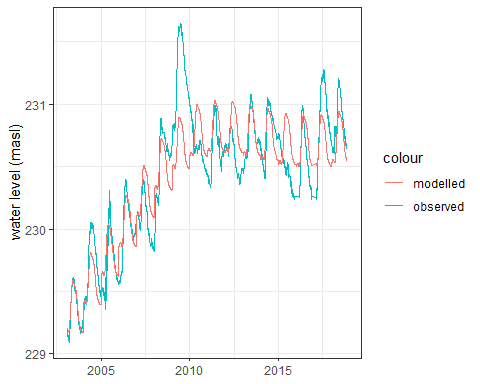
[Marchildon et.al. CWRA 2023 National Conference poster](https://golang.oakridgeswater.ca/pages/2023/Marchildon_et_al-CWRA%202023%20National%20Conference-poster.pdf)

## Complete GAM formulation

where is the row of a model matrix , is the parameter vector. can be leveraged to handle pumping influences; is the cyclic basis function for month at time multiplied by some parameter ; and is the error term.

## Example in ggplot

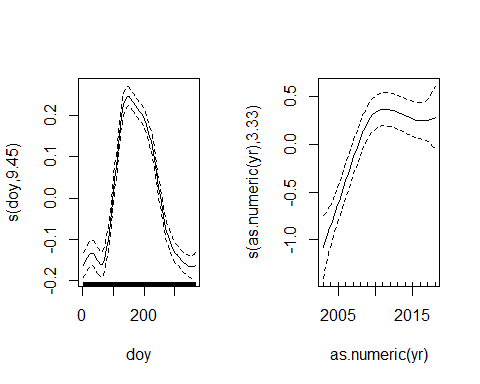
cdf <- gamm(data=exampldf, Val~s(doy,bs="cc",k=12)+s(as.numeric(yr),bs="cr"), correlation=corAR1(form=~1|yr))  
exampldf %>% ggplot(aes(date,Val)) +  
 theme\_bw() +  
 geom\_line(aes(colour='observed')) +  
 geom\_line(data=data.frame(Val = cdf$gam$fitted.values, date = exampldf$date), aes(colour='modelled')) +  
 labs(x=NULL, y="water level (masl)")



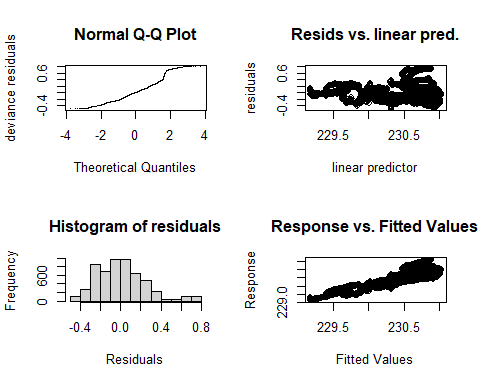
print(summary(cdf$gam))

##   
## Family: gaussian   
## Link function: identity   
##   
## Formula:  
## Val ~ s(doy, bs = "cc", k = 12) + s(as.numeric(yr), bs = "cr")  
##   
## Parametric coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 230.41501 0.05817 3961 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df F p-value   
## s(doy) 9.446 10.00 101.6 <2e-16 \*\*\*  
## s(as.numeric(yr)) 3.330 3.33 19.6 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.791   
## Scale est. = 0.062726 n = 5501

layout(matrix(1:2, nrow = 1))  
plot(cdf$gam, scale=0)



gam.check(cdf$gam)



##   
## 'gamm' based fit - care required with interpretation.  
## Checks based on working residuals may be misleading.  
## Basis dimension (k) checking results. Low p-value (k-index<1) may  
## indicate that k is too low, especially if edf is close to k'.  
##   
## k' edf k-index p-value   
## s(doy) 10.00 9.45 1.04 0.99   
## s(as.numeric(yr)) 9.00 3.33 0.33 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

intervals(cdf$lme,which = "var-cov")

## Approximate 95% confidence intervals  
##   
## Random Effects:  
## Level: g   
## lower est. upper  
## sd(Xr - 1) 0.0003575794 0.0008996196 0.002287185  
## Level: g.0   
## lower est. upper  
## sd(Xr.0 - 1) 0.0001811636 0.001539592 0.01378532  
##   
## Correlation structure:  
## lower est. upper  
## Phi 0.9976847 0.9988042 0.9993826  
## attr(,"label")  
## [1] "Correlation structure:"  
##   
## Within-group standard error:  
## lower est. upper   
## 0.1800430 0.2504520 0.3483956

cdf$gam$sig2/cdf$gam$sp

## s(doy) s(as.numeric(yr))   
## 0.0008996196 0.0015395917

*simultaneous interval for a penalized spline in a fitted Generalized additive model (GAM)*

## see also

[**Groundwater and Flooding**](gwflooding.html): investigating the causal nature of groundwater flooding through simulation.