Modelling the relationship between groundwater depth and NDVI using time series regression with Distributed Lag Models

Margaret Phiri [1*], Yegnanew A. Shiferaw [2] and Solomon G. Tesfamichael [1]

[1] Department of Geography, Environmental Management and Energy Studies, University of Johannesburg, Auckland Park Kingsway Campus, PO Box 524, Johannesburg 2006, South Africa; [2] Department of Statistics, University of Johannesburg, Auckland Park Kingsway Campus, PO Box 524, Johannesburg 2006, South Africa

Abstract

Groundwater plays a key role in hydrological processes, including in determining aboveground vegetal growth characteristics and species distribution. This study aimed at estimating time-series data of Normalized Difference Vegetation Index (NDVI) using groundwater depth as a predictor in two land cover types: grassland and shrubland. The study also investigated the significance of past (lagged) groundwater and NDVI in estimating the current NDVI. Results showed that lagged groundwater depth and vegetation conditions influence the amount of current NDVI. It was also observed that first lags of groundwater depth and NDVI were significant predictors of NDVI in grassland. In addition, first and second lags of NDVI were consistently significant predictors of NDVI in shrubland. This shows the importance of vegetation type when modelling the relationship between groundwater depth and NDVI.

Keywords: Groundwater depth; Landsat NDVI; Time-series analysis; Distributed Lag Models

1. Introduction

Groundwater plays an important role in the growth of aboveground vegetation (Fan *et al*., 2016; Rodriguez *et al*., 2016; Werner *et al*., 2017). This is particularly the case in arid areas where rainfall and surface water levels are low or unpredictable (Wada *et al*., 2010; Jin *et al*., 2011; Orellana *et al*., 2012; Liu *et al*., 2013). Knowledge of groundwater dynamics can therefore be linked to different vegetation characteristics such as vegetation type (Cooper *et al*., 2006), change in species composition (Chen *et al*., 2006), evapotranspiration (Yin *et al*., 2015) and vegetation decomposition (Wiedermann *et al*., 2017). Furthermore, information about vegetation dynamics can inform underground water fluctuations (Le Maitre *et al*., 1999) and recharge zones or amounts (Yin *et al*., 2015; Yeh *et al*., 2016). It is therefore logical to quantify the relationship between groundwater and aboveground vegetation, although this relationship can be complicated by other factors. Jeong *et al*. (2018), for example, showed the effect of precipitation on groundwater fluctuations. In a study to

147

investigate the effects of water table fluctuations on specific yield, Chinnasamy *et al*. (2018) acknowledged the difficulty of groundwater monitoring due to geologic and related porosity properties, while earlier, Leung *et al*. (2011) showed the combined effects of soil climate and vegetation on spatio-temporal dynamics of groundwater. Ghose *et al*. (2018) showed the importance of evapotranspiration and runoff in influencing groundwater level in an arid environment. Another important factor that complicates the relationship between groundwater and vegetation amount is land use type. For example, Huang *et al*. (2018) observed that older orchard farms depleted groundwater more than younger orchards and farmlands.

Field-based assessment of vegetation-groundwater relationships involves measurements using dug wells and boreholes; this approach requires considerable logistical investment, particularly for continuous monitoring purposes. Remote sensing on the other hand offers a cost- and time efficient alternative, and has been used for groundwater inventories and potential siting efforts. This has been implemented by integrating image interpretation and landscape physical characteristics such as topography, drainage patterns and lineaments (e.g., Elmahdy and Mohamed, 2014; Sakala *et al*., 2014; Shekhar and Pandey, 2015). Another common approach for characterising groundwater is through interpretation of aboveground vegetation patterns. Such patterns can be described using remote sensing techniques that are capable of producing explicit spatio-temporal maps that reflect changes in vegetation characteristics (Barron *et al*., 2014; Gou *et al*., 2015). Zhu *et al*. (2015), for example, quantified the effect of groundwater depth on vegetation growth using the normalized difference vegetation index (NDVI) derived from MODIS and AVHRR images in arid environments. Their study applied Pearson's correlation analysis, and found significant relationships between the two during drought periods. One of the contributions of the study was to investigate the relationships using time-series observations at different locations. However, the study focused on drought periods and it is unknown if such a relationship applies to other seasons too. Fu and Burgher (2015) related tree patch dynamics represented by Landsat NDVI and environmental/hydrological data (rainfall, temperature, groundwater depth, surface flow variables) using regression tree analysis in riparian vegetation. The analysis showed that groundwater depth had a major influence on NDVI values, with areas having shallower groundwater depth corresponding to higher NDVI values. Using NDVI derived from MODIS data, Jin *et al* (2016) showed the dependence of short shrubs on groundwater depth, while deep-rooted shrubs that accessed water from deeper water tables were affected less. The number of samples used by Jin *et al*. (2016) was quite high; however the observations were based on spatial distributions rather than time-series data. Similar applications have been implemented in numerous studies to map groundwater-dependent ecosystems worldwide (e.g. Barron *et al*., 2014; Gou *et al*., 2015; Gow *et al*., 2016). Although spatial analysis is one of the core advantages of remote sensing, it is crucial to build fundamental relationships using time-series data that eliminate other sources of variation due to spatial characteristics. In addition, the study by Jin *et al*. (2016) did not attempt to build quantitative relationships to estimate groundwater dependence of plants based on NDVI.

In this study, we aimed to relate groundwater depth and NDVI data using time-series regression. Specifically the study investigated the performance of distributed lag models by testing the effects of different past (lagged) values of independent (groundwater depth) and dependent (NDVI) variables. The investigation was implemented using time-series data at two sites (one on grassland and the other on shrubland).

2. Data

The study was undertaken in the Gauteng Province of South Africa (Figure 1). Selection of sampling locations was determined by the availability of sufficient number of matching observations of groundwater depth and cloud-free Landsat data. Two groundwater monitoring stations, one representing grassland and the other shrubland cover type, were sampled for the study (Figure 1). Identification of these land cover types was informed by the South African National Land-Cover map produced from Landsat imagery acquired in 2013 and 2014 (DEA, 2015). Few woody plants were observed in the grassland while the shrubland had considerable amount of wood, although low shrubland according to the National Land-Cover map has an approximately 15% woody canopy cover. A forest land was considered as a third land cover type; however it was eliminated due to insufficient time-series data. The selected stations (boreholes) are used solely for groundwater monitoring purposes, and no other uses such as consumption for domestic and irrigation purposes are allowed from them. Thus, anthropogenic interventions that would have influenced aboveground vegetation characteristics (growth patterns and plant types) and the NDVI-groundwater relationships were avoided. We did confirm the land cover types through field surveys. Instead, we used visual observations of historical images of Google Earth to ascertain that the selected land cover types did not change over the years, except for the amount of cover for each. Google Earth is a cost-effective verification method that can be used to verify other remote sensing based classifications (e.g., DEA, 2015). One important challenge in the selection of these stations was the difficulty of categorizing a land cover strictly as shrubland; as a result, the cover also included certain amount of grassland. Such confusion between the two classes was also noted by DEA (2015), which attributed the uncertainty to classification and interpretation errors. However, it was assumed for the purposes of the study that dominance of larger plants suppressed the contribution of grasses to the overall spectral properties captured by the remotely-sensed data.
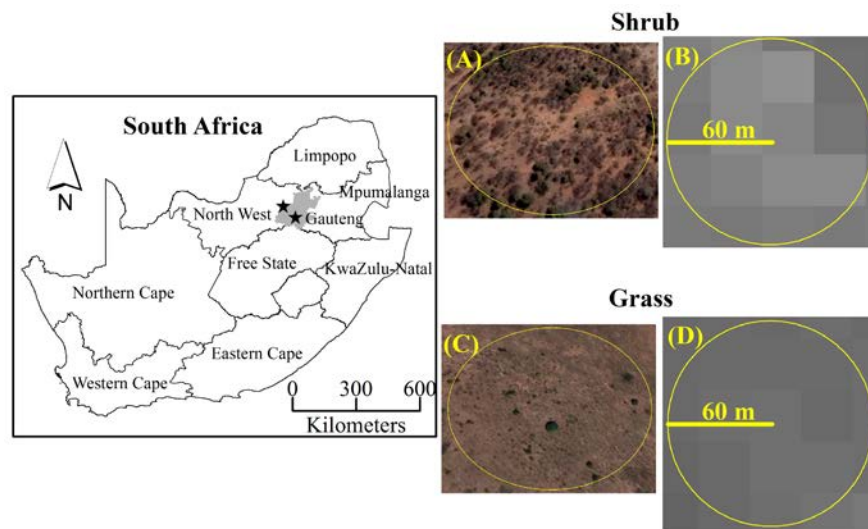
Figure 1. Map of borehole location used in the study and land cover types. (A) and (C) show Google Earth images of the two land cover types; (B) and (D) show Landsat NDVI of respective land cover types of one time.

Data showing water table depths were obtained for each station from the National Groundwater Archive of the Department of Water Affairs and Sanitation, South Africa. Although observations at different locations in the country in some areas are made bi-weekly, the data provided for this study was less regular. The frequency of the data for this study was also dictated by the availability of matching Landsat data within the same month. We limited the data between August 2012 and November 2016 for grassland and December 2012 and February 2017 for shrubland, due to large number of observations within these periods. A monthly-scale analysis was sought for the study; however the original data did not have such an ideal and matching frequency for Landsat and groundwater data. There were a maximum of 5 and 4 months of missing Landsat data and 3 and 5 missing groundwater per year for grassland and shrubland, respectively; however no two missing data were consecutive. We used imputation by mean value to estimate the missing value to generate a roughly monthly time-series data (Moritz, 2016). The total number of observations for grassland and shrubland were 52 and 51, respectively.

The NDVI (Equation 1) for both grassland and shrubland was quantified from Landsat imagery for the same months as groundwater observations. Both the red and Near Infrared in the equation had 30 m spatial resolution on the ground for all the Landsat sensors (Thematic Mapper (TM), Enhanced TM (ETM+) and Operational Land Imager (OLI)) used in the study. In all cases, cloud-free data were downloaded making atmospheric correction unnecessary. Prior to computing NDVI, radiometric calibration was applied to convert from pixel digital number to top-of-atmosphere reflectance; showed the amount of reflectance relative to the sun's radiation incident on the surface of the Earth and was achieved by using re-calibration coefficients or constants supplied in the metadata of each Landsat image (Chander *et al*., 2009).

$$NDVI = \frac{(\text{Near Infrared} - \text{Red})}{(\text{Near Infrared} + \text{Red})} \qquad [1]$$

For each date, mean NDVI was computed for pixels falling within a 60 m radius of each station. Use of an aggregated value covering a larger spatial area was preferred to focusing on the pixel at the borehole. This was done to increase the reliability of the information, instead of basing interpretation on a single pixel that could be incorrectly placed away from the borehole. In addition, considering a neighbourhood of pixels agrees with our classification of land cover type that sought to assign land cover name based on dominance over a relatively large spatial area. This avoided having two land cover types in close proximity to the borehole. Figure 2 shows the trends of groundwater depth and rainfall per land cover type.

3. Methods

There are different statistical models available in the literature to model the relationship between variables. Among these, the autoregressive distributed lag (ARDL) model has been used for decades to model the relationship between variables in a single time series setup. In the ARDL model, the lag coefficients forced to model the behaviour of a response variable *y* under the impact of an explanatory variable *x*. The ARDL model is the major workhorse in dynamic single equation regressions (Uwe and Jürgen, 2005).

3.1 Linear regression models

The linear regression model is a widely used modelling approach throughout the physical, natural and social sciences. Due to its relative simplicity and well-known properties, it is one of the fundamental (supervised) machine-learning algorithms. Modelling begins by fitting linear regression (Equation 2) in the time series context by assuming NDVI as a dependent variable y_t and depth of groundwater as an independent variable x_t. This relation is expressed through the linear regression model

$$y_t = \beta_0 + \beta_1 x_t + u_t, \ t = 1,\ldots,n \qquad [2]$$

where β_0 and β_1 are unknown fixed regression coefficients, and u_t is a random error term.

3.2 Finite autoregressive distributed lag model

The *ARDL* model is a flexible infinite distributed lag model (Equation 3). The general form of a ARDL (p,q) is:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \cdots + \beta_p x_{t-p} + \gamma_1 y_{t-1} + \cdots + \gamma_q y_{t-1} + u_t \qquad [3]$$

In this study, we used up to three past (lagged) times of both groundwater and NDVI as important factors influencing a relationship at a given time. As such, NDVI was estimated based on the immediate lag of groundwater and NDVI (ARDL (1,1)); the two immediate lags of groundwater and NDVI (ARDL (2,2)) as well as the three immediate lags of groundwater and NDVI (ARDL (3,3)).

3.3 Distributed lag models with Koyck transformation

While the basic Distributed Lag Models (DLMs) handle finite cases, the Koyck transformation is used to deal with infinite DLMs. The Koyck DLM (Equation 4) specifically uses as inputs the independent variable at a finite time and the immediate lagged value of the dependant variable. The model is given as follows:

$$y_t = \delta_1 + \delta_1 y_{t-1} + \delta_3 x_t + u_t.$$ [4]

where u_t is a stationary error term with $E(u_t) = 0$, $Var(u_t) = \sigma^2$, $Cov(u_t, u_s) = 0$, $for\ t \neq s$.

3.4 Fitted model accuracy assessment

Mean absolute scaled error (MASE) is a well-established accuracy assessment technique to measure the resultant fitted models in time-series regression analysis (Franses, 2016) and can be quantified for all fitted models used in this study (Equations 5 & 6). MASE is computed as

$$MASE = mean(|q_t|).$$ [5]

where q_t is a scaled error defined as

$$q_t = \frac{y_t - \hat{y}_t}{\frac{1}{n-1}\sum_{i=2}^{n}|y_t - y_{t-1}|},$$ [6]

where y_t and $\hat{y}_t$ represent observed and predicted NDVI, respectively. If MASE < 1, the fitted model gives smaller errors and thus is deemed acceptable (Hyndman and Koehler, 2006).

Furthermore R-square, adjusted R-square, residual analysis, log-likelihood as well as information criteria including Akaike information criteria (AIC) and Bayesian information criteria (BIC) were used to compare the models. Analysis of residuals in particular was used to determine if there was any departure from the assumptions of the random errors; in a way, this analysis confirms if the residuals are independent, and variance of the residuals is constant. The residual is defined as the differences between the observed and fitted NDVI value. The resultant residuals were then plotted against the corresponding fitted values to check the characteristics of the variance. Akaike (1974) proposed the well-known AIC in order to evaluate models given as follows

$$AIC = 2k - 2log\mathcal{L},$$ [7]

where k is the number of parameters in the model and $\mathcal{L}$ is the value of the likelihood function. Similarly, Bayesian information criterion (BIC) has a similar form

$$BIC = klog(T) - 2log\mathcal{L},$$ [8]

where *T* is the sample size. In addition to AIC and BIC, Adjusted R-square also used to choose the best fitted model.

4. Results and discussions

Table 1 presents the results obtained using the simple linear regression model, Koyck's model, ARDL(1,1), ARDL(2,2) and ARDL(3,3) models. The linear regression, which simply used groundwater depth as the sole predictor, was included to show the relative performances of the other approaches that took into consideration lagged data of groundwater depth or NDVI. The result for grassland shows that groundwater depth was not a significant estimator of NDVI using the linear regression model. The Koyck model returned a better result particularly by identifying the first lag of NDVI as a significant predictor, although groundwater depth was not significant. In all the ARDL models, the first lags of both groundwater depth and NDVI were significant predictors. In general, the model fit statistics including (adjusted) R-squared, the significance level of F-statistic (p values), AIC, BIC, log-likelihood and MASE MASE confirmed the superiority of ARD models.

As an example, let $NDVI_t$ denote the NDVI at time t, and $Depth_t$ denote the depth of the ground water at time t. Based on the adjusted R squared, AIC, BIC and log likelihood function, the ARDL1 model fits the data very well at the grassland location. Then, the fitted ARDL1 model in the grassland location takes the form (with values in parentheses representing standard errors of the coefficient estimates):

$$\widehat{NDVI}_t = \underset{(0.582)}{0.286} + \underset{(0.009)}{0.018} \; Depth_t + \underset{(0.136)}{0.690} \; NDVI_{t-1} - \underset{(0.007)}{0.021} \; Depth_{t-1} - \underset{(0.142)}{0.06} \; NDVI_{t-2}.$$

The coefficient on lagged NDVI is positive and large. That is, for a unit increase in the first lag of NDVI, NDVI will increase by approximately 0.690. Conversely, the coefficient on lagged depth is negative, indicating that NDVI decreases by 0.021 as the water level of immediate past drops by a unit.

None of the higher order lags (second and third) of the independent (groundwater depth) and dependent (NDVI) were significant in predicting NDVI in grassland. This was not unexpected for the grassland cover type. Grasses, with shallow root systems, are known to be responsive to soil moisture fairly in a short-period of time, compared to deep-rooted woody plants (Koide and Koike, 2012). Chen *et al*. (2015) also noted the sensitivity of herbaceous plant characteristics to groundwater level changes. That is, they show vigour quickly when there is moisture in the soil and become senescent quite quickly during the onset of dry conditions. Therefore a considerable lag time may not have the effect on grass condition of the present time.

The relationship between groundwater depth and shrubland was significant using linear regression, as was for the other regression approaches. However, the accuracy assessments show the

linear model to be inferior (MASE = 0.913 vs. MASE < 0.9) for the other approaches. The performances of the Koyck and ARDL models in shrubland were generally similar to the ones in grassland. Although using a different modelling approach, Leung *et al*. (2011) also reported differences in groundwater level between grassland and shrubland, due to differences in stomatal resistance and root properties. Similarly, Jin *et al*. (2011) noted marked relationships between groundwater level and NDVI of grassland, woodland and scrubland respectively; however, their study did not attempt to develop a predictive model. A notable difference observed in our study was the significance of the second lag of dependent variable (NDVI) in predicting NDVI in shrubland. Unlike in grassland, the first lag of groundwater depth was not a significant predictor. It is highly likely that the vigour of a shrubland (which is represented by the NDVI) stays for relatively longer time than grassland; therefore the significance of lagged NDVI remains relevant even if groundwater depth recedes. It is also important to note that shrubland may include evergreen vegetation more so than does grassland. In both the grassland and shrubland, the third lag of groundwater depth was not a significant predictor. Although we included the third lag in our analysis, we did not anticipate to see significant contribution considering the fact that the time frequency of the data was not constant. Instead, we included the third lag as a demonstration that a reliable significance could be identified under ideal time-series data.

Table 1. Model performances of the different time-series regression analyses and using three time-lags. The standard errors are reported in parentheses.

[a]*Regressor*		Grassland					Shrubland				
	Linear	*Koyck*	*ARDL 1*	*ARDL 2*	*ARDL 3*	*Linear*	*Koyck*	*ARDL 1*	*ARDL 2*	*ARDL3*
Intercept	-0.294 (0.563)	2.482 (1.607)	0.286 (0.582)	0.517 (0.785)	1.424 (1.109)	0.329***(0.067)	0.171 (0.148)	0.311** (0.099)	0.324*** (0.106)	0.289** (0.112)
x_t	0.006 (0.008)	-0.033 (0.022)	0.018** (0.009)	0.016 (0.01)	0.015 (0.011)	-0.024* (0.012)	-0.005 (0.025)	-0.028* (0.016)	0.028* (0.016)	0.015(0.018)
y_{t-1}		0.544** (0.162)	0.69*** (0.136)	0.661*** (0.152)	0.635*** (0.152)		0.3* (0.161)	0.33** (0.148)	0.333** (0.15)	0.358** (0.149)
x_{t-1}			-0.021*** (0.007)	-0.019** (0.009)	-0.025** (0.011)			0.007 (0.015)	0.01 (0.017)	-0.007 (0.021)
y_{t-2}			-0.06 (0.142)	-0.044 (0.148)	-0.093 (0.154)			-0.337** (0.149)	-0.343** (0.152)	-0.358** (0.151)
x_{t-2}				-0.004 (0.008)	0.003 (0.01)				-0.005 (0.016)	-0.011 (0.018)
x_{t-3}					-0.012 (0.008)					0.016 (0.016)
p-value	0.4671	1.281e-05	8.411e-07	1.545e-05	2.23e-05	0.7839	1.441e-05	4.112e-05	7.015e-05	4.835e-05
R-squared	1.06%	37.46%	47.98%	48.24%	53.76%	0.15%	37.77%	38.99%	44.99%	52.61%
Adjusted R-	0.00%	34.85%	44.66%	42.36%	45.87%	0.00%	35.12%	35.01%	38.59%	44.32%

squared										
AIC	-82.69	-101.52	-108.91	-102.44	-100.70	-83.00	-99.09	-98.08	-96.62	-96.58
BIC	-76.83	-93.79	-99.26	-89.06	-83.67	-74.21	-91.45	-88.52	-83.38	-79.74
LogLik	44.34	54.76	59.46	58.22	59.35	43.00	53.55	54.04	55.31	57.29
MASE	1.063	1.217	0.803	0.804	0.901	0.913	0.888	0.867	0.863	0.872

[a] x=independent variable (groundwater depth); y=dependent variable (NDVI); t-1, t-2, t-3 represent first, second and third lags, respectively.

*, ** and *** denote the significance at the 10%, 5% and 1% levels, respectively

The fitted values and residual plots against time are given in Figure 2. Generally, the ARDL models show better fit to the observed lines than do the linear and to a lesser extent the Koyck models for both grassland (Figure 2A) and shrubland (Figure 2B). The finesses were particularly more accurate for grassland, agreeing with other statistical metrics in Table 1. The residuals plotted against time for grassland (Figure 2C) and shrubland (Figure 2D) confirm the fitted models are adequate since the horizontal band and the residuals fluctuate more or less in a random fashion within this band. Residuals plotted against the corresponding fitted values in grassland are shown in Figure 3. The high residual values in the lower fitted range compared to the rest for the linear model shows the weakness of this specific model. In contrast, the deviations from the fitted values for the other models do not exhibit fan-shaped or cone-shaped pattern, indicating the absence of heteroscedasticity – a desirable characteristic of an adequate model (Weisberg, 1985; Montgomery and Peck, 1992). Similar observations are noted for modelling in the shrubland, except that the linear model, too, did not show heteroscedasticity (Figure 4). Considering the poor statistical reports of the linear model for shrubland (Table 1), the rather favourable distributions of residuals (Figure 4) could be due to the significant coefficient estimates of the intercept and groundwater level, separately.

Previous studies focussed primarily on correlation analysis between NDVI and groundwater or other climatic variables such as precipitation. The study by Zhu *et al*. (2015) applied Pearson's correlation analysis and reported mixed results between groundwater and NDVI with three of six sites returning significant correlations. However it is difficult to make a direct comparison with our study since we used regression analysis. Zhu et al's study had another limited similarity with ours in that they considered up to three lag times (months), though the lag was applied to precipitation that was related to NDVI. We therefore consider our application of two lag times sensible, even if the variable of interest was different (groundwater).
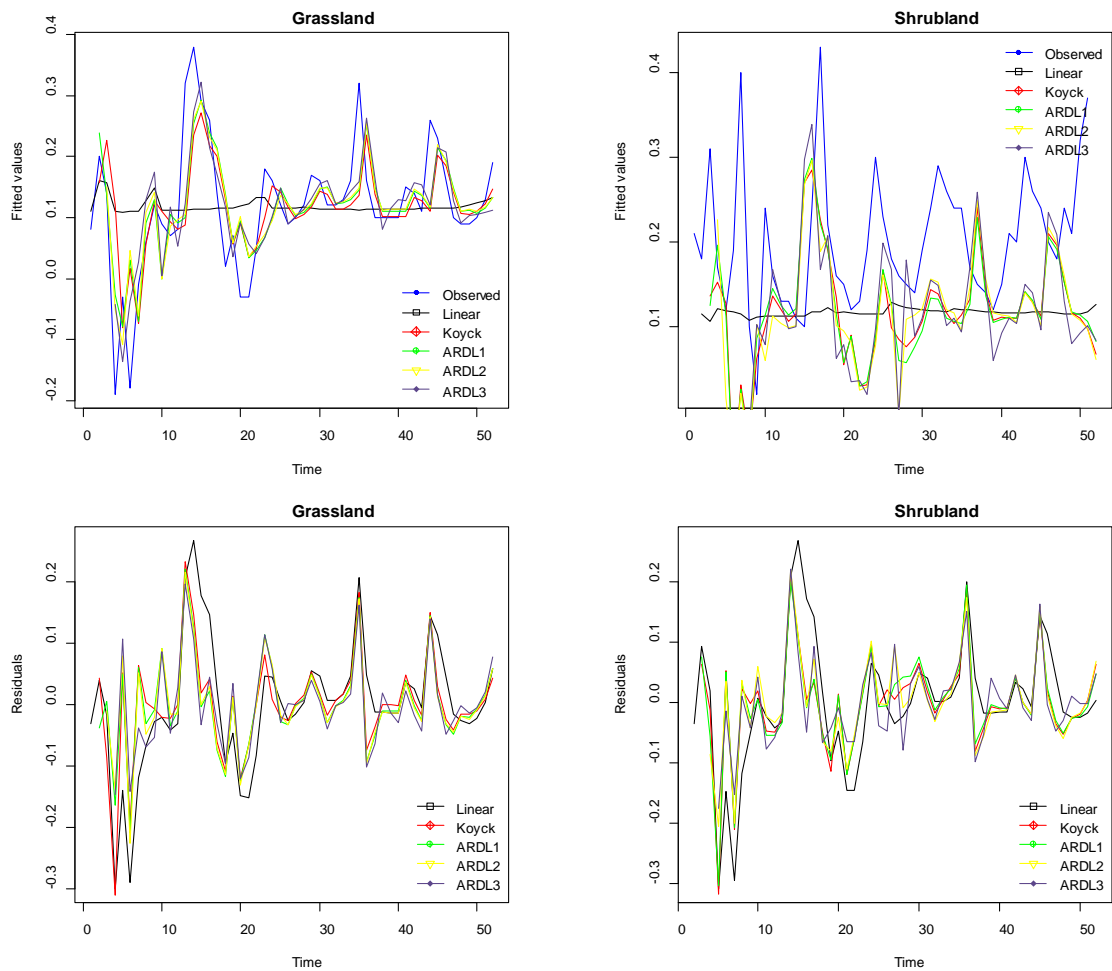
Figure 2. Fitted values against time for linear, Koyck, ARDL1, ARDL2 and ARDL3 models (A & B); residuals against time for linear, Koyck, ARDL1, ARDL2 and ARDL3 models (A & B).
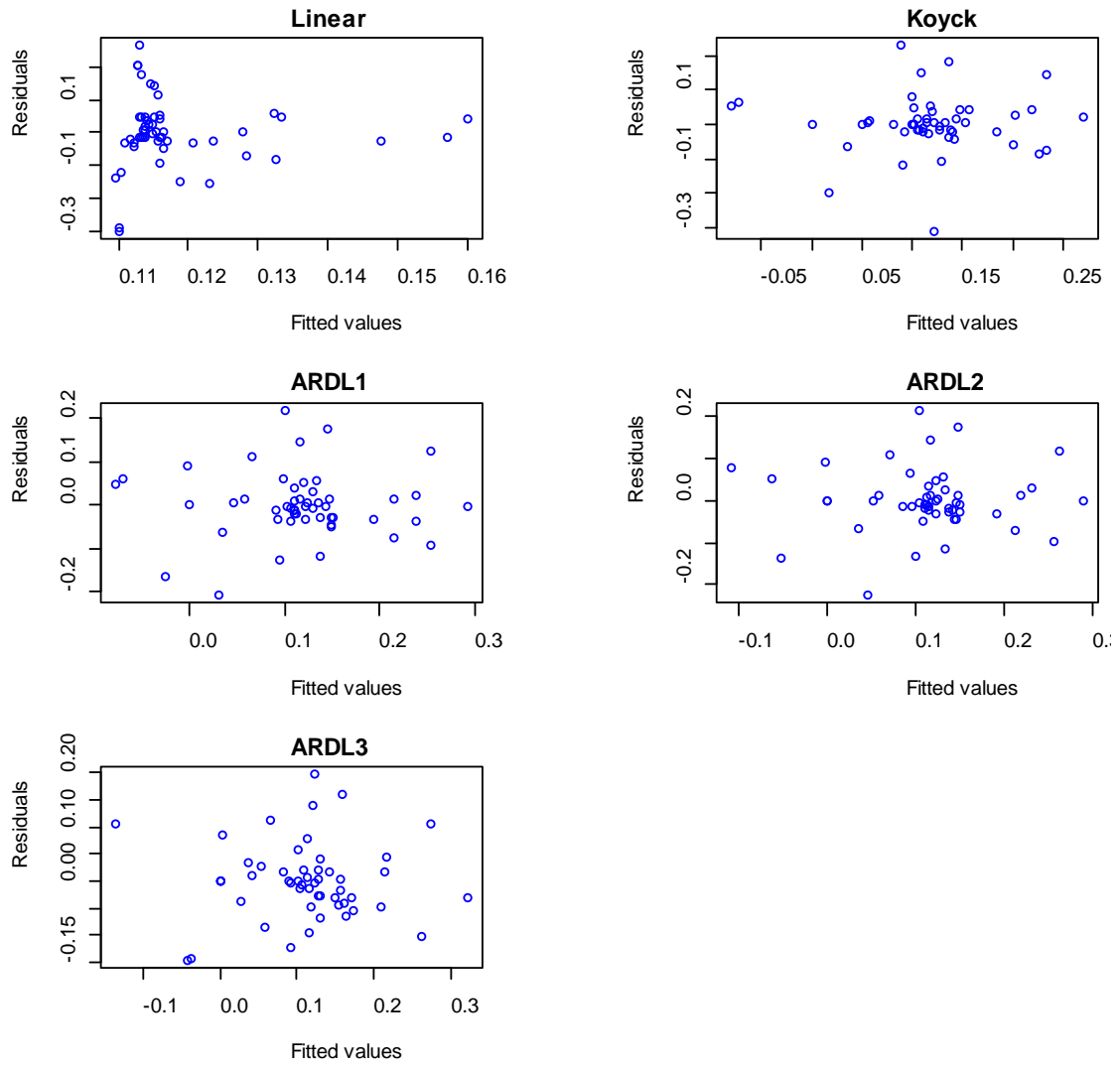
Figure 3: Residual plots versus fitted values for linear, Koyck, ARDL1, ARDL2 and ARDL3 models for grassland location.
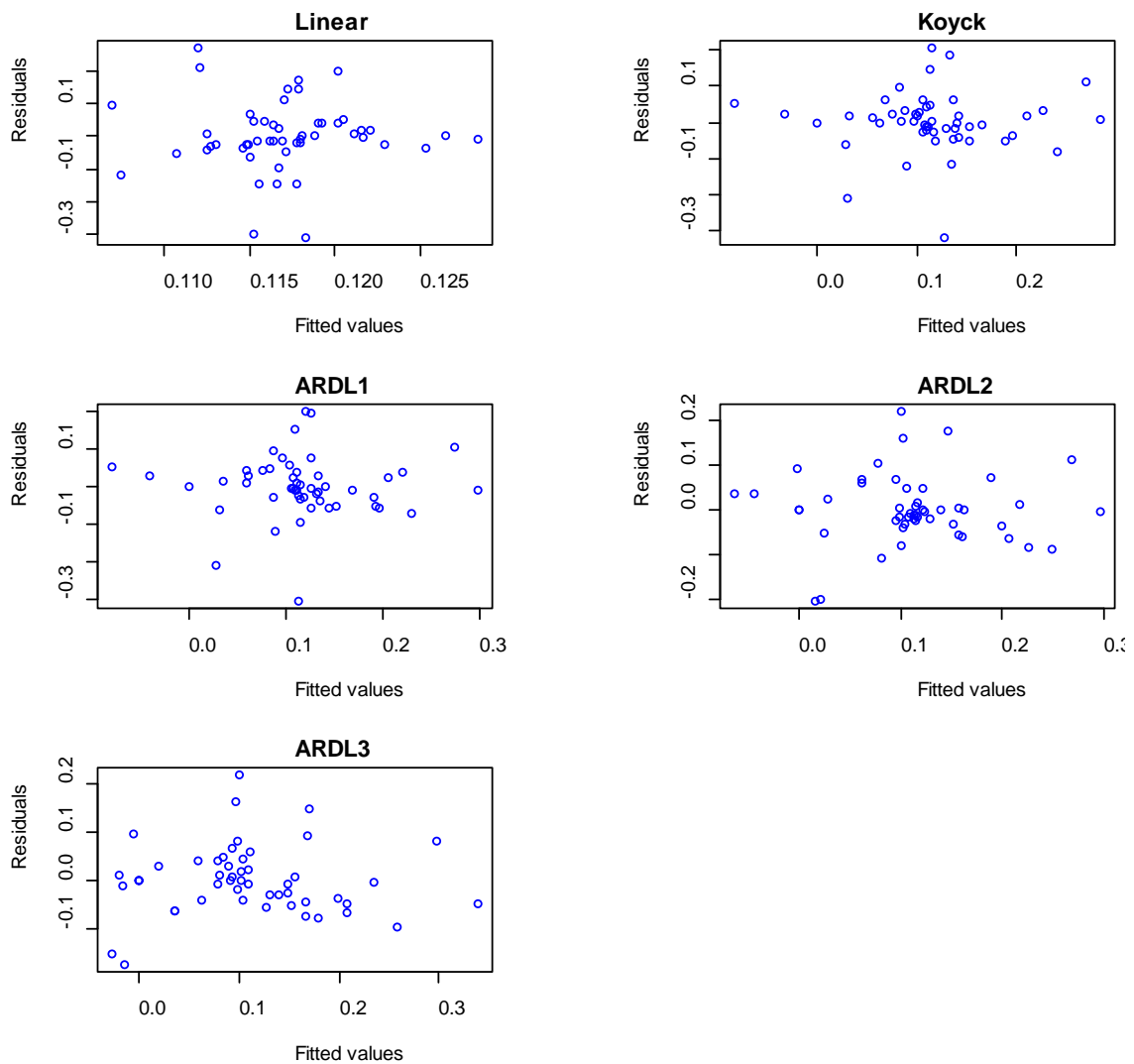
Figure 4: Residual plots versus fitted values for linear, Koyck, ARDL1, ARDL2 and ARDL3 models for shrubland location

5. Conclusions

In this paper, we studied the relationship between NDVI and depth of groundwater. For grassland and shrubland locations, the first lag of NDVI has a significant effect on the current value of NDVI. Therefore, modelling the relationship between NDVI and depth of groundwater involve both present and past values of this variable. This agrees with previous studies, for example Zhu *et al.* (2015) who correlated lagged precipitation with NDVI. The adjusted R-squared, AIC, BIC, log likelihood and MSAE results suggested that ARDL1 and ARDL2 models are performing better than the other competitive models considered in this paper for the grassland and shrubland locations, respectively. The study also showed the importance of factoring in vegetation type when building a relationship between groundwater depth and NDVI with shrubland showing greater dependence on relatively distant times of vegetation condition. One of the major limitations of the studies was the inconsistency in the time frequency of both the Landsat and the ground water data, forcing the study to apply imputations. In addition the study focussed only on two areas in order to investigate the utility of previous response and explanatory variables to predict a response variable. Time

frequency can be achieved by utilizing alternative data sources such as MODIS data that provides data at multiple spatial and temporal scales. Employing such data however must be matched by appropriate groundwater data that represent large spatial area. Regarding sampling, the approach tested in this study can be extended at larger spatial areas such as provincial and regional level assessments, in which several groundwater observations can be used. Applying the approach at larger spatial extents suits well with the use of such data as MODIS that provide reliable time-frequency compared to what would be offered by Landsat data collection.

6. **Acknowledgments**

7. **References**

Akaike, H 1974, 'A new look at the statistical model identification', *IEEE Transactions on Automatic Control*, vol. 19, pp. 716 – 723.

Barron, OV, Emelyanova, I, Van Niel, TG, Pollock, D & Hodgson, G 2014, 'Mapping groundwater-dependent ecosystems using remote sensing measures of vegetation and moisture dynamics', *Hydrological Processes*, vol. 28, pp. 372-385.

Chander, G, Markham, BL & Helder, DL 2009, 'Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors', *Remote Sensing of Environment*, vol. 113, pp. 893-903.

Chen, Y., Li, W, Xu, C, Ye, Z & Chen, Y 2015, 'Desert riparian vegetation and groundwater in the lower reaches of the Tarim River basin', *Environmental Earth Sciences*, vol. 73, pp. 547-558.

Chen, YN, Zilliacus, H, Li, WH, Zhang, HF & Chen, YP 2006, 'Ground-water level affects plant species diversity along the lower reaches of the Tarim river, Western China', *Journal of Arid Environments*, vol. 66, pp. 231-246.

Chinnasamy, P, Maheshwari, B, Dillon, P, Purohit, R, Dashora, Y, Soni, P & Dashora, R 2018, 'Estimation of specific yield using water table fluctuations and cropped area in a hardrock aquifer system of Rajasthan, India', *Agricultural Water Management*, vol. 202, pp. 146-155.

Cooper, DJ, Sanderson, JS, Stannard, DI & Groeneveld, DP 2006, 'Effects of long-term water table drawdown on evapotranspiration and vegetation in an arid region phreatophyte community', *Journal of Hydrology*, vol. 325, pp. 21-34.

DEA 2015, '2013–2014 South African national land cover data set' Prepared by Geoterraimage, South Africa for the Department of Environmental Affairs, South Africa.

Elmahdy, SI & Mohamed, MM 2014, 'Groundwater potential modelling using remote sensing and GIS: a case study of the Al Dhaid area, United Arab Emirates', *Geocarto International*, vol. 29, pp. 433-450.

Fan, J, Ostergaard, KT, Guyot, A, Fujiwara, S & Lockington, DA 2016, 'Estimating groundwater evapotranspiration by a subtropical pine plantation using diurnal water table fluctuations: implications from night-time water use', *Journal of Hydrology*, vol. 542, pp. 679-685.

Franses, PH 2016, 'A note on the Mean Absolute Scaled Error', *International Journal of Forecasting,* vol. 32, pp. 20-22.

Fu, B & Burgher, I 2015, 'Riparian vegetation NDVI dynamics and its relationship with climate, surface water and groundwater', *Journal of Arid Environments*, vol. 113, pp. 59-68.

Ghose, D, Das, U & Roy, P 2018, 'Modeling response of runoff and evapotranspiration for predicting water table depth in arid region using dynamic recurrent neural network', *Groundwater for Sustainable Development*, vol. 6, pp. 263-269.

Gou, A, Gonzales, S & Miller, GR 2015, 'Mapping Potential Groundwater-Dependent Ecosystems for Sustainable Management', *Groundwater*, vol. 53, pp. 99-110.

Gow, LJ, Barrett, DJ, Renzullo, LJ, Phinn, SR & O'Grady, AP 2016, 'A detection problem: sensitivity and uncertainty analysis of a land surface temperature approach to detecting dynamics of water use by groundwater-dependent vegetation', *Environmental Modelling and Software*, vol. 85, pp. 342-355.

Huang, Y, Chang, Q & Li, Z 2018, 'Land use change impacts on the amount and quality of recharge water in the loess tablelands of China', *Science of the Total Environment*, vol. 628–629, pp. 443-452.

Hyndman, RJ & Koehler, AB 2006, 'Another look at measures of forecast accuracy', *International Journal of Forecasting*, vol. 22, pp. 679-688.

Jeong, J, Park, E, Han, WS, Kim, K-Y, Suk, H & Jo, SB 2018, 'A generalized groundwater fluctuation model based on precipitation for estimating water table levels of deep unconfined aquifers', *Journal of Hydrology*, vol. 562, pp. 749-757.

Jin, X, Liu, J, Wang, S & Xia, W 2016, 'Vegetation dynamics and their response to groundwater and climate variables in Qaidam Basin, China', *International Journal of Remote Sensing*, vol. 37, pp. 710-728.

Jin, XM, Schaepman, ME, Clevers, JGPW, Su, ZB & Hu, GC 2011, 'Groundwater depth and vegetation in the Ejina Area, China', *Arid Land Research and Management*, vol. 25, pp. 194-199.

Koide, K & Koike, K 2012, 'Applying vegetation indices to detect high water table zones in humid warm-temperate regions using satellite remote sensing', *International Journal of Applied Earth Observation and Geoinformation*, vol. 19, pp. 88-103.

Le Maitre, DC, Scott, DF & Colvin, C 1999, 'A review of information on interactions between vegetation and groundwater', *Water SA*, vol. 25, pp. 137-152.

Leung, LR, Huang, M, Qian, Y & Liang, X 2011, 'Climate-soil-vegetation control on groundwater table dynamics and its feedbacks in a climate model', *Climate Dynamics*, vol. 36, pp. 57-81.

Liu, B, Zeng, FJ, Arndt, SK, He, JX, Luo, WC & Song, C 2013, 'Patterns of root architecture adaptation of a phreatophytic perennial desert plant in a hyperarid desert', *South African Journal of Botany*, vol. 86, pp. 56-62.

Montgomery, DC & Peck, EA 1992, '*Introduction to Linear Regression Analysis*, 2nd Ed'. Wiley, New York.

Moritz, S 2016, 'Time series missing value imputation, R Foundation for Statistical Computing', URL: https://cran.r-project.org/web/packages/imputeTS/imputeTS.pdf.

Orellana, F, Verma, P, Loheide II, SP & Daly, E 2012, 'Monitoring and modelling water-vegetation interactions in groundwater-dependent ecosystems', *Reviews of Geophysics*, vol. 50, viewed 02 May 2018, <http://dx.doi.org/10.1029/2011RG000383>.

Rodriguez, J, Ustin, S, Sandoval-Solis, S & O'Geen, AT 2016, 'Food, water, and fault lines: remote sensing opportunities for earthquake-response management of agricultural water', *Science of the Total Environment*, vol. 565, pp. 1020-1027.

Sakala, E, Tessema, A & Nyabeze, PK 2014, 'Regional interpretation of aeromagnetic data for groundwater exploration in Capricorn District, Limpopo, South Africa', *International Journal of Modelling and Simulation*, vol. 34, pp. 36-42.

Shekhar, S & Pandey, AC 2015, 'Delineation of groundwater potential zone in hard rock terrain of India using remote sensing, geographical information system (GIS) and analytic hierarchy process (AHP) techniques', *Geocarto International*, vol. 30, pp. 402-421.

Uwe, H & Jürgen, W 2005, 'Autoregressive distributed lag models and cointegration', Diskussionsbeiträge des Fachbereichs Wirtschaftswissenschaft der Freien Universität Berlin, No. 2005/22, Freie Univ., Fachbereich Wirtschaftswiss., Berlin.

Wada, Y, van Beek, LPH, van Kempen, CM, Reckman, JWTM, Vasak, S & Bierkens, MFP 2010, 'Global depletion of groundwater resources', *Geophysical Research Letters*, vol. 37, L20402, doi:10.1029/2010GL044571.

Weisberg, S 1985, *Applied Linear Regression*, 2nd Ed. Wiley, New York.

Werner, AD, Sharp, HK, Galvis, SC, Post, VEA & Sinclair, P 2017, 'Hydrogeology and management of freshwater lenses on atoll islands: Review of current knowledge and research needs', *Journal of Hydrology*, vol. 551, pp. 819-844.

Wiedermann, MM, Kane, ES, Potvin, LR & Lilleskov, EA 2017, 'Interactive plant functional group and water table effects on decomposition and extracellular enzyme activity in Sphagnum peatlands', *Soil Biology and Biochemistry*, vol. 108, pp. 1-8.

Yeh, HF, Cheng, YS, Lin, HI & Lee, CH 2016, 'Mapping groundwater recharge potential zone using a GIS approach in Hualian River, Taiwan', *Sustainable Environment Research*, vol. 26, pp. 33-43.

Yin, L, Zhou, Y, Huang, J, Wenninger, J, Zhang, E, Hou, G & Dong, J 2015, 'Interaction between groundwater and trees in an arid site: potential impacts of climate variation and groundwater abstraction on trees', *Journal of Hydrology*, vol. 528, pp. 435-448.

Zhu, L, Gong, H, Dai, Z, Xu, T & Su, X 2015, 'An integrated assessment of the impact of precipitation and groundwater on vegetation growth in arid and semiarid areas', *Environmental Earth Sciences*, vol. 74, pp. 5009-5021.