Dynamic factor analysis of surface water management impacts on soil and bedrock water contents in Southern Florida Lowlands

I. Kisekka, K. W. Migliaccio, R. Muñoz-Carpena, B. Schaffer, Y. C. Li

Published Version Information

Citation: Kisekka, I., Migliaccio, K. W., Muñoz-Carpena, R., Schaffer, B., & Li, Y. C. (2013). Dynamic factor analysis of surface water management impacts on soil and bedrock water contents in Southern Florida Lowlands. Journal of Hydrology, 488, 55-72.

Digital Object Identifier (DOI): doi:10.1016/j.jhydrol.2013.02.035

Publisher's Link: http://www.sciencedirect.com/science/article/pii/S0022169413001583

Dynamic Factor Analysis of Surface Water Management Impacts on Soil and Bedrock Water Contents in Southern Florida Lowlands

I. Kisekka[a], K.W. Migliaccio[a], R. Muñoz-Carpena[b], B. Schaffer[a], Y.C. Li[a]

University of Florida, IFAS Tropical Research and Education Center 18905 SW 280th St Homestead, FL 33031[a]

University of Florida, Agricultural and Biological Engineering Department, P. O. Box 110570, Gainesville, FL 32611[b]

[a]Corresponding author: Kati W. Migliaccio; Tel.: +1 305-246-7001 x288; fax: +1 305-246-7003.

E-mail address: klwhite@ufl.edu.

Abstract

As part of the C111 spreader canal project, structural and operational modifications involving incremental raises in canal stage are planned along one of the major canals (i.e., C111) separating Everglades National Park and agricultural production areas to the east of the park. This study used Dynamic Factor Analysis (DFA) as an alternative tool to physically based models to explore the relationship between different hydrologic variables and the effect of proposed changes in surface water management on soil and bedrock water contents in south Florida. To achieve the goal, objectives were to: (1) use DFA to identify the most important factors affecting temporal variation in soil and bedrock water contents, (2) develop a simplified DFA based regression model for predicting soil and bedrock water contents as a function of canal stage and (3) assess the effect of the proposed incremental raises in canal stage on soil and bedrock water contents. DFA revealed that 5 common trends were the minimum required to describe unexplained variation in the 11 time series studied. Introducing canal stage, water table evaporation and net recharge resulted in lower Akaike information criterion (AIC) and higher Nash-Sutcliffe (C_{eff}) values. Results indicated that canal stage significantly ($t > 2$) drives temporal variation in soil and bedrock water contents, which was represented as scaled frequency while net surface recharge was significant in 7 out of the 11 time series analyzed. The effect of water table evaporation was not significant at all sites. Results also indicated that the most important factor influencing temporal variation in soil and bedrock water contents in terms of regression coefficient magnitude was canal stage. Based on DFA results, a simple regression model was developed to predict soil and bedrock water contents at various elevations as a function of canal stage and net recharge. The performance of the simple model ranged from good (C_{eff} ranging from 0.56 to 0.74) to poor (C_{eff} ranging from 0.10 to 0.15), performance was better at sites with smaller depths to water table (< 1 m) highlighting the effect of micro-topography on soil and bedrock water content dynamics. Assessment of the effect of 6, 9 and 12 cm increases in canal stage using the simple regression model indicated that changes in temporal variation in soil and bedrock water contents were negligible (average<1.0% average change) at 500 to 2000 m from C111 (or low elevations) which may be attributed to the near saturation conditions already occurring at these sites. This study used DFA to explore the relationship between soil and bedrock water dynamics and surface water stage in shallow water table environments. This approach can be applied to any system in which detailed physical modeling would be limited by inadequate information on parameters or processes governing the physical system.

Key words: Soil water content, bedrock water content, scaled frequency, Dynamic Factor Analysis, canal stage, water table

Abbreviations: DFA, dynamic factor analysis; SF, scaled frequency; R_{net}, net surface recharge; MWT, mean water table elevation; S177T, C111 canal stage; SFWMD, South Florida Water Management District; AIC, Akaike information criterion; BIC, Bayesian information criterio; VIF, variance inflation factor; NGVD29, National Geodetic Vertical Datum of 1929.

1. Introduction

In an attempt to correct some of the undesired consequences of south Florida's extensive drainage canal network on the region's ecosystem, an environmental restoration project named the Comprehensive Everglades Restoration Plan (CERP) is currently under implementation. CERP was approved by the United States Congress under the Water Resources Development Act (2000). One of the 68 components that comprise CERP is the C111 spreader canal project whose goal is to reduce the impacts of C111 (i.e., reduce groundwater seepage into C111) on Everglades National Park (ENP) and Taylor Slough which is a natural drainage feature that conveys water to Florida while maintaining existing levels of flood protection in the adjacent agricultural and urban areas (U.S. Army Corps of Engineers [USACP] and South Florida Water Management District [SFWMD], 2009). As part of the C111 spreader canal project, structural modifications and operational adjustments involving incremental raises in canal stage are planned along one of the major canals (i.e., C111) separating ENP and agricultural production areas to the east of the canal. The increase in canal stage will occur by changing surface water management at the gated spillway located at structure named S18C (Fig. 1) in the form of incremental raises in canal stage of up to 12 cm.

It is anticipated that the planned rise in C111 canal stage will affect water table levels in the adjacent agricultural areas. Earlier research indicated that there is substantial interaction between the highly permeable Biscayne aquifer and water level in canals (Genereux and Slater, 1999). The hydraulic connection between Biscayne aquifer and canal C111 causes the shallow water table system to fluctuate with respect to changes in canal stage. Using the drain to equilibrium assumption, Barquin et al. (2011) showed that water table elevation in the Biscayne aquifer significantly influenced soil and bedrock water contents in a fruit orchard with soil and bedrock formations that are very similar to our current study site. Therefore, raising water table elevation could result in increased soil and bedrock water contents or greater saturation of the root zone which could affect the production of winter vegetables predominately grown in this area. Saturation of the root zone could impact yield potential by impairing root growth due to anoxia, reducing stomatal conductance, and reducing net CO_2 assimilation (Schaffer, 1998). In addition to physiological stress, having the soils saturated could render movement of machinery difficult and also impact growing season and market dates. However, it is not known to what extent the proposed structural modifications and operational adjustments along canal C111 would impact water table elevations and thus soil and bedrock water contents in agricultural areas east of the canal.

Vegetable production in Miami-Dade County, a substantial proportion of which is located along the extensive eastern boundary of ENP, is a significant contributor to both the local and state economies. According to the 2007 Census of Agriculture from the US Department of Agriculture (USDA, 2007), the total value of vegetables produced in Miami-Dade County was over 128 million dollars in 2007. Green beans, sweet corn, squash, tomatoes and sweet potato are the dominant vegetables grown in the area. There is need to quantify the impacts of hydrological modifications and surface water management on agricultural land use at field scale because large regional hydrology models have discretization that might not be suitable for resolving small scale micro-topographic differences within the landscape.

Long term monitoring and exploratory analysis of soil and limestone bedrock water contents could characterize the effect of various drivers on the temporal variability of water contents. The soils in the agricultural areas east of C111 were created from scarification of the underlying limestone bedrock hence they are very shallow and have high gravel content. Three main stresses that influence soil water content that could be included in exploratory analysis are 1) canal stage, which affects water table elevation; 2) rainfall, and 3) evapotranspiration. While these stresses may be assessed using physically based models of vadose zone flow and transport, implementation of unsaturated flow models (e.g., WAVE [Vanclooster et al., 1995] or HYDRUS [Šimůnek, et al., 2008]) is not an easy task since they contain numerous parameters and processes that have to be quantified (Ritter et al., 2009). In very gravelly and shallow soils such as those in south Miami-Dade County, quantifying parameters such as hydraulic conductivity for use in Richards' equation is further complicated by having porous gravely soils that are not homogeneous. Previous applications of WAVE, for example, in gravely soils of south Florida have indicated that a detailed description of soil hydraulic properties (e.g., using dual porosity) could result in improved robustness of vadose zone models (Duwig et al., 2003; Muñoz-Carpena et al., 2008). Therefore the success of applying physically based models to simulate soil and bedrock water dynamics depends largely on proper conceptualization of location specific processes and proper measurement or estimation of parameters. In this context, complementary exploratory tools such as Dynamic Factor Analysis (DFA) which are not processes based are desired as simpler preliminary exploratory tools that could also be used for preliminary predictions of the impact of surface water management decisions on land use.

A comprehensive description of DFA and modeling can be found in Zuur et al. (2003). For purposes of aiding discussion, we only provide a brief description of this technique. DFA is a dimension reduction multivariate time series analysis technique that is used to estimate underlying common patterns (common trends) in short time series as well as the effect of explanatory variables on response variables. The advantage of DFA over other traditional dimensional reduction techniques (e.g., Factor Analysis or Principal Component Analysis) is that DFA accounts for the time component. This allows the underlying hidden effects driving the temporal variation in the observed time series data to be detected (Zuur et al., 2003). DFA does not require observed time series to be long and stationary. Although non-stationarity could be handled through de-trending, trends in the times series could hold necessary information required to explain the temporal dynamics in the observed variable (Ritter et al., 2009). In addition, DFA can handle missing values in the observed time series (i.e., DFA does not require data sets to be regularly spaced). Missing values in observed time series data sets are not uncommon especially when time series data are obtained from unattended automatic data logging field instruments (e.g., multi-sensor capacitance probes for soil water monitoring).

DFA applications are documented in literature from several disciplines (e.g., Geweke, 1977; Márkus et al., 1999; Zou and Yu, 1999; Zuur et al., 2003; Zuur and Pierce, 2004; Muñoz-Carpena et al., 2005; Ritter and Muñoz-Carpena, 2006; Zuur et al., 2007; Ritter et al., 2009; Kaplan et al., 2010a; Kaplan and Muñoz-Carpena, 2011). Thus, we only provide a brief review of the most relevant examples. Ritter and

Muñoz-Carpena (2006) applied DFA and modeling to study interactions between surface water and groundwater levels within the Frog Pond agricultural area located west of canal C111 in south Florida (Fig.1). Their results indicated that the two canals surrounding the Frog Pond area had the greatest influence on temporal changes in water table elevation. Their study did not address the issue of the impact of surface management decisions on soil water content. Soil water is a major concern for vegetable growers in south Florida due to the impact saturated or near saturated soil conditions have on planting dates and yield losses (Fig. 1).

Others have applied DFA and modeling to study soil water dynamics. Ritter et al. (2009) applied DFA to analyze temporal changes in soil water status of a humid, subtropical, evergreen forest in Canary Islands, Spain. Kaplan and Muñoz-Carpena (2011) applied DFA to study the complementary effects of surface and groundwater on soil water dynamics in a coastal flood plain. Thus, DFA was successfully used to identify unexplained variability in observed hydrologic time series and to assess the effect of selected explanatory variables on response variables (observed time series of interest).

The difference between our study and prior studies is that we applied DFA to investigate the effect of surface water management in canals on soil water dynamics in an agricultural area with very shallow very gravely loam soils, and unlike in the previous studies we also considered not only the effects of potential evaporation (ET_o) but also the effect of water table evaporation given the shallow water table. We then attempted to develop a simple model, using information from the DFA, to predict soil water content from easily measured variables such as canal stage and recharge (i.e., difference between rainfall and evapotranspiration). Canal stage was selected instead of water table elevation since water table elevation data in our study area are less complete due to the limited period of record and the limited number of continuously monitored groundwater wells. Canal stage has been monitored for a longer period of record and has no foreseeable end of data collection, thus it is a more reliable measurement for long-term use. We assumed that at any given time, water table elevation is approximately equal to canal stage. We concede that at certain times this assumption might not hold e.g., immediately after or during storm events; however, due to the high permeability of the aquifer and the daily time step used, the assumption holds for the majority of the time.

The goal of this study was to use DFA and modeling to investigate how the proposed raises in canal stage along C111 could impact soil and bedrock water contents in low lying farmlands located between canals C-111 and C-111E. The specific objectives were to: (1) apply DFA to identify the most important factors affecting temporal variation in soil and bedrock water contents, (2) develop a simplified DFA based regression model for predicting soil and bedrock water contents as a function of canal stage, and (3) use the developed simple regression model to predict the impact of proposed incremental raises in canal stage on soil and bedrock water contents at various elevations and distances from the canal.

2. Materials and methods

2.1 Experimental site

The study was conducted in southern Miami-Dade County, Homestead, Florida, United States in a small agricultural area approximately 17 km^2 (Fig. 1). The area is located east of ENP between SFWMD canals C111 and C111E which are planned to experience increases in canal stage under the C111 spreader canal project. Canal stage upstream in the two canals is controlled by a remotely operated spillway at S177 and a culvert at S178, respectively (Fig. 1). C111 is the larger of the two canals and the two join to become a single canal at the southern end of the study area which is managed using a gated spillway at S18C. It is proposed that stage will be increased by modifying operation of S18C and thus affect canal stage in the reach of C111 between S177 and S18C. The hydrogeological system at the study site consists of the Biscayne aquifer which is a highly permeable shallow unconfined aquifer with hydraulic conductivities reported to exceed 10,000 m/day, which explains the high connectivity between the canals and the aquifer (Chin, 1991). The shallow nature of the water table implies that evaporation from the groundwater could impact soil water content. The topography at this site is essentially flat with elevation ranging approximately between 1.2 to 2.0 m above sea level NGVD 29. The climate is subtropical with dry season (November to May), which is the growing season for vegetables, and wet season (June to October). Approximately two thirds of all the rain (average annual rainfall ranges between 1100 to 1524 mm) is received during the wet season months.

The soil at the study site is very shallow (10 to 20 cm) with underlying limestone bedrock. According to Nobel et al. (1996), the soils east of C111 vary and could be classified as either Krome and Chekika very gravely loam (loamy skeletal, carbonatic, hyperthermic, Lithic Undorthents), or Biscayne Marl (loamy, carbonatic, hyperthermic) based on their physical characteristics. We performed particle size analysis using a standard 2-mm sieve and determined that the soils contain on average of 45% fine fractions and 55% gravel. Color analysis using the Munsell soil color charts (Munsell soil charts, 2000) and the color guide in Noble et al. (1996) identified the study site soils to be broadly characterized as Chekika soil series.

Three monitoring sites were used in this study located at 500, 1000 and 2000 m along a transect perpendicular to canal C111, the three sites also had varying topographies and represented areas expected to experience the greatest impact from the proposed raises in canal stage. Sites were selected to capture differences in soil texture within our study area; this was done with a soil survey map and site visits. Sites were also selected to ensure they were in privately owned agricultural low lying lands that were expected to be impacted by the rises in water table elevation. For each site: i) GPS coordinates and elevation data were collected, ii) groundwater wells were constructed and each was equipped with level loggers (Levelogger, Gold Solinst Canada Ltd., 35 Todd Rd, Georgetown, Ontario, Canada) to record water table elevation every 15 minutes, iii) multi-sensor capacitance probes (MSCP) (EnviroScan probes, Sentek Technologies, Ltd., Stepney, Australia) were installed at each site to monitor soil and bedrock water contents. Monitoring site locations are shown in Fig. 1; elevations are shown in Fig.2. Differences in the length of times series at the three sites was due to differences in the dates of installation of the EnviroScan probes (i.e., probes could only be installed when water was at least 50 cm below the ground surface) and relocation of the probes due to initial poor installation. Site T500 was installed on August 25, 2010, while sites T1000 and T500 were installed on January 21, 2011.

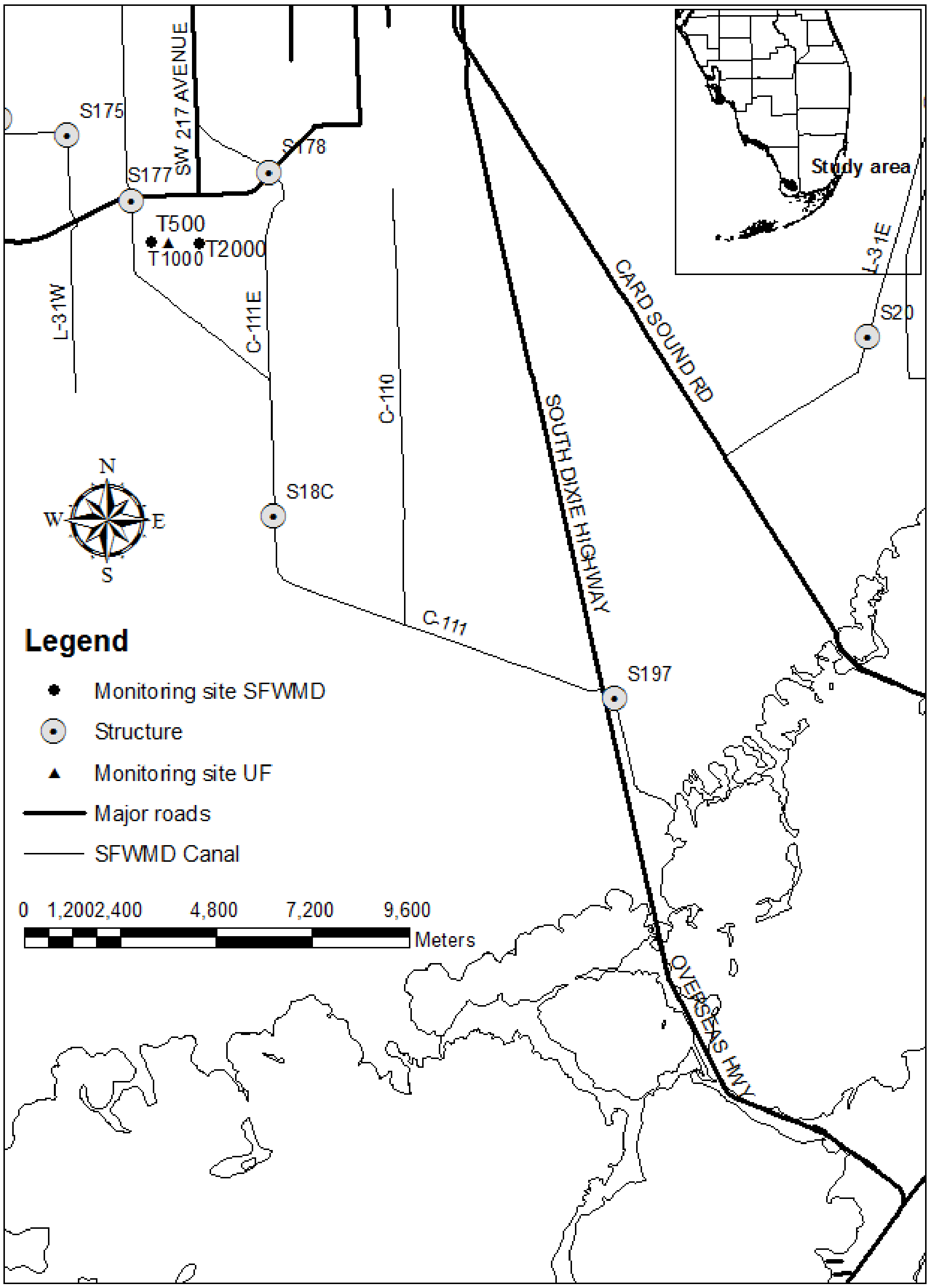

Figure 1. . Map of the study area showing Everglades National Park, Taylor Slough, Florida Bay, SFWMD canal network and low lying agricultural areas east of canal C111 in south Florida

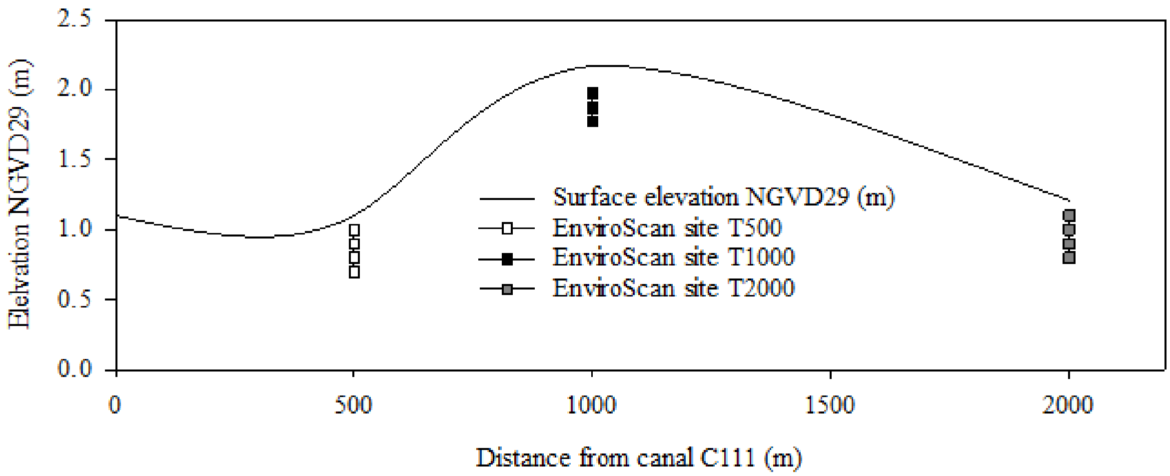

Figure 2. Showing a topographic changes along transect T and the elevation of the EnviroScan sensors at the three sites.

2.2 Soil and bedrock water contents monitoring

Two EnviroScan probes were installed at each site for a total of six. Each access tube with a diameter of 50.5 mm housed four sensors positioned at various elevations as shown in Fig. 2. The elevations correspond to 10, 20, 30 and 40 cm from the ground surface at each site. The top 20 cm typically represent the scarified soil layer which is used for crop production and the lower 20 cm represent the underlying limestone bedrock in which plant roots cannot penetrate. To minimize the problem of air pockets, we used fast setting cement slurry between the access tube and the soil. The purpose of installing two EnviroScan probes at the same location was to ensure that at least one probe was functioning at any given time. Due to the shallowness of the limestone bedrock at all the study sites, a motorized drill was required to bore a hole that held the access tube in a vertical position. Water content data were logged every 15 minutes and were downloaded weekly and averaged daily.

EnviroScans are an example of capacitance based sensors which measure frequency of an oscillating electrical circuit. The oscillator is coupled electrically to capacitive elements that are made of two metal cylindrical electrodes. The electrode system is arranged so the soil becomes part of the dielectric medium affected by the fringing electromagnetic field. Volumetric soil water content affects the electrical permittivity of the soil which in turn affects the capacitance causing the oscillation frequency to shift (IAIA, 2008) since the soil dielectric constant is a combination of mineral particles (2-4), water (80), and air (1). According to Dean et al. (1987) the oscillatory frequency from the capacitance soil water sensor could be expressed eq. (1):

$$F = \frac{1}{2\pi\sqrt{L}}\left(\frac{1}{C}+\frac{1}{C_b}+\frac{1}{C_c}\right)^{1/2} \tag{1}$$

Where C_b is the total base capacitance and C_c is the total collector capacitance and these represent capacitances of internal circuit elements to which the electrodes are connected, L is the inductance of the coil in the circuit, and C is the capacitance of the soil access tube system. Therefore capacitance of the soil access tube system, C, can be expressed as a function of the soil dielectric constant (ε) and a value g representing the geometry of the sensor as shown in eq.(2).

$$C = g\,\varepsilon \tag{2}$$

Differences in oscillatory frequency among sensors at the same soil and bedrock water contents were eliminated by normalizing the oscillatory frequency values using values of frequency when the sensor was surrounded by water and air. The normalized oscillatory frequency is known as the scaled frequency (*SF)* and is estimated as in eq. 3. The manufacture default calibration equation (eq. 4) can be used to convert scaled frequency to volumetric soil water content (θ)

$$SF = F - F_a / F_w - F_a \tag{3}$$

$$\theta = (0.792 * SF - 0.0226)^{2.475} \tag{4}$$

where F is the oscillatory frequency value measured by the EnviroScan sensor, F_a is frequency value when the EnviroScan probe is surrounded by air, and F_w is the frequency value when the EnviroScan probe is surrounded by water. To avoid location specific calibration for each sensor, we use SF as surrogate for θ for investigating the effect of various factors on soil and bedrock water contents and thus did not use eq. (4). This approach was successfully applied by Ritter et al. (2009) when studying the effect of various factors on hydrologic fluxes in a forest top soil using refractive index from time-domain reflectometry (TDR) as a surrogate for volumetric soil water content. Gabriel et al. (2010) observed that the manufacturer's calibration equation overestimated volumetric soil water compared to the locally developed calibration equation. However, they noted that despite the overestimation of volumetric soil water content, the manufacturer's equation was able to reproduce temporal soil water dynamics. Therefore, if the goal is to measure relative changes in water content the manufacturer's default calibration equation is sufficient.

2.3 Measurement and estimation of hydrologic variables

Hydrologic variables including canal stage, water table elevation NGVD29 m, rainfall (*P*), potential evapotranspiration (ET_o) and groundwater evaporation (*E*) were measured or estimated to assess their influence on soil and bedrock water content time series.

2.3.1 Canal stage

Canal stage data were measured at the S177 spillway for headwater (S177H) and tail water (S177T) every 15 minutes but daily averages were used. Canal stage data were measured by the SFWMD and are publically available from the online environmental database (DBhydro; http://www.sfwmd.gov/dbhydroplsql/show_dbkey_info.main_menu). During the first phase of the C111 spreader canal project, the main operational adjustments will involve incrementally raising canal stage at S18C (Fig. 1) which will result in increased stage in the reach of C111 between the spillways at S177 and S18C.

2.3.2 Water table elevation

Water table elevation data were collected from three observation wells constructed at the three monitoring sites. Water table elevation was measured by the University of Florida (UF) every 15 minutes and averaged daily using a multi parameter pressure transducer at T1000 (Levelogger, Gold Solinst Canada Ltd., 35 Todd Rd, Georgetown, Ontario, Canada). Atmospheric corrections were included using a STS Barologger (Solinst Canada Ltd) in the well at T1000 (Fig. 1). Data were downloaded from the well weekly and as a quality control procedure, water table elevations were also measured manually with a Model 102 Laser water level well meter (Solinst, Canada Ltd). Wells T2000 (C111AE) and T500 (C111AW) were installed and operated by the SFWMD and published on DBHydro.

2.3.3 Rainfall

Gauge adjusted Next Generation Radar (NEXRAD) rainfall data used in this study were obtained from the SFWMD. The United States National Weather Service operates two NEXRAD sites close to the study site (i.e., KBYX in Key West, FL and KAMX in Miami, FL) that provide 2 km x 2 km NEXRAD rainfall data. There are tradeoffs between rainfall estimated by rain gauges and NEXRAD. Rain gauges (e.g., tipping buckets) provide accurate point estimates of rainfall which are acceptable for frontal related rainfall events. However, in South Florida where most of the rainfall is received in summer and summer rainfall is dominated by conventional or tropical rainfall forming processes, rain gauges may fail to accurately represent the orientation of the rainfall front or fail to capture the entire rainfall event (Pathak, 2008). On the other hand, measurement of rainfall by NEXRAD relies on the raindrop reflectivity which could be affected by factors such as raindrop size and microwave signal reflection by other particles in the atmosphere. Skinner et al. (2008) showed that the best of the two measurement methods is realized by using rain gauge or tipping bucket data to adjust NEXRAD values.

2.3.4 Ground surface potential evapotranspiration

Ground surface reference evapotranspiration (ET_o) was computed from micrometeorological data (i.e., solar radiation, temperature, relative humidity and wind speed) obtained from a Florida Automated

Weather Network (FAWN; http://fawn.ifas.ufl.edu/) station located approximately 10 km northeast of the study site at the Tropical Research and Education Center, Homestead, FL. The American Society of Civil Engineers (ASCE) standardized Penman–Monteith equation was used to estimate ET_o values (ASCE, 2005). We assumed a crop with the following characteristics transpiring at a potential rate: crop height (0.12 m), albedo (0.23), active leaf area index (1.44), and well illuminated leaf stomatal resistance (100.8 s/m). We applied the tool REF-ET (Allen, 2011) to calculate the ASCE standardized ET_o from weather data.

2.3.5 Evaporation from the water table

Flux due to water table evaporation may influence soil and bedrock water contents. Previous studies have shown that when canal influences are negligible, direct evaporation from the water table significantly contributes to water table declines in the Biscayne aquifer (Merrit, 1996; Chin, 2008). Two types of models are available to estimate evaporation from a water table: physically based models and empirically based models. In this study, the latter was used because the former requires detailed data such as coefficient of diffusion of water vapor through the soil and vapor pressure above the soil surface which were not collected. Empirical models simply relate water table evaporation rate to the depth of the water table below the ground surface and are used in groundwater studies (e.g., MODFLOW uses this approach; Chin, 2008). We used a model similar to that proposed by McDonald and Harbaugh (1988) (eq. (5)). Chin (2008) modified eq. (5) and obtained eq. (6) for south Florida conditions.

$$\frac{E}{E_0} = \left(1 - \frac{d}{d_{cr}}\right), \quad d_{cr} = 100*(170 + 8T), \quad d < d_{cr} \tag{5}$$

$$\frac{E}{E_0} = \begin{cases} 1 & d \le d_0 \\ 1 - \dfrac{d - d_0}{d_{cr}} & d_0 < d < d_{cr} \\ 0 & d \ge d_{cr} \end{cases} \tag{6}$$

where E is water table evaporation [mm/day], E_0 (same as ET_o) is the potential evaporation rate at the ground surface [mm/day], d is the depth of the water table below the ground surface [m], d_{cr} is the critical depth below which evaporation ceases [m], T is annual average air temperature [°C] which is approximately 25°C in south Florida, d_0 is water table depth above which water table evaporation proceeds at potential rate i.e., at the rate similar to the ground surface evapotranspiration [m]. Chin (2008) proposed parameters d_0 and d_{cr} in eq. (6) at each observation well can be estimated from the least squares best fit of eq. (7) and the parameters described as eq. (8) and (9).

$$\frac{E}{E_0} = \alpha - \beta\, d \tag{7}$$

$$d_0 = \frac{\alpha - 1}{\beta} \tag{8}$$

$$d_{cr} = \frac{\alpha}{\beta} \tag{9}$$

2.4 Dynamic factor analysis

DFA uses eq. (10) to describes a set of N observed time series (Lütkepohl, 1991; Zuur et al., 2003; Ritter and Muñoz-Carpena, 2006). The goal in DFA is to keep M as small as possible while still obtaining a good model fit. Including relevant explanatory variables helps to reduce some of the unexplained variability in the observed time series.

$$s_n(t) = \sum_{m=1}^{M} \gamma_{m,n}\, \alpha_m(t) + \mu_n + \sum_{k=1}^{K} \beta_{k,n} v_k(t) + \varepsilon_n(t) \tag{10}$$

$$\alpha_m = \alpha_m(t-1) + \eta_m(t) \tag{11}$$

where $s_n(t)$ is a vector containing the set of N time series being modeled (response variables), α_m (t) is a vector containing the common trends (same units as the response variables), $\gamma_{m,n}$ are factor loadings or weighting coefficients that indicate the importance of each of the common trends to each response variable (unitless), μ_n is a constant level parameter for shifting time series up or down, $\nu_k(t)$ is a vector containing explanatory variables, and $\beta_{k,n}$ are weighting coefficients for the explanatory variables (regression parameters) which indicate the relative importance of explanatory variables to each response variable (inverse units to convert $\nu_k(t)$ into response variable units) , and $\varepsilon_n(t)$ and $\eta_m(t)$ are independent, Gaussian distributed noise with zero mean and unknown diagonal covariance matrix. The elements in the covariance matrix represent information that cannot be explained by the common trends or the explanatory variables. The unknown parameters $\gamma_{m,n}$ and μ_n were estimated using the Expectation Maximization (EM) algorithm that is described in Dempster et al. (1977) and Shumway and Stoffer (1982). The common trends in eq. (11) were modeled as a random walk (Harvey, 1989) and were predicted using the Kalman filter and EM algorithms. The regression parameters in eq. (10) are estimated using the same procedure as used in linear regression (Zuur et al., 2003). DFA was implemented using a statistical package called Brodgar Version 2.5.6 (Highland Statistics Ltd., Newburgh, UK).

The results from the DFA were interpreted in terms of the canonical correlations ($\rho_{m,n}$), factor loading ($\gamma_{m,n}$), regression parameters ($\beta_{k,n}$) and agreement between modeled and observed soil and bedrock water contents (i.e., expressed as scaled frequency). The goodness-of-fit between modeled and observed soil and bedrock water contents were quantified using the Nash-Sutcliffe coefficient of efficiency (C_{eff}; Nash and Sutcliffe, 1970), the Akaike's Information Criteria (*AIC*; Akaike, 1974) and the Bayesian information criterion (*BIC*). C_{eff} provides an estimate of how well a model predicts an observed data set, while *AIC* and *BIC* are relative measures of the goodness-of-fit of a statistical model. A model with the C_{eff} closest to 1 and lowest *AIC* and BIC is the preferred DFA model. Cross correlations between the soil and bedrock water content time series and common trends were measured using $\rho_{m,n}$. In our study $\rho_{m,n}$ close to unity implied that the common trend was highly associated with water content time series. Typically canonical correlations are classified as follows: $|\rho_{m,n}|$>0.75, 0.5-0.75, and 0.3-0.5 as high, moderate, and weak correlations, respectively. The influence of the explanatory variables on water content time series were quantified using the magnitude of the $\beta_{k,n}$ coefficients and their associated standard errors which were used with a *t-test* to assess whether the response variable and explanatory variables were significantly related.

DFA was implemented sequentially by varying the number of common trends *M* until a minimum *AIC* and *BIC* and C_{eff} closest to one were achieved (Zuur et al., 2003). After identifying the minimum *M*, different combinations of explanatory variables were introduced into the analysis until a combination of common trends and explanatory variables that resulted in the most parsimonious model with best good-of-fit indicators was achieved. The procedure followed here is similar to that described by Ritter et al. (2009).

2.4.1 Explanatory variables

Soil and bedrock water content time series are autocorrelated (Kaplan and Muñoz-Carpena, 2011) while evapotranspiration and rainfall time series are not. For example, soil and bedrock water contents at time *t* will depend on antecedent soil and bedrock water contents at time *(t-1)* whereas the rainfall today does not depend on rainfall yesterday. Therefore in order to relate the soil and bedrock water content time series and evapotranspiration and rainfall time series, we calculated a new variable called net cumulative recharge (R_{net}) using eq. 12.

$$R_{net} = \sum_{t=1}^{t} P_t - \sum_{t=1}^{t} E_{o\,t} \tag{12}$$

where P_t is the total rainfall for day t (mm) and E_{ot} is the potential evapotranspiration on day t (mm/day). Cumulative water table evaporation was also used instead of daily values. To minimize multi-colinearity of explanatory variables, we used mean water table elevation instead of water table elevation at each well. Before proceeding with the DFA, multi-colinearity of explanatory variables was quantified by computing variance inflation factors (VIFs) for each explanatory variable (Zuur et al., 2007).

2.5 Simple predictive regression model for soil water content

The simple regression model was developed from a DFA model having the minimum number of common trends required to explain underlying common patterns in the eleven time series and explanatory variables with significant influence on modeled soil water and bedrock water content time series. To enable practical use of the simple model, DFA was performed again for the identified model using non-normalized/non-standardized time series. After estimating the parameters through DFA the common trends were ignored in the model to derive a simple expression relating identified significant explanatory variables and soil and bedrock water contents. The period from August 25, 2010 to December 2011 was used to develop the regression model while the data from December 01, 2011 to June 30, 2012 was used to validate the new simple model. The developed simple model was then applied to predict the impact of a 6, 9 and 12 cm increase in canal stage on soil and bedrock water contents at the study sites.

3. Results and discussion

3.1 Visual exploratory analysis of experimental time series

Visual inspection of soil and bedrock water content time series expressed as SF indicates that there were some common patterns in the temporal variation of soil and bedrock water contents at the three sites (T500, T1000 and T2000) along the transect perpendicular and east of canal C111. From February 2011 to July 2011, soil and bedrock water contents gradually decreased at all monitoring elevations and all sites (Fig. 3). The gradual decrease in soil and bedrock water contents corresponded to the decline in canal stage and water table elevation (Fig. 4). The period from April to August was characterized by pronounced drying and wetting cycles at all sites. The wetting or spikes in soil and bedrock water contents in this period correspond to the start of the rains while the drying cycles correspond to the increasing potential evapotranspiration during the same period (Fig. 4). The period from late March to July corresponds to the end of the growing season and beginning of the wet season. From August 2011 to February 2012, soil and bedrock water increased corresponding to stage operation criteria within the canal network that enhances water storage in the system.

However, there were observed differences in temporal soil and bedrock water variability at the three monitoring sites along the transect. Site T500 which is the shallowest and closest to the canal exhibited lack of temporal variation in bedrock water content at elevations less than 0.9 m NGVD29 while soil water content at 1.0 m NGVD29 exhibited temporal variation in the same period probably due to irrigation during the growing season. Site T1000 (i.e., approximately 1000 m from canal C111) exhibited the least increase in water content between March 2011 and June 2012. Unlike sites T500 and T2000, the trends in soil and bedrock water contents at T1000 were not identical to the temporal variation in canal stage or water table elevation suggesting micro-topography within the field might be affecting soil and bedrock water contents since this site had the highest elevation along the transect (Fig. 2). At site T2000 (i.e., approximately 2000 m from canal C111), soil and bedrock water contents for the periods between August 2010 to March 2011 and August 2011 to February 2012 were similar characterized by small temporal variation similar to those exhibited at site T500. Sites T500 and T2000 have very similar elevation (1.1 and 1.2 m NGVD29 respectively) implying that topography or ground surface elevation might exert a stronger influence on temporal variation of soil and bedrock water contents compared to distance from the canal. Differences also existed at the different monitoring elevations with bedrock water content generally higher at the lowest elevation at each site. Other reasons for observed differences in water content at the different sites could be a combination of several factors such as differences in soil surface conditions, soil and limestone bedrock heterogeneity (specifically differences in soil water retention and unsaturated hydraulic conductivity) and differences in the environments surrounding the EnviroScan access tubes.

All the hydrologic variables monitored (Fig. 4) exhibited seasonal variations with rainfall increasing during wet season (May to October) resulting in increased water table elevation and canal stage. ET_o also increased during the wet season. In turn, decreased depth to water table and increased ET_o resulted in increased *E* (Fig. 4).The water table evaporation parameters for eq. (6) were computed following the procedure described by Chin (2009) in which steady declines in water table elevation particularly in the dry season when canal stage was maintained relatively constant are assumed to be caused by water table evaporation. Using data from a total of six wells (i.e., the 3 wells along transect T and 3 additional wells approximately 1 km north of the transect) within the vicinity of the study area, we obtained an average critical depth of 1.94 m which is within the range of 1.5 to 2.9 earlier reported by Chin (2009). We obtained a value of 0.59 m for the depth above which water table evaporation proceeds at the potential rate which is approximately half the average value of 1.4 reported by Chin (2009). The water table elevations at the three monitoring sites were very similar and also corresponded to the temporal variations in canal stage on the tail water side of the spillway at S177.

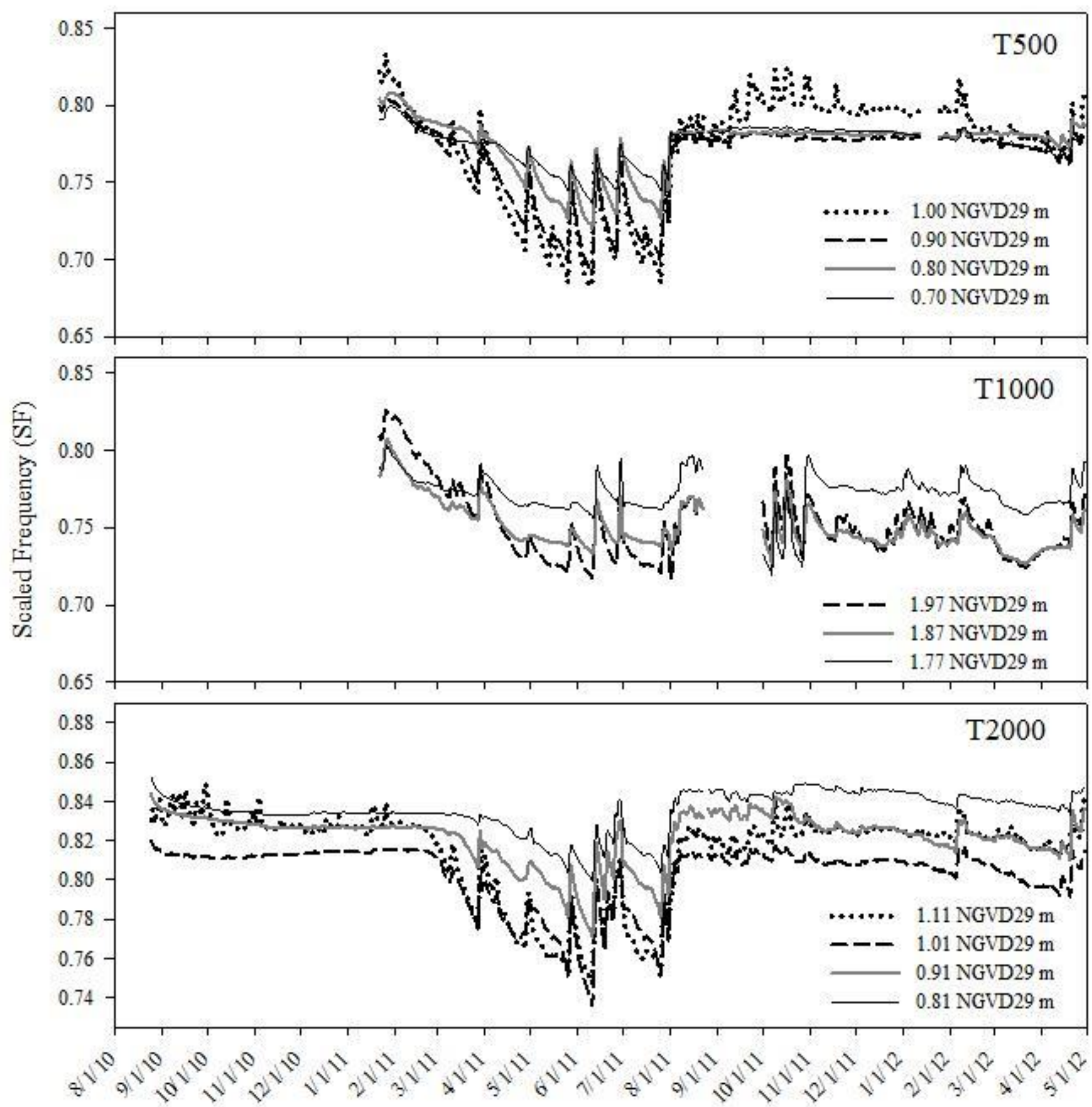

Figure 3. Temporal variation in scaled frequency (i.e., soil and bedrock water contents) at three sites (i.e., T500, T1000 and T2000 with soil and bedrock water contents monitored at different elevations using EnviroScan probes) along a transect perpendicular to C111 on the tail water side of the spillway at structure S177 during the period August 2010 to June 2012.

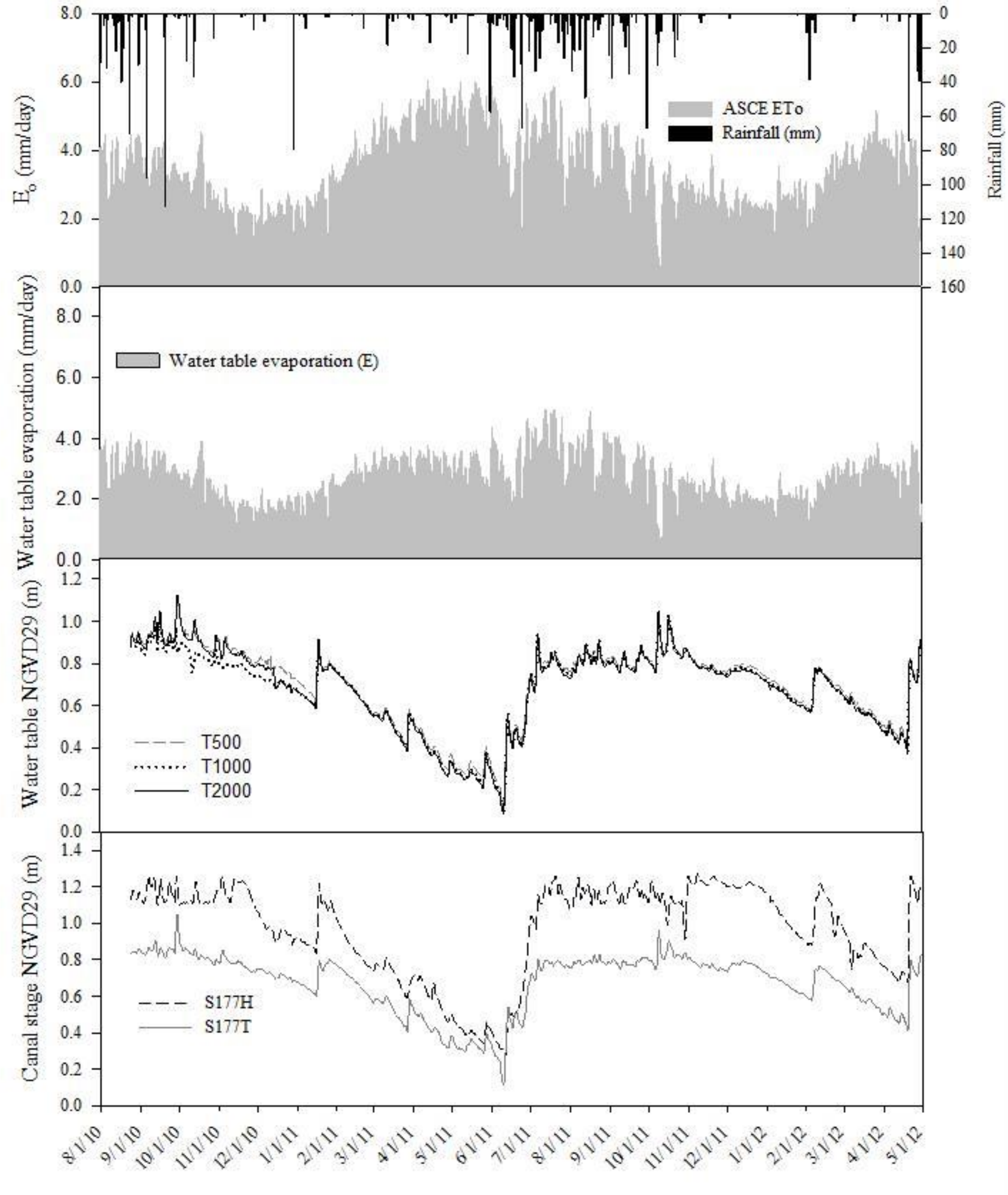

Figure 4. Temporal variation in hydrologic factors evaluated for their influence on soil and bedrock water contents at the study site during the period August 2010 to June 2012.

3.2 Response and explanatory variables

Visual inspection indicated that seasonality affects temporal variation of both response variables (i.e., soil and bedrock water contents at different elevations) and explanatory variables (i.e., ET_o, rainfall P, water table elevation, E and canal stage). We attempted to remove seasonality effects through seasonal standardization following procedures described by Salas (1993), but this approach was abandoned since it resulted in poor model fit compared to the models in which seasonal effects were assumed to be masked in the common trends (i.e., average $C_{eff} < 0.7$ and $C_{eff} > 0.9$, respectively). The poor model fit could be attributed to loss of information resulting from seasonal standardization. Ritter et al. (2009) also reported improved DFA model fit after back transforming refractive index data from a TDR as a surrogate for soil water content compared to seasonally standardized refractive index.

To facilitate interpretation of factor loadings and comparison of regression parameters as suggested by Zuur et al. (2004), all the time series were normalized. Therefore, the DFA results presented in reference to objective 1 are based on normalized time series data. Prior to performing the DFA, multicollinearity in explanatory variables was quantified by calculating Variance Inflation Factor (VIFs) for each explanatory variable. Threshold VIF of 5 was set as the highest, high values of VIF indicate multicollinearity in the explanatory variables which makes interpretation of regression results difficult (Ritter et al., 2009). As expected there was high multi-colinearity between water table elevation time series for different wells (VIFs > 30), but this was considerably reduced when mean water table elevation at the three sites was used instead (i.e., VIFs < 2). There was also high multi-colinearity between headwater and tail water canal stages at S177 (VIFs > 8) implying that these two time series could not be used as explanatory variables in the same DFA model. Mean water table elevation was also correlated to canal stage S177 (VIFs >10) probably due to the high hydraulic connectivity between C111 and Biscayne aquifer. The correlation coefficient between canal stage and water table elevation time series was greater than 0.9.

3.3 Common trends

We developed the DFA model by exploring common trends and explanatory variables in relation to the 11 observed water content time series. Results of the DFA model selection are summarized in Table 1. We used the *AIC*, the *BIC* (which penalizes more strongly for over parameterization than the *AIC*) and the C_{eff} statistic for deciding which of the DFA models with zero explanatory variables best described the response time series. Ten was the maximum number of common trends used to describe common variability in the 11 response water content time series. However, the goal of DFA is to minimize the number of common trends while maintaining a good model fit. Several models consisting of fewer numbers of common trends and noise were tested and model 4 with five common trends was determined to be the model with the minimum number of common trends required to describe the 11 response time series. Model 4 was selected since using M>5 resulted in negligible improvement in model goodness-of-fit measures while increasing the number of parameters to be interpreted. The three common trends with high ($\rho_{m,n}$ |>0.75) to moderate (0.5< $\rho_{m,n}$ <0.75) canonical correlations particularly at sites T500 and T2000 are shown in Fig. 5. Common trends 2 and 3 exhibited minor cross correlation with water content time series as measured by $\rho_{m,n} < 0.5$ at all the sites and in the interest of brevity are not presented.

Visually, the unexplained variation in soil and bedrock water contents described by the common trends in Fig. 5 is similar to the seasonal variation of soil and bedrock water contents at sites T500, T1000 and T2000 for the period August 2010 to August 2011. There was greater uncertainty as shown by a large (95%) confidence interval from August 25, 2010 to January 21, 2011 which is due to missing data for sites T500 and T1000 during this period. The first common trend exhibited high positive ($|\rho_{1,n}| \geq 0.75$) correlation with soil and bedrock water content time series at sites T500 and T2000 with low surface elevation (1.1 and 1.2 m NGVD29, respectively) compared to the moderate to weak correlation at site T1000 with ground surface elevation of 2.17 m NGVD29. Indicating that in addition to other factors, such as irrigation during the growing season, micro-topography within the field influences temporal variations in soil water content as it governs the effect exerted by the water table.

Table 1. Dynamic Factor Analysis (DFA) models tested based on the following goodness-of-fit measures: AIC, BIC and C_{eff}

Model	No. of common trends	Explanatory variables	No. of parameters	AIC[1]	BIC[2]	C_{eff} [3]
		Step I (DFA model with K=0)				
1	2	None	98	-2690.50	-2041.75	0.68
2	3	None	107	-4654.23	-3945.90	0.84
3	4	None	115	-5830.21	-5068.92	0.88
4	**5**	**None**	**122**	**-6901.47**	**-6093.84**	**0.97**
5	6	None	128	-7028.76	-6181.40	0.97
6	8	None	137	-7263.94	-6357.01	0.97
		Step II (DFA model with K>0)				
7	5	R_{net}[4],	133	-7018.644	-6138.193	0.97
8	5	R_{net}, E[5]	144	-7797.525	-6844.255	0.98
9	5	S177T[6]	133	-7340.981	-6460.530	0.97
10	5	S177T, R_{net}	144	-7542.680	-6589.410	0.97
11	5	R_{net}, E, MWT[7]	155	-8052.436	-7026.346	0.98
12	5	MWT, R_{net}	144	-7444.030	-6490.761	0.97
13	**5**	**R_{net}, E, S177T**	**155**	**-7922.346**	**-6896.257**	**0.98**

[1]AIC Akaike information criterion
[2]BIC Bayesian Information Criterion
[3]C_{eff} Nash-Sutcliffe coefficient calculated based all the nine observed time series
[4]Cumulative net surface recharge
[5]R_{net} Cumulative water at table evaporation
[6]S177T Canal stage in C111
[7]MWT Mean water table evaporation

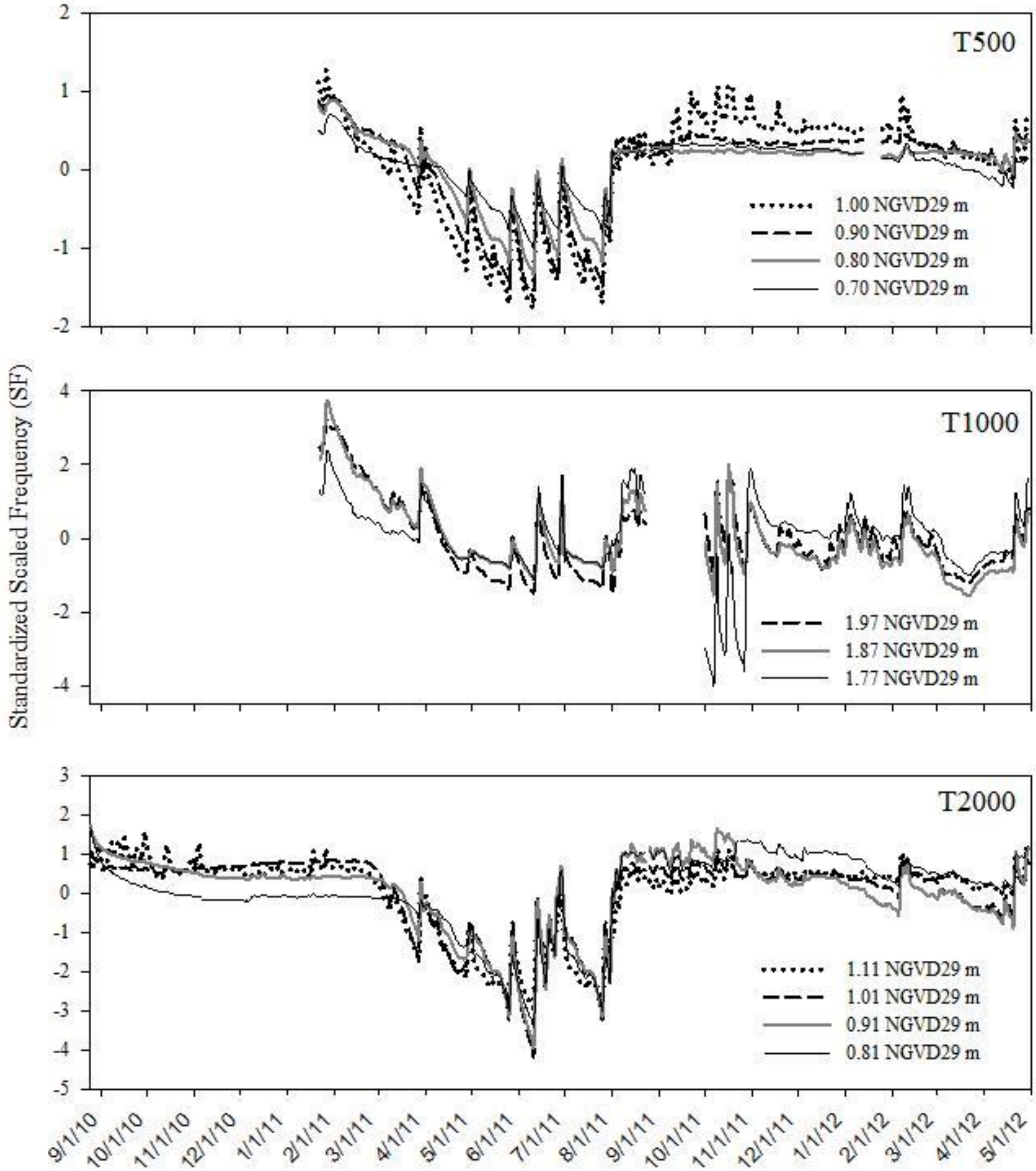

Figure 5. Common trends with 95% confidence interval describing unexplained temporal variation in scaled frequency as a surrogate for soil and bedrock water content and the canonical correlation for quantifying the correlation between water time series and the common trends, in the nomenclature for site

names the number represents distance from the canal in m, and the numbers in the parenthesis represent elevation NGVD 29 m.

3.4 Relative contribution of explanatory variables

Introducing net surface recharge, water table evaporation, and mean water table elevation or C111 canal stage to model 4 resulted in the best models (11 and 13). Inclusion of explanatory variables in the DFA model also produced regression parameters ($\beta_{k,n}$) and since response and explanatory variables were normalized, the regression parameters were used to quantify the relative influence of each explanatory variable on the modeled soil and bedrock water content time series. It is worth noting that substituting mean water table elevation in model 11 with canal stage as in model 13 resulted in *AIC* and *BIC* that were not substantially different and similar goodness-of-fit indicator (Table 1). Since part of the motivation for this research was to assess the effect of canal stage management on soil and bedrock water contents, further analysis was made on model 13 because canal stage data have a more consistent record compared to water table elevation data. At the study site, canal stage can be used as a good approximation to water table elevation due to the high permeability of the aquifer.

Model 13 fitted plots are shown in Figs. 6 to 8; these figures indicate that DFA modeling was successfully applied to describe temporal variations in soil and bedrock water contents at all three monitoring sites and elevations ($C_{eff} > 0.9$). Results in Table 2 indicate that net surface recharge (R_{net}) had a significant influence (*t value >2*) on the temporal variation of soil and bedrock water contents at sites T500, T1000, and T2000 but was not significant at lower elevations at sites T1000 and T2000 as shown (*t value <2*). The significance of R_{net} could be attributed to rainfall (*P*) patterns in the study area in which two thirds of the *P* was received in the wet season (SFWMD, 2011) and these large amounts of net water input to the vadose zone are sufficient to maintain soil and limestone bedrock near saturation, while absence of *P* in the dry season was responsible for the dry conditions. Lack of significance at lower elevations at sites T1000 and T2000 could be attributed to heterogeneity in soils and bedrock (e.g., differences in hydraulic conductivity), and differences in surface cover which influence ET_o.

Water table evaporation was found to not significantly influence temporal variation of soil and bedrock water contents (*t value <2*) at all the sites monitored. The non-significant effect of water table evaporation on soil and bedrock water content could be attributed to the fact that there is sufficient water for evaporation due to the shallow water table. However, the negative effect was stronger at site T1000, the negative effect is due to the fact that water table evaporation is a net loss from the vadose zone system. The small positive water table evaporation regression coefficient at T1000 and T2000 (Table 2) could be attributed to computational numerical errors. These results are worth highlighting given the fact that meteorological based methods for estimating ET_o like Penman Monteith equation are criticized for ignoring evaporation from the shallow water table meaning they might under estimate total ET_o losses. These observations could be attributed to that fact ET_o in such cases is not limited by water availability but by available energy only.

C111 canal stage on the tail water side at the S177 spillway (Fig. 1) had the strongest influence on soil and bedrock water content temporal variations (*t value >7*) for most sites. This finding is significant because it confirms the hypotheses that the shallow water table and canal stage are highly connected and that canal stage can be used to predict soil water content at a given location. From a hydrologic perspective, these results were expected because in this case canal stage is used an approximation for the shallow water table which serves as the lower boundary condition for the vadose zone and therefore regulates available storage during the rainy season. Based on the relative magnitudes of the regression coefficients (Table 2), the overall contribution of canal stage on the respective soil and bedrock water content time series is higher than that of net recharge.

The factor loadings ($\gamma_{1,n}$) for the five common trends are shown in Table 2, these represent the influence of each common trend on the modeled soil and bedrock water content time series at the different monitoring sites and elevations. Since the time series in the DFA were normalized, the coefficients $\beta_{k,n}$ and $\gamma_{1,n}$ can be compared (Zuur and Pierce, 2004). The results indicate that trend 1 was very critical for describing unexplained variation in soil water dynamics at site T2000, while common trends 2 and to a lesser extent 3 were more critical for describing unexplained variation in soil water content at site T1000. Site T500 was sufficiently described by the explanatory variables and constant level parameters given their magnitudes were larger compared to the $\gamma_{1,n}$. Trends 4 and 5 had minor effects at all the monitoring sites.

Overall at all the sites, compared to regression coefficients and the constant level parameters, common trends had less influence on soil and bedrock water dynamics. However, since the values of the factor loadings are not zero (i.e., they account for some unexplained variability) especially at T2000 and site T1000, this implies that the information provided by the hydrologic variables used as the explanatory variables in the DFA models only account for part of the unexplained variability in the temporal variation of the soil and bedrock water contents. Other information such as irrigation, differences in soil surface conditions, differences in the environment surrounding the EnviroScan access tube, and variation in soil hydraulic properties not considered in this study might account for part of the remaining unexplained variability.

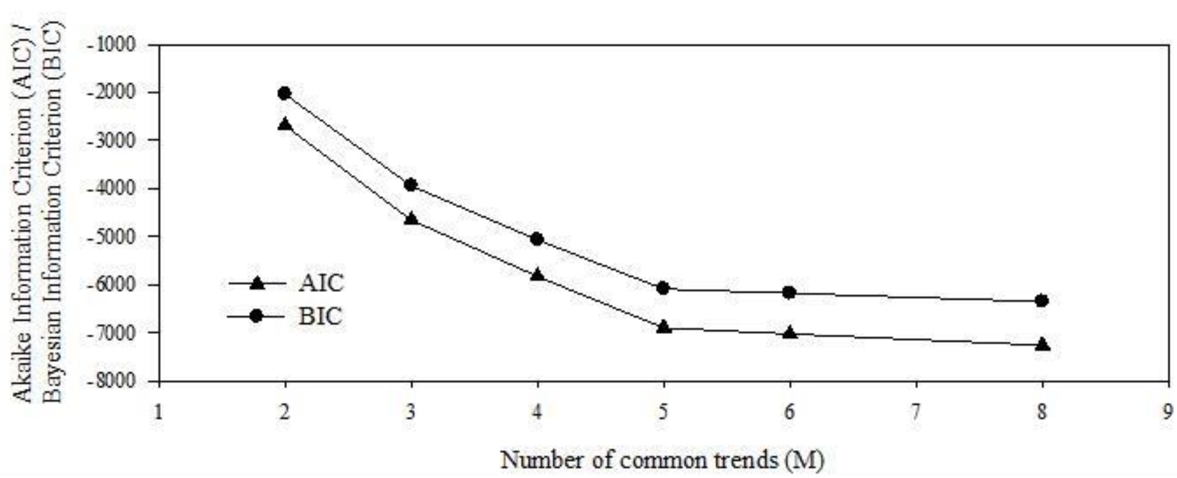

Figure 6. Fitted Dynamic Factor Model (DFM) and observed temporal variation in scaled frequency (used as a surrogate for soil and rock water) in gravely loam soils and limestone bedrock at a site located 500 m along a transect from C111 and the numbers in the parentheses indicate elevations.

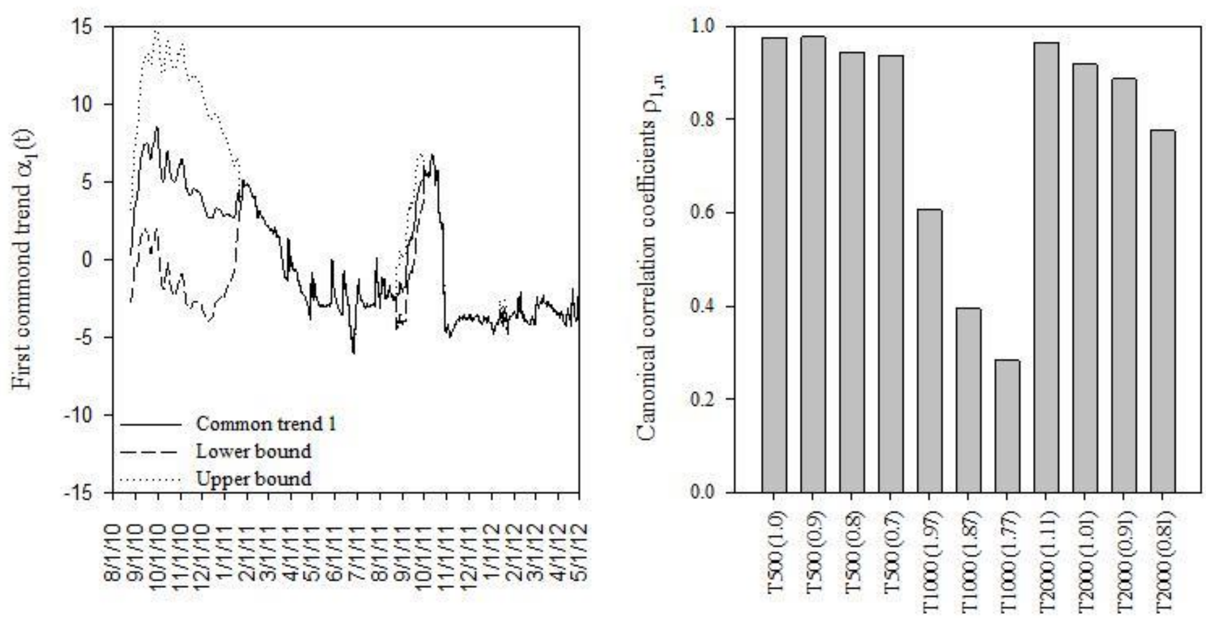

Figure 7. Fitted Dynamic Factor Model (DFM) and observed temporal variation in scaled frequency (used as a surrogate for soil and rock water) in gravely loam soils and limestone bedrock at a site located 1000 m along a transect from C111 and the numbers in the parentheses indicate elevations.

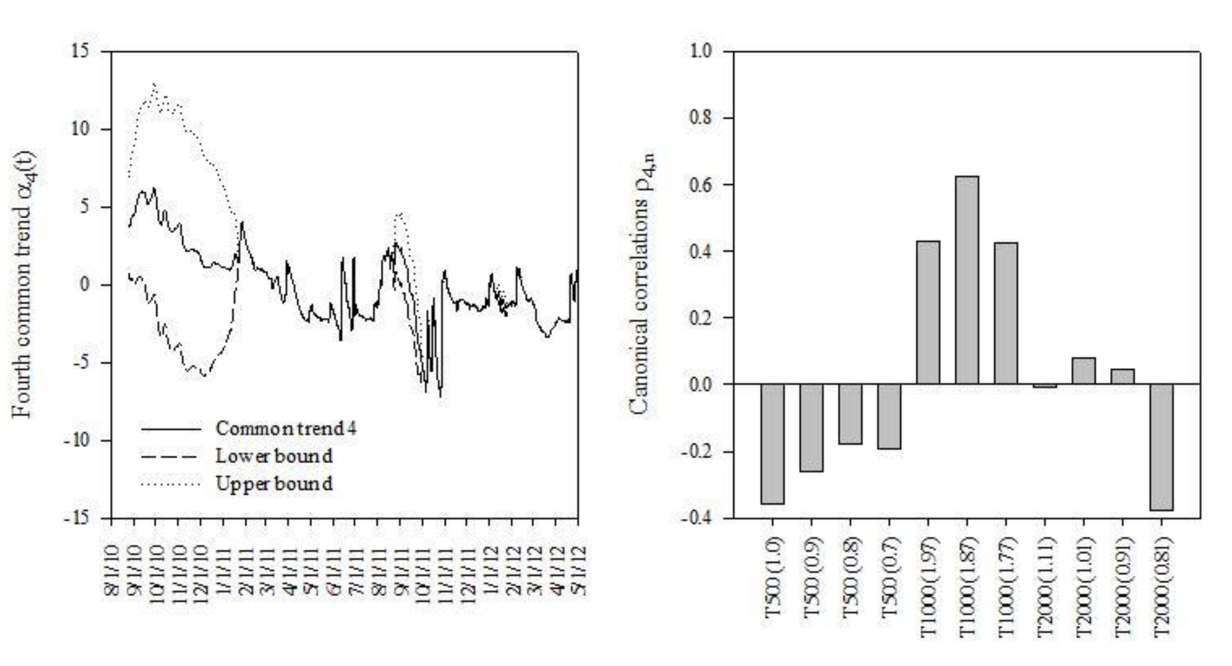

Figure 8. Fitted Dynamic Factor Model (DFM) and observed temporal variation in scaled frequency (used as a surrogate for soil and rock water) in gravely loam soils and limestone bedrock at a site located 2000 m along a transect from C111 and the numbers in the parentheses indicate elevations.

Table 2. Dynamic Factor Analysis results for model 13 with 5 common trends and 3 explanatory variables

s_n	$\gamma_{1,n}$	$\gamma_{2,n}$	$\gamma_{3,n}$	$\gamma_{4,n}$	$\gamma_{5,n}$	μ_n	β_{Rnet}	β_E	$\beta_{C111stage}$	C_{eff}
[1]T500 (1.0)	0.05	0.02	0.04	-0.02	-0.03	0.28 (0.6)	0.34 (6.9)	0.00 (0.0)	0.24 (8.8)	0.93
T500 (0.9)	0.05	0.06	0.03	-0.04	-0.05	0.37 (0.5)	0.24 (3.5)	-0.14 (-0.3)	0.29 (8.3)	0.94
T500 (0.8)	0.04	0.06	0.03	-0.03	-0.04	0.34 (0.6)	0.20 (3.2)	-0.17 (-0.5)	0.22 (7.5)	0.90
T500 (0.7)	0.03	0.02	0.01	0.00	-0.01	0.13 (0.6)	0.18 (7.1)	-0.09 (-0.7)	0.13 (9.1)	0.90
T1000 (1.97)	0.04	0.16	0.13	-0.02	0.00	0.95 (0.9)	0.47 (3.1)	-0.53 (-0.7)	0.62 (8.7)	0.85
T1000 (1.87)	0.04	0.20	0.11	0.01	0.01	0.82 (0.8)	0.38 (2.1)	-0.61 (-0.8)	0.70 (8.5)	0.81
T1000 (1.77)	0.01	0.50	0.01	0.00	0.00	0.00 (0.0)	0.44 (1.1)	0.23 (0.1)	0.77 (4.6)	0.67
T2000 (1.11)	0.10	0.04	0.06	-0.07	0.01	0.07 (0.1)	0.13 (2.0)	0.05 (0.1)	0.50 (11.6)	0.99
T2000	0.13	0.05	0.06	-0.02	0.06	-0.09 (-0.1)	0.03 (0.3)	-0.03 (0.0)	0.68 (13.2)	0.90

(1.01)										
T2000 (0.91)	0.17	0.03	0.06	0.01	-0.01	-0.12 (-0.1)	0.05 (0.4)	-0.22 (-0.3)	0.71 (12.4)	0.93
T2000 (0.81)	0.16	0.04	-0.03	-0.02	-0.02	-0.31 (-0.3)	0.08 (0.8)	0.02 (0.0)	0.46 (8.8)	0.96

γ Factor loading corresponding to common trend 1to 5 and observation, *n*= 1,2, 3…., 11
µ Constant level parameter in dynamic factor model with associated *t-value* in parenthesis
ß Regression parameter corresponding to the 3 explanatory variables (net recharge [R_{net}], water table evaporation [E], and canal stage in C111 [C111stage]) with associated t-value in parenthesis
C_{eff} is Nash-Sutcliffe coefficient
[1]Site name nomenclature; T is refers to transect name T, number refers to distance from canal and number in parentheses refers to elevation NGVD29 m
n number of observations

3.5 Predicting soil and bedrock water contents using a simplified dynamic factor analysis based model

To enable practical application of the DFA model, the common trends and two of the exploratory variables included in model 13 were used in a new DFA model with non-standardized time series. This new model was referred to as model 14. To further simplify model 14, we ignored the common trends to derive a simple model that predicts soil and bedrock water contents as function of net recharge and canal stage expressed as eq. 13

$$SF(X,Z,t) = \beta_{Rnet}(X,Z)\ R_{net}(t) + \beta_{C111}(X,Z)\, S177T(t) + \mu(X,Z) \qquad (13)$$

where *SF(X,Z,t)* is the SF at distance *X* from the canal, at elevation *Z*, and time *t*, other terms in are previously described and varies with elevation and distance from the canal. The coefficients for eq. 13 at all the sites and monitoring elevations are obtained from Table 3. The C_{eff} in Table 3 are calculated based on eq. 13 with common trends removed. As expected, performance of the simple model (eq. 13) was lower as shown by the reduction in C_{eff} (Table 3 and Figs. 9 to 10) compared to the DFA models that include common trends particularly for site T1000.

Since factor loadings are not zero for all the trends (Table 3), this suggests that the explanatory variables (net recharge and canal stage) used in the DFA model are not sufficient to explain all the observed variations in the soil and bedrock water content time series. This is particularly true at site

T1000 which is affected by 4 out of the 5 common trends. Common trend number 2 appears to affect all the sites, it probably masks common variation such seasonal changes in rainfall, evapotranspiration and canal stage. Other common trends had minor effects at sites at all the other sites particularly at site T1000. The difference in response at site T1000 could be attributed to differences in elevation as shown in Fig. 2, site T1000 has a higher surface elevation and hence larger depth to water table.

The results in Table 3 also underscore the point that the effect of canal stage is stronger at low elevation sites T500 and T2000 compared to T1000. Thus, proper interpretation of modeling results in this area requires accurate quantification of micro-topography. Model performance ranged from good at sites T500 and T2000 to poor at site T1000 with root mean square error (RMSE) ranging from 0.005 to 0.01. Figs. 9 to 10 show model performance during the calibration and validation periods, after removing the common trends, it can be seen that the simple model misses the peaks but is able to generally predict the temporal variation in soil and rock water content. The simple model (eq. 13) could be improved by using location specific water table elevation since canal stage is simply a good approximation of the mean water table elevation. Another simple sigmoidal regression model to predict soil and bedrock water contents from canal stage proposed by Kaplan et al. (2010a) was tried but later abandoned due to lower C_{eff} (i.e., averaging 0.2). This approach is based on the physical concept of drain to equilibrium. However, for our study site this condition was hard to achieve since during the dry season irrigation was taking place and in the rainy season there was frequent rainfall hence by removing data points corresponding to rainfall or irrigation, very few data points were left to develop a useful sigmoidal model for predicting soil and bedrock water content from canal stage.

Table 3. Dynamic Factor Analysis results for model 14 with 5 common trends and 2 explanatory variables implemented with non-standardized time series

s_n	$\gamma_{1,n}$	$\gamma_{2,n}$	$\gamma_{3,n}$	$\gamma_{4,n}$	$\gamma_{5,n}$	μ_n	β_{Rnet}	$\beta_{c111stage}$	C_{eff}	C_{eff}
[1]T500 (1.0)	-0.003	0.000	0.000	0.000	0.000	0.72	0.14	0.06	0.73	0.70
T500 (0.9)	-0.001	-0.004	0.000	0.000	0.000	0.72	0.11	0.04	0.61	0.62

T500 (0.8)	-0.001	-0.004	0.000	0.000	0.000	0.75	0.09	0.02	0.51	0.56
T500 (0.7)	-0.001	-0.002	0.002	0.000	0.000	0.76	0.07	0.01	0.81	0.74
T1000 (1.97)	0.003	-0.005	-0.002	0.000	0.001	0.73	0.10	0.02	0.61	0.15
T1000 (1.87)	0.002	-0.003	-0.001	0.000	0.001	0.74	0.05	0.01	0.51	0.13
T1000 (1.77)	0.001	-0.003	0.001	-0.002	0.000	0.77	0.02	0.00	0.25	0.11
T2000 (1.11)	0.000	-0.003	-0.002	0.000	0.000	0.76	0.08	0.06	0.70	0.61
T2000 (1.01)	0.000	-0.003	0.000	0.000	0.001	0.76	0.05	0.05	0.60	0.67
T2000 (0.91)	0.000	-0.002	0.000	0.000	0.001	0.77	0.03	0.04	0.67	0.63
T2000 (0.81)	-0.001	-0.001	0.000	0.000	0.001	0.80	0.02	0.02	0.65	0.61

γ Factor loading in the dynamic factor model
μ Constant level parameter in dynamic factor model
ß Regression parameter corresponding to the 2 explanatory variables (net recharge [R_{net}], and canal stage in C111 [C111stage])
[1]Nash-Sutcliffe coefficient are calculated after ignoring common trends
[2]Nash-Sutcliffe coefficient during validation

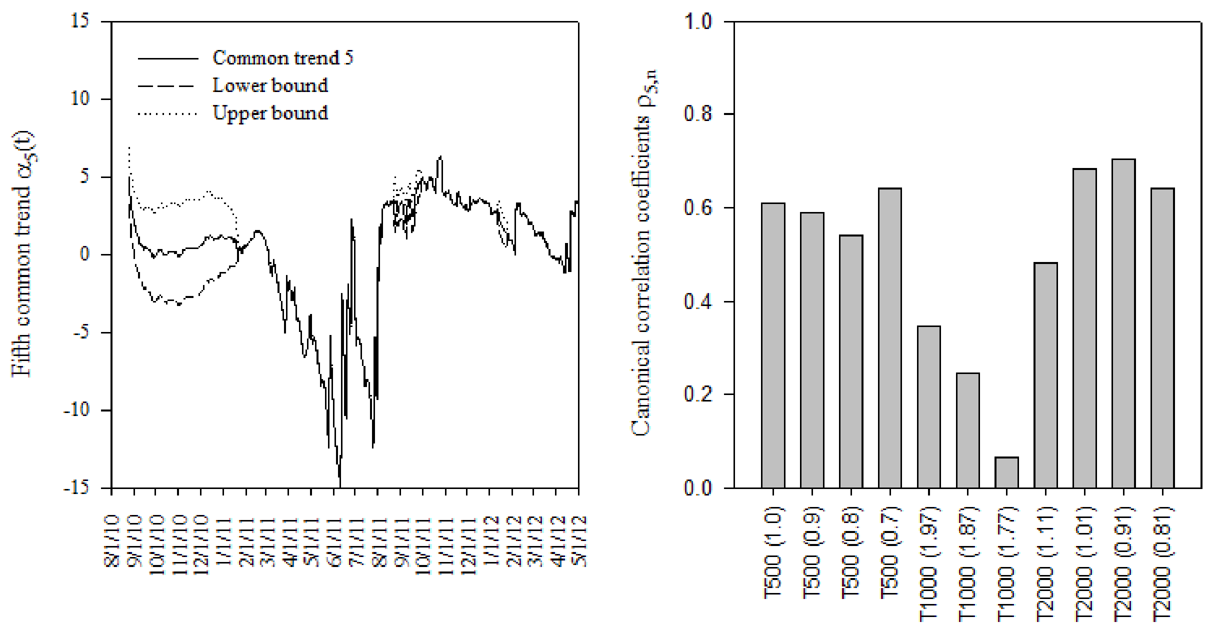

Figure 9. Performance of a simple model for predicting scaled frequency (used as a surrogate for soil and bedrock water content) as a function of canal stage and net recharge at specific elevations in parentheses NGVD29 at a site located 500 m along transect T from C111.

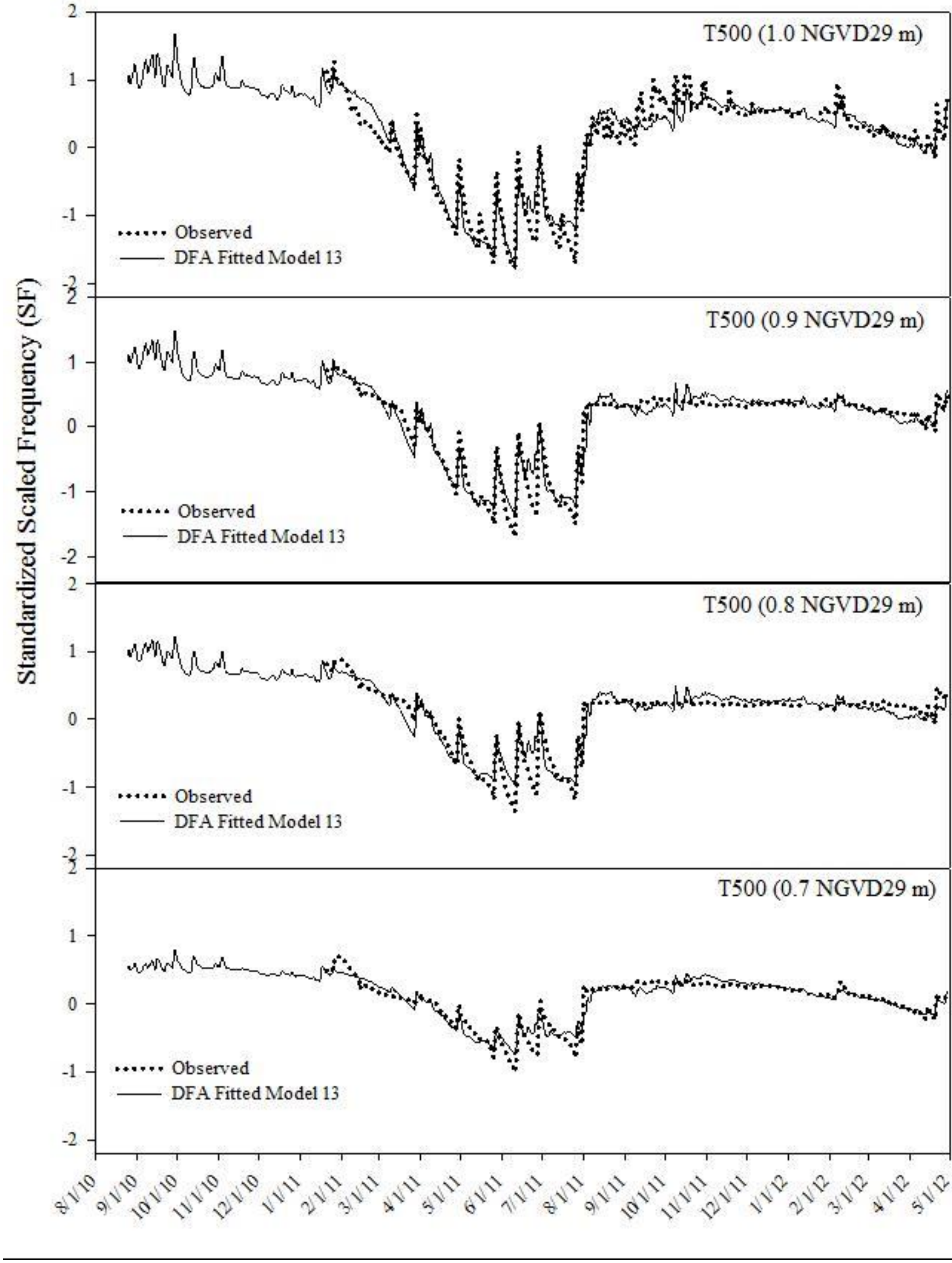
T500 (1.0 NGVD29 m)
T500 (0.9 NGVD29 m)
T500 (0.8 NGVD29 m)
T500 (0.7 NGVD29 m)
Observed
DFA Fitted Model 13
Standardized Scaled Frequency (SF)
2
1
0
-1
-2
8/1/10
9/1/10
10/1/10
11/1/10
12/1/10
1/1/11
2/1/11
3/1/11
4/1/11
5/1/11
6/1/11
7/1/11
8/1/11
9/1/11
10/1/11
11/1/11
12/1/11
1/1/12
2/1/12
3/1/12
4/1/12
5/1/12

Figure 10. Performance of a simple model for predicting scaled frequency (used as a surrogate for soil and bedrock water content) as a function of canal stage and net recharge at specific elevations in parentheses NGVD29 at a site located 2000 m along transect T from C111.

3.6 Assessing the impact of proposed operational changes in C111 canal stage management on soil and bedrock water contents

The low lying agricultural areas east of canal C111 are anticipated to experience the greatest impact from the proposed changes in C111 stage operation (i.e., canal stage increases of 6, 9, and 12 cm); a simple DFA based regression model eq. 13 was proposed to predict the soil and bedrock water contents as a function of canal stage. We considered the period from January 01, 2012, to June 30, 2012 for the analysis. Increases in canal stage were computed by simply adding the proposed incremental rises in canal stage to the daily canal stage recorded at S177T while P and ET_o from the original dataset were not changed.

The results from using this simplified DFA based model (Figs. 11 and 12) indicate that the proposed increases in canal stage were predicted to have changes in daily mean SF for the study period (i.e., which is used as a surrogate for soil and bedrock water contents) of <1% at all sites and all elevations monitored. The range in daily SF differences was 0.065 to -0.024 and 0.075 to -0.041 at sites T500 and T2000 respectively, which indicates that the simple model over predicted and under predicted SF on certain days during the study period. However, note that the daily differences in SF are not substantially large, this may be attributed to already high values of soil and bedrock water contents observed in the area. On an event basis the potential to flood or saturate the root zone would depend on the size of the storm and storm contingency planning for lowering of canal stage in anticipation of heavy storms. Since we showed using DFA that soil and bedrock water contents were significantly affected by canal stage and net recharge.

The simple model used in this evaluation was more accurate at sites T500 and T2000 and therefore results at these two sites would be considered with less uncertainty. Soil and bedrock water responses to incremental raises in canal stage were not computed for site T1000 since results at this site would be considered less accurate (greater uncertainty) because model performance was very poor at this site. Figs. 11 and 12 show that changes in soil and bedrock water contents were more noticeable at the highest elevation. However, at the lowest elevations monitored the difference between mean SF before and after all increments was zero at T500. These observations could be attributed to the fact that low elevation sites are normally close to saturation. For example, at site T500 (0.7) when water elevation was above the sensor (implying saturated conditions), SF was recorded as 0.786 compared to average SF of 0.775 for the study period meaning small changes in water table may not result in substantial changes in soil water content since the pores are already near saturation.

It is worth noting that the simple model developed above should be applied with the following limitations in mind. The model does not account for water input from irrigation and therefore would under predict soil and bedrock water content during the growing season, the model also uses canal stage as an approximation for water table elevation at a specific location although the two are usually close there may be deviations especially after large rainfall events, it ignores water content drivers that were masked in the common trends, and lastly the simple model ignores the effect of *E* which might vary based on micro-topography within the field as well as differences in land surface cover conditions. Finally, although the simplified DFA based model is empirical in nature, the results suggest it can be used as a preliminary tool to relate the potential impacts of surface water management decisions on soil and bedrock water contents in low lying farmlands adjacent to canal C111. This is because during the duration of the study, we able to capture a wide range of variation in canal stage and water table elevation e.g., on June 10, 2011 we recorded canal stage and groundwater table elevation of 0.14 m NGV29 which is lower than the optimum design stage of 0.6 m for the reach of C111 between S 18C and S177 under current canal stage operational criteria. During the summer of 2011 (on October 09, 2011) we recorded canal stage and groundwater levels as high as 0.9 and 1.02 m NGVD29 which is close to the level supposed to trigger the spillway to open at S177 under current operational criteria.

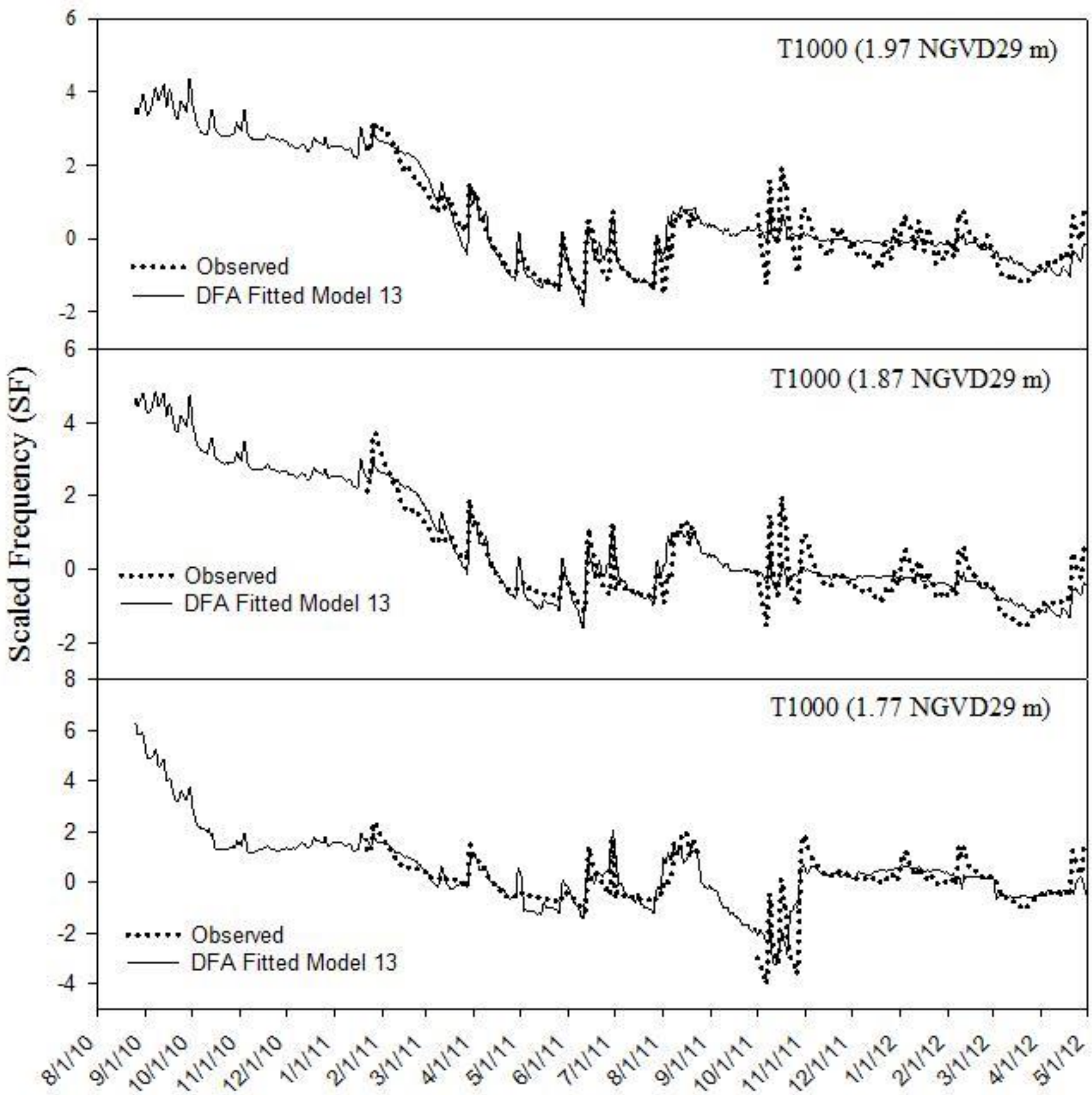

Figure 11. Boxplots showing soil and rock water content as measured using scaled frequency at site T500 before and after 6, 9 and 12 cm increase canal at structure S18C along C111.

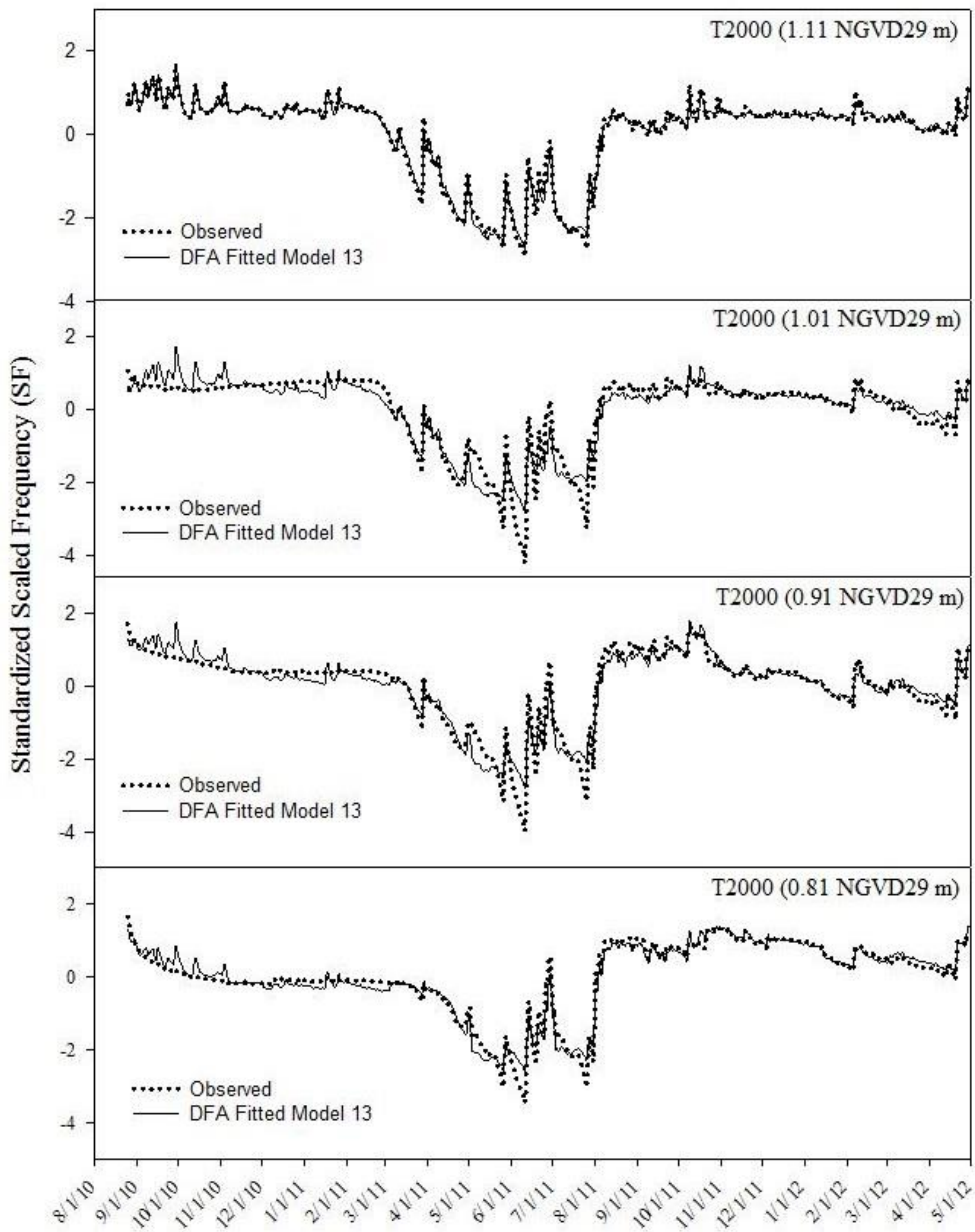

Figure 12. Boxplots showing soil and rock water content as measured using scaled frequency at site T2000 before and after 6, 9 and 12 cm increase canal at structure S18C along C111.

4.0 Summary and Conclusions

The response of soil and bedrock water contents to incremental raises in canal stage proposed under the C111 spreader canal project whose goal is to restore the hydrology of ENP while maintaining flood protection in the adjacent agricultural areas was investigated using DFA. The study objectives were to use DFA to identify the important factors driving temporal variation in soil and bedrock water content above the shallow water table at the study site, develop a simple model for predicting soil water content as a function of canal stage and assess the effect of the proposed incremental raises in canal stage on soil and bedrock water contents. Five was the minimum number of common trends required to account for the unexplained variation in the eleven observed soil and bedrock water content time series while producing an acceptable model fit. Introduction of explanatory variables i.e., net recharge, water table evaporation, and canal stage or water table elevation to the DFA model resulted in lowering AIC and BIC values while C_{eff} values did not substantially change. Evaluation of the regression coefficients indicated that net recharge and canal stage had significant effects on temporal variation of soil and bedrock water contents while the effect of water table evaporation was non-significant. Based on the magnitude of the regression coefficients, canal stage had the greatest influence on the temporal variation of soil and bedrock water contents at all elevations and distances from the canal at the locations monitored. The effect of canal stage and mean water table elevation in the DFA model was similar confirming the high hydraulic connectivity between the canal and Biscayne aquifer.

Based on the high connectivity between surface water in the canal and Biscayne aquifer, a simple DFA based regression model (DFA model in which the common trends were removed), was developed to predict soil and bedrock water contents as a function of canal stage and net recharge at various elevations. The performance of the simplified regression model was described as good to acceptable at sites with low elevation (i.e., water table elevation within 1m from the ground surface) and poor at the location at with water table depth greater than 1.5 m. These findings highlight the effect of micro-topography within the field on soil water content. The study also revealed that factor loadings were not zero for all the common trends suggesting that the explanatory variables (net recharge and canal stage) used in the DFA model are not sufficient to explain all the observed variations in the soil and bedrock water content time series.

The effect of the proposed 3 incremental raises in canal stage on soil and bedrock water content was simulated using the developed simple DFA based regression model for a total of 181 days beginning January 01, 2012. The results based on the data collected indicate that the proposed raises in canal stage would result in negligible changes in average soil and bedrock water contents at low elevations monitored in this study based. Changes in soil water content near the ground surface were more noticeable. The DFA based regression model developed is limited in its prediction ability to the range of canal elevations and net recharge by which it was developed. The uncertainty in predictions could be minimized by continuously updating the regression coefficients and constant level parameters as more data on response and explanatory variables are collected. The results of the regression model could be further evaluated using physically based modeling approaches. The approach used in this study could be applied to any system in which detailed physical modeling would be limited by inadequate information on parameters or processes governing the physical system.

Acknowledgements

The authors would like to thank the South Florida Water Management District for providing the funding for this study, the University of Florida IFAS Tropical Research and Education Center, and the University of Florida Agricultural and Biological Engineering Department, Mr. Vito Strano and Mr. Sam Accursio for allowing us to use their lands and Mrs. Tina Dispenza for her contribution towards data collection and processing. We also like to thank the three anonymous reviewers whose comments and suggestions greatly improved this manuscript.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automat. Control 19, 716–723.

ASCE, 2005., The ASCE Standardized Reference evapotranspiration Equation. Task Committee on Standardization of Calculation of Reference ET. Environment and Water Resources Institute of ASCE. 200 p.

Allen, R., 2011. REF-ET: Reference Evapotranspiration Calculation Software. User Manual.

Barquin, L.P., Migliaccio, K.W., Muñoz-Carpena, R., Schaffer, B., Crane, J.H., Li, Y.C., 2011. Shallow Water Table Contribution to Soil-Water Retention in Capillary Fringe of a Very Gravelly Loam Soil of South Florida. Vadose Zone J, 10:1–8.

Chin, D., 1991. Leakage of clogged channels that partially penetrate surficial aquifers. ASCE J. Hydraulic Engineering. 117, 467-488.

Chin, D., 2008. Phenomenological models of hydrologic processes in south Florida. J. Hydrol. 349, 230–243.

Dean, T.J., Bell J.P. Baty A.J.B., 1987. Soil moisture measurement by an improved capacitance technique. Part 1: sensor design and performance. Journal of Hydrology 93:67.

Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. J.R. Stat. Soc. Ser. B 39, 1–38.

Duwig, C., Normand, B., Vauclin, M., Vachaud, G., Green, S.R., Becquer, T., 2003. Evaluation of the WAVE model for predicting nitrate leaching for two contrasted soil and climate conditions. Vadose Zone J. 2, 76–89.

Gabriel, J. L., Lizaso J.I., Miguel, Q., 2010. Laboratory versus Field Calibration of Capacitance Probes. Soil Sci Soc Am J. 74, 593-601.

Genereux, D., Slater, E.,1999. Water exchange between canals and surrounding aquifer and wetlands in the Southern Everglades, USA. Journal of Hydrology 219 (1999), 153–168.

Geweke, J.F., 1977. The dynamic factor analysis of economic time series models. In: Aigner, D.J., Goldberger, A.S. (Eds.), Latent Variables in Socio-economic Models. North-Holland, Amsterdam, pp. 365–382.

Harvey, A.C., 1989. Forecasting, structural time series models and the Kalman filter. Cambridge Univ. Press, New York.

Kaplan, D., Muñoz-Carpena, R., Wan, Y., Hedgepeth, M., Zheng, F., Roberts, R., Rossmanith, R., 2010a. Linking river, floodplain, and vadose zone hydrology to improve restoration of a coastal river impacted by saltwater intrusion. J. Environ. Quality 39 (5), 1570–1584. doi:10.2134/jeq2009.0375.

Kaplan, D., Muñoz-Carpena, R., Ritter, A., 2010b. Untangling complex groundwater dynamics in the floodplain wetlands of a southeastern U.S. coastal river. Water Resour. Res. 46, W08528-10. doi:10.1029/2009WR009038.

Kaplan, D., Muñoz-Carpena R., 2011. Complementary effects of surface water and groundwater on soil moisture dynamics in a degraded coastal floodplain forest. J. of Hydrol. 398, 221–234.

Lütkepohl, H., 1991. Introduction to multiple time series analysis. Springer- Verlag, Berlin.

Márkus, L., Berke, O., Kovács, J., Urfer, W., 1999. Spatial prediction of the intensity of latent effects governing hydrogeological phenomena. Environmetrics. 10, 633-654.

McDonald, M.G., Harbaugh, A.W., 1988. A modular three-dimensional finite-difference ground-water flow model. Techniques of Water Resources Investigations of the United States Geological Survey, Book 6, Chapter A1, US Geological Survey, Reston, Virginia.

Merritt, M.L., 1996. Simulation of the water table altitude in the Biscayne aquifer, Southern Dade County, Florida, water years1945-89. Water-Supply Paper 2458, US Geological Survey.

Muñoz-Carpena, R., Ritter, A., Li, Y.C., 2005. Dynamic factor analysis of groundwater quality trends in an agricultural area adjacent to Everglades National Park. J. Contam. Hydrol. 80, 49–70.

Muñoz-Carpena, R., Ritter, A., Bosch, D.D., Schaffer, B., Potter, T.L., 2008. Summer cover crop impacts on soil percolation and nitrogen leaching from a winter corn field. J. Agricultural Water Management. 95, 633–644.

Munsell Soil Color Charts., 2000. Revised Edition. Greta G. Macbeth. New Windsor, NY.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models: Part 1-A. Discussion of principles. J. Hydrol. 10, 282–290.

Nobel, C.V., Drew, R.R.W., Slabaugh, J.D., 1996. Soil survey of Dade County Area, Florida, U.S. Department of Agriculture, NRCS Report, Washington, DC.

Pathak, S. C., 2008. South Florida Water Management District. South Florida Environmental Report Volume I. Appendix 2-1. Pg. 64. 3301 Gun Club Road, West Palm Beach, FL. Available online at:http://my.sfwmd.gov/portal/page/portal/pg_grp_sfwmd_sfer/portlet_sfer/tab2236041/volume1/vol1_table_of_contents.html.

Ritter, A., Regalado, C.M., Muñoz Carpena R., 2009. Temporal common trends of topsoil water dynamics in a humid subtropical forest watershed. Vadose Zone J, 8(2), 437–449.

Ritter A., Muñoz-Carpena, R., 2006. Dynamic factor modeling of ground and surface water levels in an agricultural area adjacent to Everglades National Park. J. Hydrol. 317, 340-354.

Salas, J.D., 1993. Analysis and modeling of hydrologic time series. p. 19.1–19.72. In D.R. Maidment (ed.) Handbook of hydrology. McGraw-Hill, New York.

Schaffer, B.,1998. Flooding Responses and Water-use Efficiency of Subtropical and Tropical Fruit Trees in an Environmentally-sensitive Wetland. Annals of Botany 81, 475–481. SFWMD, 2011. Past and Projected Trends in Climate and Sea Level for South Florida. Hydrologic and Environmental Systems Modeling. 3301 Gun Club Road West Palm Beach, Florida

Shumway, R.H., Stoffer D.S., 1982. An approach to time series smoothing and forecasting using the EM algorithm. J. Time Ser. Anal. 3, 253–264.

Šimůnek, J., van Genuchten, M. Th., Šejna, M., 2008. Development and applications of the HYDRUS and STANMOD software packages, and related codes, Vadose Zone J.7 (2), 587-600.

Skinner, C., Bloetscher, F., Pathak, C.S., 2008. Comparison of NEXRAD and Rain Gauge Precipitation Measurements in South Florida. J. of Hydrologic Engineering. 14(3), 248-260.

U.S. Army Corps of Engineers, and South Florida Water Management District, 2009. Comprehensive Everglades Restoration Plan: C-111 spreader canal western project: Draft integrated project implementation report and environmental impact statement.

USDA Soil Conservation Service. U.S. Department of Agriculture Handbook 18.

USDA National Agricultural Statistics Service for Miami-Dade County, Florida 2007. Available at: http://www.agcensus.usda.gov/Publications/2007/Online_Highlights/County_Profiles/Florida/cp12086.pdf

Vanclooster, M., Viaene, P., Diels, J., Christiaens, K., 1995. WAVE: A mathematical model for simulating water and agrochemicals inthe soil and vadose environment. Reference and user's manual (release2.0). Institute for Land and Water Management, Katholieke UniversiteitLeuven, Leuven, Belgium.

Water Resources Development Act, 2000. PUBLIC LAW 106–541—DEC. 11, 2000. http://www.fws.gov/habitatconservation/omnibus/wrda2000.pdf Retrieved July-23-2010.

Zou, S., Yu, Y., 1999. A dynamic factor model for multivariate water quality time series with trends. J. of Hydrol 178 (1–4), 381–400.

Zuur, A.F., Ieno, E.N., Smith, G.M., 2007. Analysing ecological data. Springer- Verlag, Berlin.

Zuur, A.F., Pierce, G.J., 2004. Common trends in Northeast Atlantic squid time series. J. Sea Res. 52, 57–72.

Zuur, A.F., Fryer, R.J., Jolliffe, I.T., Dekker, R., Beukema, J.J., 2003. Estimating common trends in multivariate time series using dynamic factor analysis. Environmetrics. 14 (7), 665–685.