Journal of Hydrology 398 (2011) 221–234

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Complementary effects of surface water and groundwater on soil moisture dynamics in a degraded coastal floodplain forest

D. Kaplan [1], R. Muñoz-Carpena *

Agricultural and Biological Engineering Dept., University of Florida, 287 Frazier Rogers Hall, PO Box 110570, Gainesville, FL 32611-0570, USA

ARTICLE INFO

Article history:
Received 27 May 2010
Received in revised form 15 December 2010
Accepted 17 December 2010
Available online 23 December 2010
This manuscript was handled by P. Baveye, Editor-in-Chief

Keywords:
Dynamic factor analysis
Soil moisture
Vadose zone
Surface water
Groundwater
Floodplain

SUMMARY

Restoration of degraded floodplain forests requires a robust understanding of surface water, groundwater, and vadose zone hydrology. Soil moisture is of particular importance for seed germination and seedling survival, but is difficult to monitor and often overlooked in wetland restoration studies. This research hypothesizes that the complex effects of surface water and shallow groundwater on the soil moisture dynamics of floodplain wetlands are spatially complementary. To test this hypothesis, 31 long-term (4-year) hydrological time series were collected in the floodplain of the Loxahatchee River (Florida, USA), where watershed modifications have led to reduced freshwater flow, altered hydroperiod and salinity, and a degraded ecosystem. Dynamic factor analysis (DFA), a time series dimension reduction technique, was applied to model temporal and spatial variation in 12 soil moisture time series as linear combinations of common trends (representing shared, but unexplained, variability) and explanatory variables (selected from 19 additional candidate hydrological time series). The resulting dynamic factor models yielded good predictions of observed soil moisture series (overall coefficient of efficiency = 0.90) by identifying surface water elevation, groundwater elevation, and net recharge (cumulative rainfall–cumulative evapotranspiration) as important explanatory variables. Strong and complementary linear relationships were found between floodplain elevation and surface water effects (slope = 0.72, $R^2 = 0.86$, $p < 0.001$), and between elevation and groundwater effects (slope = −0.71, $R^2 = 0.71$, $p = 0.001$), while the effect of net recharge was homogenous across the experimental transect (slope = 0.03, $R^2 = 0.05$, $p = 0.242$). This study provides a quantitative insight into the spatial structure of groundwater and surface water effects on soil moisture that will be useful for refining monitoring plans and developing ecosystem restoration and management scenarios in degraded coastal floodplains.

1. Introduction

Ecosystem restoration is often undertaken with the goal of returning a degraded (i.e., impacted, invaded, perturbed, altered, etc.) plant community to an earlier, more "natural" state (e.g., Pottier et al., 2009). Hydrological regime is often the primary environmental sieve (i.e., filter or barrier) (Harper, 1977) controlling seed germination, seedling recruitment, and long-term maintenance of plant species and communities, particularly in wetlands (van der Valk, 1981). Accordingly, wetland restoration efforts are usually built upon a foundation of hydrological restoration, whereby historical hydrological regimes and connections are reestablished in order to provide well-timed freshwater flows (Middleton, 2002), nutrients (Junk et al., 1989), and (where appropriate) the sediment required for accretion (DeLaune et al., 1994). A robust understanding of site hydrology is therefore vital for meeting restoration goals.

Hydrological monitoring and modeling efforts in support of wetland restoration usually focus on surface water (e.g., Wang, 1987), and less frequently, groundwater (e.g., Jung et al., 2004), but overwhelmingly overlook hydrological conditions in the vadose (unsaturated) zone, which largely dictate seed germination and seedling survival for many wetland plant species (Middleton, 1999). In addition to surface water performance measures like hydroperiod, restoration plans that rely on plant recruitment from existing seed banks, extant populations, or re-seeding must also ensure that restored areas experience the appropriate soil moisture regime to facilitate germination of desired species. Given the specific life-cycle requirements of many wetland plant species (Burns and Honkala, 1990; Conner, 1988; Conner et al., 1986; 1987; South Florida Water Management District [SFWMD], 2006), the success of floodplain forest restoration efforts relies on an accurate

Abbreviations: DFA, dynamic factor analysis; DFM, dynamic factor model; WTE, water table elevation; SWE, surface water elevation; R_{net}, net recharge; SFWMD, South Florida Water Management District; C_{eff}, Nash and Sutcliffe coefficient of efficiency; AIC, Akaike's information criterion; VIF, variance inflation factor.

* Corresponding author. Tel.: +1 352 392 1864x287; fax: +1 352 392 4092.
E-mail address: carpena@ufl.edu (R. Muñoz-Carpena).
[1] Present address: Ecohydrology Laboratory, School of Forest Resources and Conservation, University of Florida, 319 Newins-Ziegler Hall, PO Box 110410, Gainesville, FL 32611-0410, USA.

0022-1694/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2010.12.019

understanding of the relationships between vadose zone, surface water, and groundwater hydrology. However, finding direct relationships between basic hydrological inputs can be difficult due to interactions between surface water, groundwater, and porewater in variably saturated matrices with heterogeneous soils, vegetation, and topography (e.g., Gardner et al., 2002; Langevin et al., 2005). In particular, as noted by Rodriguez-Iturbe et al. (2007), quantifying wetland soil moisture dynamics in humid regions is especially vexing due to the complex interdependencies of climate, soil, vegetation, and stochastic water table variation.

Collection of long-term, high-resolution data serves to describe temporal soil moisture dynamics (i.e., magnitude, range, daily, seasonal and interannual variation, etc.) and spatial variation (e.g., Kaplan et al., 2010a). However, the intrinsic stochasticity of hydrological processes complicates the identification of hydrological fluxes that contribute to this observed variation. On the other hand, physically based models of the vadose zone (e.g., reviews in Šimůnek et al. (2003) and Vachaud et al. (1993) are useful exploratory tools to improve our understanding of these complex hydrological processes (Ritter et al., 2009), but require extensive parameterization and often rely on simplifying assumptions to estimate model boundary (Kampf and Burges, 2010) and initial conditions (Rocha et al., 2006).

Given these limitations, an alternative method for identifying possible shared variation and explanatory relationships is required. Dynamic factor analysis (DFA), a multivariate time series dimension reduction technique, provides the means to analyze complex, non-stationary environmental systems and decomposes observed times series variation into one or more common trends (which represent unexplained variation) and any number of explanatory variables. DFA is especially useful for assessing which explanatory variables (if any) most affect the time series of interest. DFA was initially developed for economic time series (Geweke, 1977), and has been applied in ecology to identify factors affecting squid populations (Zuur and Pierce, 2004), Atlantic bluefish (Addis et al., 2008), and commercial fisheries (Erzini, 2005; Tulp et al., 2008). In water resources, it has lately been applied to analyze groundwater dynamics (Kaplan et al., 2010b; Kovács et al., 2004; Ritter and Muñoz-Carpena, 2006); groundwater quality trends (Muñoz-Carpena et al., 2005; Ritter et al., 2007); and soil moisture dynamics (Regalado and Ritter, 2009a,b; Ritter et al., 2009). DFA produces alternative dynamic factor models (DFMs), driven by measured data. Since the DFMs are based on observed data, no *a priori* information about the physical system being modeled is required.

Based on previous work on coastal floodplain hydrodynamics (Kaplan et al., 2010a,b), this research hypothesizes that the complex effects of surface water and shallow groundwater on soil moisture dynamics in a coastal floodplain forest are spatially distributed and complementary. To test this hypothesis, the objectives of this study were to identify the external hydrological factors that explain observed variation in soil moisture dynamics in a degraded coastal floodplain forest, and to quantify their spatial distribution. This was accomplished through application of DFA to 12 long-term soil moisture datasets and other hydrological variables collected in and around the Loxahatchee River (Florida, USA), a managed coastal river where watershed modifications and management over the past century have led to reduced freshwater flow, inadequate hydroperiod, and a shift towards drier plant communities (SFWMD, 2009).

2. Materials and methods

2.1. Study area

The Loxahatchee River is located on the southeastern coast of Florida, USA (26° 59′ N, 80° 9′ W; Fig. 1) and is often referred to

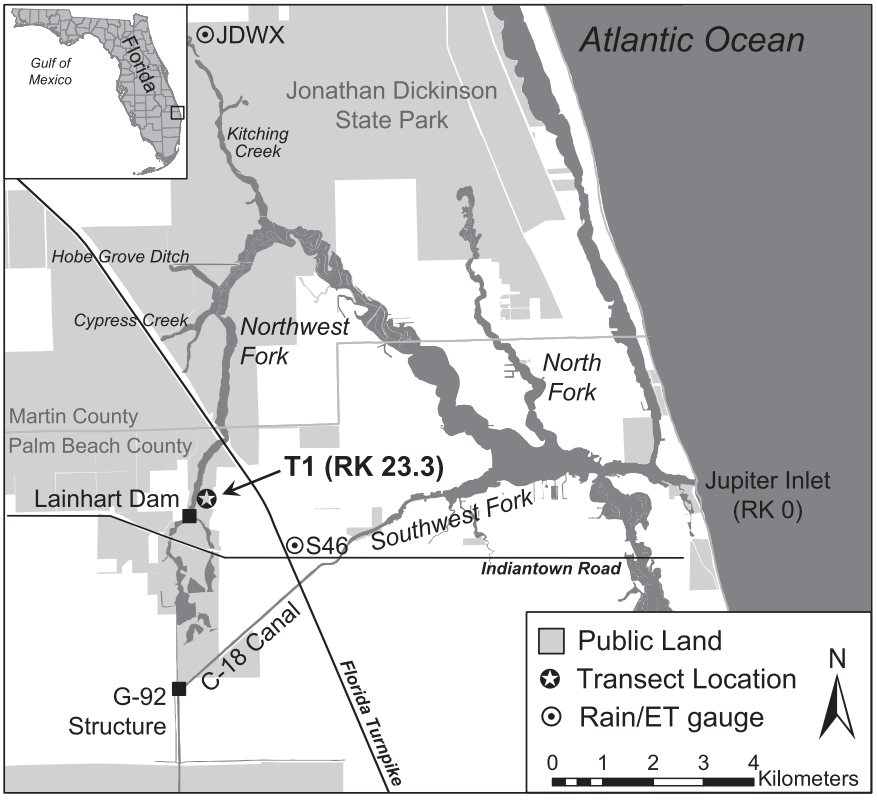

Fig. 1. The Loxahatchee River and surrounding area with experimental transect (T1), meteorological measurement locations (JDWX and S46), and major hydraulic infrastructure. Distance from river mouth indicated by river kilometer, RK.

as the "last free-flowing river in southeast Florida" (SFWMD, 2006). The upper watershed of the NW Fork is home to one of the last remnants of bald cypress (*Taxodium distichum* [L.] Rich) floodplain forest in southeast Florida, but modified watershed hydrology and management threaten this resource (SFWMD, 2006). Hydrologic changes have led to inadequate hydroperiod and soil moisture in the upstream riverine floodplain, which has shifted the system towards drier plant communities (SFWMD, 2009). Similar changes in the composition of floodplain vegetation as a result of reduced flooding frequency have been observed regionally and globally (e.g., Darst and Light, 2008; Leyer, 2005). Restoration of the Loxahatchee River is part of the Comprehensive Everglades Restoration Project (CERP), the most expensive ecological restoration project in history, with an initial budget of US$10 billion (SFWMD, 2006). Data collection and modeling efforts in the Loxahatchee River have been underway for several years (Muñoz-Carpena et al., 2008; SFWMD, 2002, 2006, 2009; VanArman et al., 2005), and have been directed at developing surface water management goals to maintain and restore the river's floodplain forest, but have largely overlooked groundwater and vadose zone hydrology in the floodplain. More recently, monitoring in the vadose zone (Kaplan et al., 2010a) has supported the development of initial relationships between surface water management and vadose zone conditions. However, improved soil moisture predictions are required to better assess the effects of restoration implementation and to guide adaptive management.

2.2. Experimental site and setup

The experimental site is a freshwater, riverine area, 23.3 km upstream of the river mouth (T1 in Fig. 1) and is not impacted by daily tides. Elevations range from 4.19 m to 1.66 m (Fig. 2; all elevations are referenced to the National Geodetic Vertical Datum of 1929, NGVD29). Soils on the higher elevation hydric hammock consist of Winder fine sand (a fine-loamy, siliceous, superactive, hyperthermic Typic Glossaqualf; (Soil Survey Staff, 1981), transitioning to sandy clay loam at depths of ~90 cm (Mortl et al., 2011). In the lower floodplain, soils are classified as fluvents – stratified entisols made up of interbedded layers of sand, clay, and organic matter, typical of areas with frequent flooding and deposition – with sand content increasing with depth (Mortl, 2006; Mortl et al., 2011). Vegetation communities in this area consist of hydric hammock at higher elevations and mature bald cypress floodplain forest (average diameter at breast height, DBH = 49 cm) at lower elevations (SFWMD, 2006). Low bald cypress recruitment and the invasion of less flood-tolerant species into the hydric hammock and riverine floodplain in this and other upstream areas have been documented (SFWMD, 2009), indicating the ecological impact of reduced moisture and shortened hydroperiod in the area.

Twelve coaxial impedance dielectric sensors (Hydra Probe, Stevens Water Monitoring Systems, Beaverton, OR, USA) measuring soil moisture, bulk electrical conductivity, and temperature were installed at four locations and three depths along a previously established vegetation survey transect perpendicular to the river (Fig. 2). Each cluster of three probes was wired to a field data logger (CR10/CR10-X, Campbell Scientific, Logan, Utah, USA), which recorded data every 30 min. Every 2–4 weeks, system batteries were changed and data were downloaded. Data collection began in September 2004 and continued through September 2008.

The Hydra probe determines soil moisture by measuring soil dielectric properties (Campbell, 1990). In this study, the temperature-corrected (25 °C) real portion of the dielectric constant was used to calculate soil moisture (θ) based on calibrations developed specifically for the soils of the Loxahatchee River floodplain by Mortl et al. (2011). When comparing θ across soils, we used the effective soil moisture, Θ_e [−], since the Winder fine sand and fluvent soils have substantially different hydraulic characteristics (Table 2 in Mortl et al. (2011). Θ_e scales θ from zero to unity and is calculated by:

$$\Theta_e = \frac{\theta - \theta_r}{\theta_s - \theta_r} \tag{1}$$

where θ is the actual (measured) soil moisture content [$m^3\ m^{-3}$], θ_r is the residual soil moisture content [$m^3\ m^{-3}$], and θ_s is the saturated soil moisture content [$m^3\ m^{-3}$].

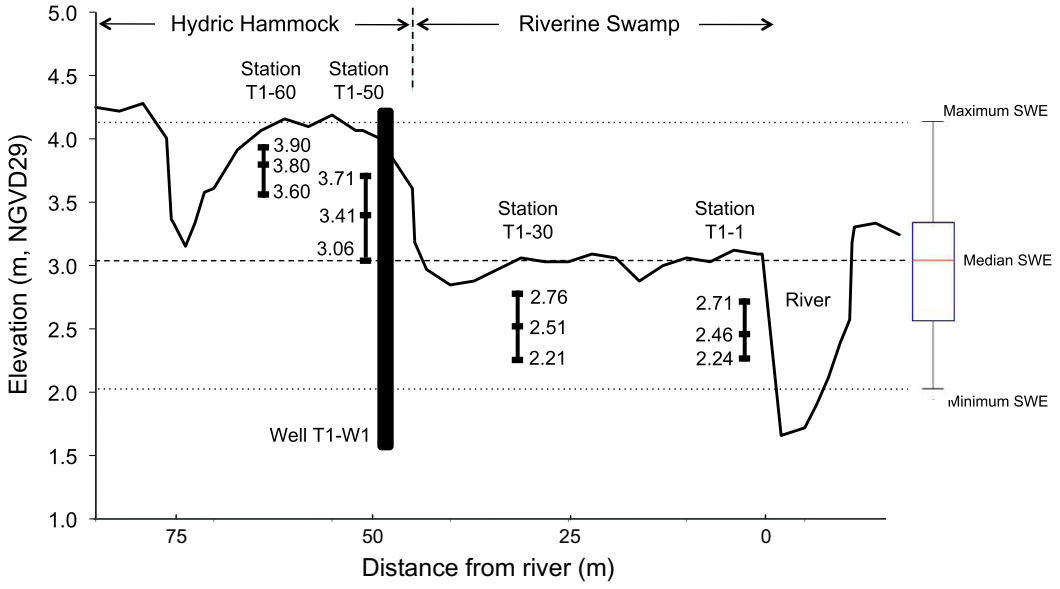

Fig. 2. Topographic cross-section of experimental transect with layout of vadose zone and groundwater monitoring instrumentation. Station names denote transect name (T1) and distance from the river (m). Probe installation elevations (m, NGVD29) listed with each station. Box and whisker plot of estimated adjacent surface water elevation (SWE) shows median, upper and lower quartile, and minimum/maximum values (no outliers less/greater than 1.5 times the interquartile range were observed). Vertical scale exaggerated ~10×.

2.3. Dynamic factor analysis

Soil moisture time series were investigated using dynamic factor analysis (DFA) (Zuur et al., 2003b). DFA is a parameter optimization and dimension reduction technique that is useful for identifying interactions between variables of interest and possible explanatory factors. With DFA, temporal variation in a set of N observed time series is modeled as a linear combination of one to M common trends, zero to K explanatory variables, a constant intercept parameter, and noise (Zuur et al., 2003b):

$$N \text{ time series} = M \text{ common trends} + \text{level parameter} + K \text{ explanatory variables} + \text{noise} \quad (2)$$

In this construction, the M common trends represent unexplained, shared variation among the N measured time series, the level parameter allows for relative shifts up and down, and the K are additional observed time series that represent explained variation. The goal of DFA is to identify one or more common trends in the set of observed time series that represent latent (unexplained) variation, minimizing the number of trends required to achieve a good fit with measured data. Appropriate explanatory variables may improve the model fit and point out which external factors most affect the response variables, improving conceptualization of the physical system that drives observed variation.

Mathematically, Eq. (2) may be written as:

$$s_n(t) = \sum_{m=1}^{M} \gamma_{m,n}\alpha_m(t) + \mu_n + \sum_{k=1}^{K} \beta_{k,n} v_k(t) + \varepsilon_n(t) \quad (3)$$

$$\alpha_m(t) = \alpha_m(t-1) + \eta_m(t) \quad (4)$$

where $s_n(t)$ is a vector containing the set of N time series being modeled (dubbed "response variables"); $\alpha_m(t)$ [same units as response variables] is a vector containing the common trends; $\gamma_{m,n}$ [dimensionless] are weighting coefficients that represent the relative importance of common trends to each response variable (dubbed "factor loadings"); μ_n [same units as response variables] is a constant level parameter that shifts series up or down; $v_k(t)$ [units vary] is a vector containing the explanatory variables; and $\beta_{k,n}$ [inverse units to convert $v_k(t)$ into response variable units] are weighting coefficients for the explanatory variables that indicate the relative importance of explanatory variables to each response variable (dubbed "regression parameters"). In this study, the response variables, $s_n(t)$, are the 12 Θ_e time series. The terms $\varepsilon_n(t)$ and $\eta_m(t)$ [same units as response variables] are independent, Gaussian noise with zero mean and unknown diagonal or symmetric/non-diagonal covariance matrix. DFMs with diagonal matrices may include a smaller number of model parameters than those with symmetric, non-diagonal matrices, but may also require a larger number of common trends to achieve adequate model fits (Zuur et al., 2003a).

Common trends, $\alpha_m(t)$, are modeled as a random walk (Harvey, 1989) and predicted with the Kalman filter/smoothing algorithm and Expectation Maximization (EM) techniques (Dempster et al., 1977; Shumway and Stoffer, 1982; Wu et al., 1996). The EM technique is also used to calculate factor loadings ($\gamma_{m,n}$) and level parameters (μ_n), while regression parameters ($\beta_{k,n}$) are modeled by linear regression (Zuur and Pierce, 2004). The $\gamma_{m,n}$ and $\beta_{k,n}$ accompanying common trends and explanatory variables allow us to identify the differential effects of common trends and explanatory variables on the soil moisture response variables. The significance of the $\beta_{k,n}$ were assessed using their magnitude and associated standard errors to compute a t-value. Relationships

Table 1
Hydrological time series used in the DFA.

Variable	Series type	No. of series	Description
Θ_e	Response	12	Effective soil moisture (–) in the floodplain root zone
SWE	Explanatory	1	Surface water elevation in the river (m, NGVD29) measured 0.45 km upstream of the experimental transect
WTE	Explanatory	1	Water table elevation (m, NGVD29) measured on the experimental transect, 50 m from the river
SWE_κ	Explanatory	8	Capped SWE (m, NGVD29), see explanation in text
WTE_κ	Explanatory	8	Capped WTE (m, NGVD29), see explanation in text
R_{net}	Explanatory	1	Cumulative net recharge (cumulative rainfall–cumulative ET, mm) calculated from rain gauge at the S46 structure and weather station in Jonathan Dickinson State Park (JDWX)

Table 2
Number of parameters, Nash–Sutcliffe coefficients of efficiency (C_{eff}), Akaike's information criteria (AIC), Bayesian Information Criterion (BIC), and Consistent Akaike's Information Criteria (CIAC) for selected dynamic factor models (DFMs). The best DFM for each model type are highlighted in bold and italics (i.e., Model I: trends and no explanatory variables; Model II: trends and explanatory variables; Model III: just explanatory variables).

DFM	Explanatory variables[a]	No. of trends	AIC	BIC	CIAC	C_{eff}	No. of parameters
Model I	0	1	15,585	16,140	16,215	0.61	75
	0	2	11,114	11,736	11,820	0.79	84
	0	3	7459	8140	8232	0.83	92
	0	4	4172	4906	5005	0.85	99
	0	***5***	***1292***	***2070***	***2175***	***0.90***	***105***
	0	6	−1949	−1135	−1025	0.91	110
	0	7	−3625	−2781	−2667	0.93	114
	0	8	−4603	−3736	−3619	0.95	117
	0	9	−6214	−5333	−5214	0.97	119
Model II	2 (SWE, $WTE_{\kappa=3.10}$)	3	2587	–	–	0.86	112
	3 (SWE, WTE, R_{net})	3	5796	–	–	0.84	122
	3 (SWE, $WTE_{\kappa=3.00}$, R_{net})	3	2724	–	–	0.88	122
	3 (SWE, $WTE_{\kappa=3.10}$, R_{net})	***3***	***2408***	–	–	***0.90***	***122***
	3 (SWE, $WTE_{\kappa=3.20}$, R_{net})	3	2937	–	–	0.87	122
Model III	***3 (SWE, $WTE_{\kappa=3.10}$, R_{net})***	***0***	***11,188***	–	–	***0.79***	***40***

[a] SWE, surface water elevation at Lainhart Dam; WTE, water table elevation; $WTE_{\kappa=x}$, water table elevation capped at $\kappa = x$; R_{net}, net recharge.

between response and explanatory variables were deemed significant for t-values >2 (Ritter et al., 2009). Relationships between response variables and common trends, on the other hand, were quantified with the canonical correlation coefficient ($\rho_{m,n}$). Values of $\rho_{m,n}$ close to unity indicated high association between the common trend and response variable. We classified the strength of these correlations into four groups: "minor" ($|\rho_{m,n}| < 0.25$); "low" ($0.25 \leqslant |\rho_{m,n}| < 0.50$); "moderate" ($0.50 \leqslant |\rho_{m,n}| < 0.75$); and "high" correlations ($|\rho_{m,n}| \geqslant 0.75$) (after Ritter et al., 2009).

2.4. Explanatory variables: meteorological, surface water, and groundwater data

Additional meteorological and hydrological variables were measured across the watershed, and a total of 31 daily time series (12 response variables and 19 candidate explanatory variables, each with 1442 daily values) were investigated for use in this analysis (Table 1). Since multi-collinearity may exist between explanatory variables measured at nearby locations, not all candidate explanatory variables could be used simultaneously. To assess the severity of multi-collinearity, we used the variance inflation factor (VIF) of each set of explanatory variables (Zuur et al., 2007), avoiding combinations of explanatory variables that resulted in VIF > 5 (Ritter et al., 2009).

Average annual precipitation in the Loxahatchee River watershed is 1550 mm year^{-1}, with approximately two-thirds falling during the wet season from May to October (Dent, 1997). Average annual evapotranspiration (ET) losses are 1140 mm year^{-1} in southern Florida (SFWMD, 2002). For this study, rainfall data were recorded at the S46 hydraulic structure on the Southwest Fork and ET data were recorded at the JDWX weather station in JDSP (Fig. 1). These data are publicly available and were downloaded from the SFWMD online environmental database, DBHYDRO (accessed at http://my.sfwmd.gov/dbhydroplsql/; stations S46_R and JDWX) and converted to daily means. Note that soil moisture data are autocorrelated (i.e., Θ_e at time t is dependent on Θ_e at $t-1$), while this is not true for rainfall and ET. To make these data potentially useful to the DFA, the difference between cumulative rainfall and cumulative ET was used to calculate a net recharge time series such that:

$$R_{\text{net},t} = \sum_{i=1}^{t} P_i - \sum_{i=1}^{t} ET_i \tag{5}$$

where $R_{\text{net},t}$ is the net recharge at time t [mm], P is precipitation [mm] and ET is evapotranspiration [mm] (Ritter et al., 2009).

Breakpoint surface water elevation (SWE) was measured at a SFWMD monitoring station (DBHYDRO station LNHRT_H) on the headwater side of Lainhart Dam (0.45 km upstream of the study area; Fig. 1) and converted to mean daily values. While SWE data were not available directly adjacent to the experimental transect, an available stage–discharge relationship between upstream SWE at Lainhart Dam and SWE at the study site (SFWMD, 2006) allowed us to estimate the range of inundation and drawdown in the floodplain (box-and-whisker plot in Fig. 2). SWE at Lainhart Dam is the primary controlled variable in restoration planning (SFWMD, 2002, 2006), and its use in this analysis allows for direct application of DFA results to proposed restoration scenarios.

Water table elevation (WTE) data were collected on T1 in a groundwater well located 50 m from the river (Fig. 2) using a multi-parameter water quality probe (TROLL 9000/9500, In-Situ Inc., Ft. Collins, CO, USA). The well was constructed of slotted 5.08-cm (nominally 2-in.) PVC pipe housed in a 20.32-cm (nominally 8-in.) PVC pipe. The screen size was 0.254 mm (nominally 0.01-in.) and the slotted section length was 0.61 m (nominally 2 ft), corresponding to slotted elevations between 1.66 and 2.27 m. WTE was measured every 30 min from September 2004 through January 2009 and converted to mean daily values. A full description of the groundwater dataset and QA/QC procedure are available in Muñoz-Carpena et al. (2008).

Unlike SWE and WTE, Θ_e time series are bound, by definition, between zero and unity (corresponding to 0% and 100% saturation, respectively; see Eq. (1)). Physically, this means that SWE and/or WTE may continue to increase after Θ_e reaches unity, after which the response variables lose dependence on these explanatory variables. To account for this, we calculated an additional set of explanatory variables that capped measured SWE and WTE variables by excluding values greater than a fixed elevation (κ [m, NGVD29]) according to:

$$\text{SWE}_{\kappa,t} = \min(\text{SWE}_t, \kappa) \tag{6}$$

$$\text{WTE}_{\kappa,t} = \min(\text{WTE}_t, \kappa) \tag{7}$$

where $\text{SWE}_{\kappa,t}$ and $\text{WTE}_{\kappa,t}$ are surface water and water table elevations [m, NGVD29] capped at fixed elevations, κ [m, NGVD29]. For this analysis we investigated values of κ ranging from the minimum and maximum soil moisture monitoring elevations in the floodplain (Fig. 2), i.e. 2.2–3.9 (in 0.1 m increments), to determine which capped time series served as the best explanatory variables in the DFA.

2.5. Analysis procedure

DFA was implemented using the Brodgar v. 2.6.5 statistical package (Highland Statistics Ltd., Newburgh, UK), which uses the "R" statistical software language, version 2.9.1 (R Core Development Team, 2009). To compare the relative importance of common trends and explanatory variables across response variables (Zuur et al., 2003b; Zuur and Pierce, 2004), all series were normalized (mean subtracted, divided by standard deviation). We carried out the DFA in three distinct steps, resulting in three models. Model I was developed by constructing a set of DFMs using an increasing number of common trends until model performance was deemed satisfactory according to goodness-of-fit indicators (Zuur et al., 2003a). Model II was developed by incorporating explanatory variables into the DFA until a combination of common trends and explanatory variables was identified that met or exceeded the goodness-of-fit indicators from Model I without exceeding the VIF criterion. The use of explanatory variables in Model II is intended to reduce the amount of unexplained variability and improved description of Θ_e in the floodplain. A final reduced model (Model III) was explored by using the explanatory variables identified in Model II to create a multi-linear model (Model III) without common trends. Model III was developed using a multiple regression procedure run in Matlab (2009b, The MathWorks, Inc., Natick, MA, USA).

DFM goodness-of-fit was quantified with the Nash–Sutcliffe coefficient of efficiency ($-\infty \leqslant C_{\text{eff}} \leqslant 1$, Nash and Sutcliffe, 1970) and Akaike's information criterion (AIC; Akaike, 1974). C_{eff} compares the variance between predicted and observed data about the 1:1 line, with $C_{\text{eff}} = 1$ indicating that the plot of predicted versus observed data matches the 1:1 line. The AIC is a statistical criterion that balances goodness-of-fit with model parsimony by rewarding goodness-of-fit but including a penalty term based on the number of model parameters. Generally, the DFM with the largest C_{eff} and smallest AIC are preferred.

Finally, to assess whether model performance had been improved through this analysis, DFMs were compared with the sigmoidal model of floodplain soil moisture based solely on SWE developed by Kaplan et al. (2010a) of the form:

$$\Theta_e = \frac{1}{1 + e^{-\left(\frac{SWE-a}{b}\right)}} \tag{8}$$

where SWE is measured at Lainhart Dam (m, NGVD29), and a and b are local parameters. Underlying Eq. (8) is the fundamental relationship describing θ as a function of soil water pressure head (ψ (e.g., Brooks and Corey, 1964; van Genuchten, 1980). Under relatively hydrostatic conditions (i.e., no inflow, outflow, or redistribution of soil water above the water table), ψ can be estimated as the distance to the water table (Skaggs, 1991). Since SWE and WTE at T1 are often coupled (see Section 3), directly linking θ to SWE is consistent with these fundamental relationships.

3. Results and discussion

3.1. Experimental time series

Fig. 3 shows hydrological time series collected on and near the experimental transect (T1). Data collected during this 4-year period represented a wide range of climatic conditions, including four wet/dry seasons; 2 years with above-average rainfall and hurricane-induced flooding (2004 and 2005); and the driest 2-year period (2006–2007) on record in south Florida in more than 75 years (Neidrauer, 2009). Rainfall and ET (Fig. 3a) followed a seasonal pattern, with wet season (May to October) rain accounting for 73–80% of yearly totals over the 4 years (mean 77%). SWE and WTE (Fig. 3b) were closely correlated in wet seasons ($r = 0.92$), but diverged during dry seasons when SWE remained impounded at a relatively constant level behind Lainhart Dam and WTE continued to decline – most notably in the summers of 2006 and 2007.

SWE and WTE dynamics were reflected in Θ_e time series (Fig. 3c–f). On the sandy hydric hammock (Fig. 3c and d), highest elevation Θ_e series (e.g., T1–60 [3.90 m]) were the most dynamic, responding quickly to rainfall and promptly draining.

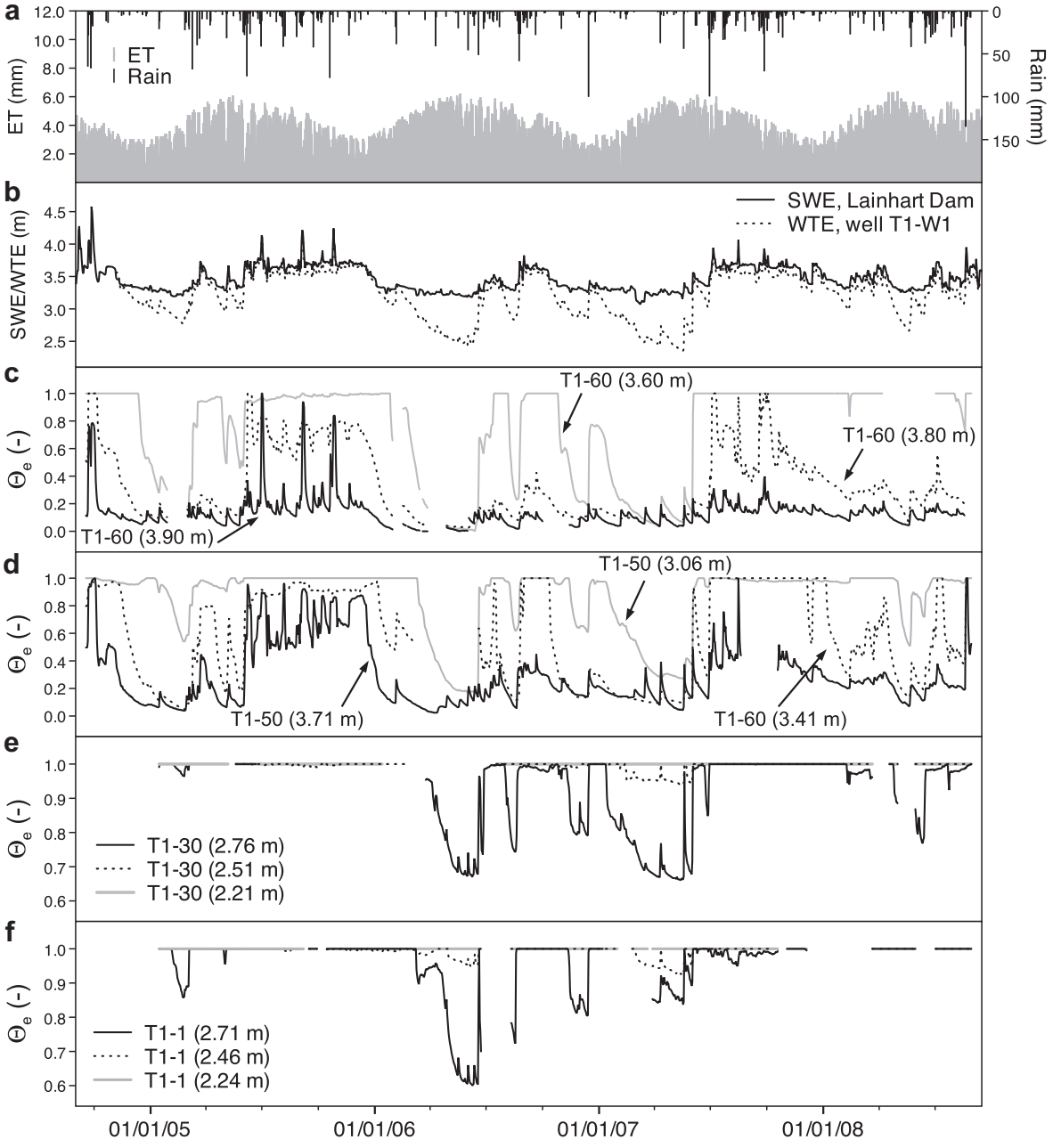

Fig. 3. Precipitation (a), evapotranspiration (ET) (a), surface water elevation (SWE) (b), water table elevation (WTE) (b), and effective soil moisture (Θ_e) (c–f) measured in and around the experimental site. Θ_e series names indicate experimental transect 1 (T1), distance from river (m) and installation elevation (m, NGVD29).

Middle elevation soils (T1–60 [3.80 m] and T1–50 [3.71 m]) showed a similar, but damped response to hydrological fluxes (rain, SWE, and WTE). Lowest elevation (i.e., deepest) soils (T1–60 [3.60 m] and T1–50 [3.06 m]) remained saturated for several months at a time, but all soils dried considerably during dry seasons. The lower floodplain was periodically inundated during the study period and Θ_e was at or close to saturation (i.e., $\Theta_e = 1$) at all depths for long periods, however surface soils (T1–30 [2.76 m] and T1–1 [2.71 m]) experienced considerable drying during all dry seasons (most markedly in 2006 and 2007; Fig. 3e and f, dark lines), with very dry conditions at surface of these highly organic and clayey soils. Lowest elevation soils (T1–30 [2.76 m] and T1–1 [2.71 m]) remained saturated for the entire study period (Fig. 3e and f, gray lines). As constant values, these series had zero variance and were thus removed from the DFA. Kaplan et al. (2010a) described these Θ_e datasets in further detail.

3.2. Dynamic factor analysis

3.2.1. Baseline DFA with no explanatory variables (Model I)

DFA was applied in three steps. First, different DFMs were obtained using $M = 1$–9 common trends and no explanatory variables to model the 10 response variables (time series for the two constantly saturated, lowest elevation probes were removed from the analysis). Initially, both diagonal and non-diagonal error covariance matrices were explored to identify the number of trends required to achieve a maximum C_{eff} and minimum AIC. However, when using a diagonal error covariance matrix, we found one or more common trends that exactly fit one or more of the response variables. This can occur with highly variable and "noisy" datasets and is referred to as a Heywood case (Highland Statistics, 2000). Since the goal of the DFA is to identify shared variation, non-diagonal error matrices were used in subsequent analyses to avoid this occurrence.

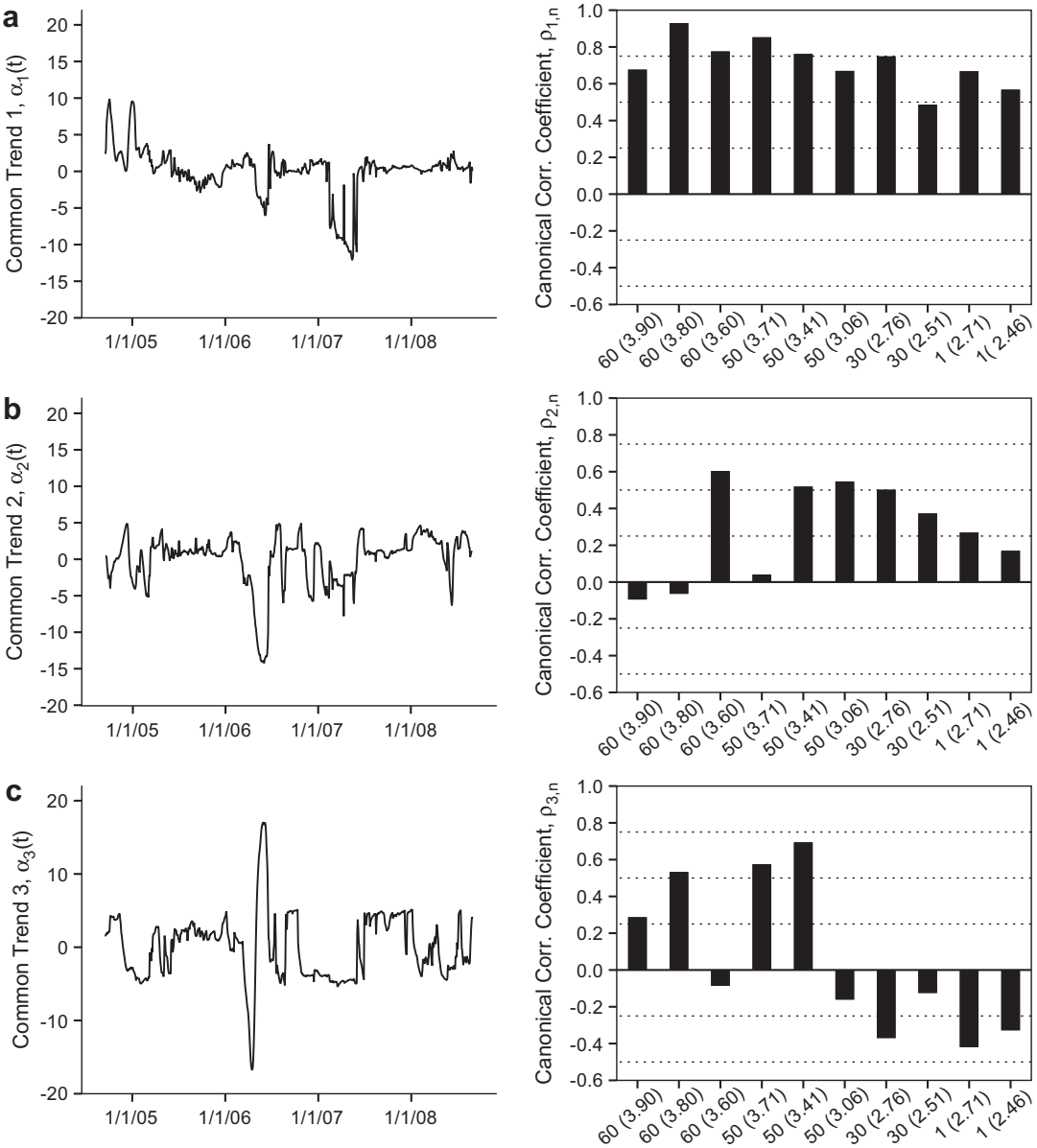

Fig. 4. The three most important trends from Model I (left) and their associated canonical correlation coefficients (right). Trend 1 (a) has high or moderate correlations with all response series; trend 2 (b) is most associated with middle and lower elevation series; trend 3 (c) is positively correlated with the four highest elevation series and negatively correlated with lower elevation series.

With a non-diagonal matrix, AIC continued to decrease and C_{eff} to increase with up to nine trends ($M = 9$). As an alternative to the AIC, we also examined the Bayesian Information Criterion (BIC; Schwarz, 1978) and Consistent Akaike's Information Criteria (CIAC; Bozdogan, 1987), which penalize additional parameters more strongly than the AIC. However, all metrics continued to decrease with up to nine trends (Table 2). That more than nine trends (representing unexplained information) were necessary to achieve the best DFM of 10 response variables suggested that multiple latent effects influence the variability of Θ_e at different depths across the floodplain. Since no inflection point in AIC, BIC, or CIAC was identified, we used C_{eff} and visual inspection as a measure of a model's goodness-of-fit. Though choice of a threshold C_{eff} is necessarily arbitrary, it is common to choose an appropriate model based on the reduction in model improvement with increased parameterization (e.g., Regalado and Ritter, 2009b). This led to the selection of the DFA with five common trends ($M = 5$; Table 2) as Model I, since addition of an extra trend had minimal impact on model performance. Model I had overall $C_{eff} = 0.90$ ($0.46 \leqslant C_{eff} \leqslant 0.99$) and AIC = 1292. Based on this selection, the objective of the subsequent DFA with explanatory variables (Model II) was to reduce the amount of unexplained variability in the DFM by achieving similar model performance using less than five common trends.

It is first instructive to examine common trends from Model I and their associated canonical correlation coefficients ($\rho_{m,n}$), since high $\rho_{m,n}$ values indicate high correlation with response variables. The three most important common trends from Model I (highest average $|\rho_{m,n}|$ across the 10 response variables) are shown in Fig. 4. Though only describing latent (unknown) variability at this stage, these trends and their patterns of correlation are useful for developing ideas about how Θ_e varies in the floodplain and where to look for the most useful explanatory variables. For example, common trend 1 (Fig. 4a, left panel) was highly ($|\rho_{m,n}| \geqslant 0.75$) to moderately ($0.50 \leqslant |\rho_{m,n}| < 0.75$) correlated with nine of the 10 Θ_e series (Fig. 4a, right panel), and appears to reflect large variation due to high water events associated with Hurricane Frances and Jeanne (which passed over the study site in 2004) and extended dry periods in the summers of 2006 and 2007. In general, correlations with common trend 2 (Fig. 4b, right panel) were weaker and positive, and highest for lower elevation Θ_e series. On the other hand, common trend 3 (Fig. 4c, left panel) was most correlated with Θ_e in higher elevation soils (Fig. 4c, right panel). This topographic distribution of $\rho_{m,n}$ suggested it would be useful to search for explanatory variables whose relative importance is split across lower and higher elevation response variables.

3.2.2. DFA with explanatory variables (Model II)

Next, explanatory variables were added to the model to reduce the number of common trends required while maintaining similar goodness-of-fit metrics as those from Model I. By adding explained

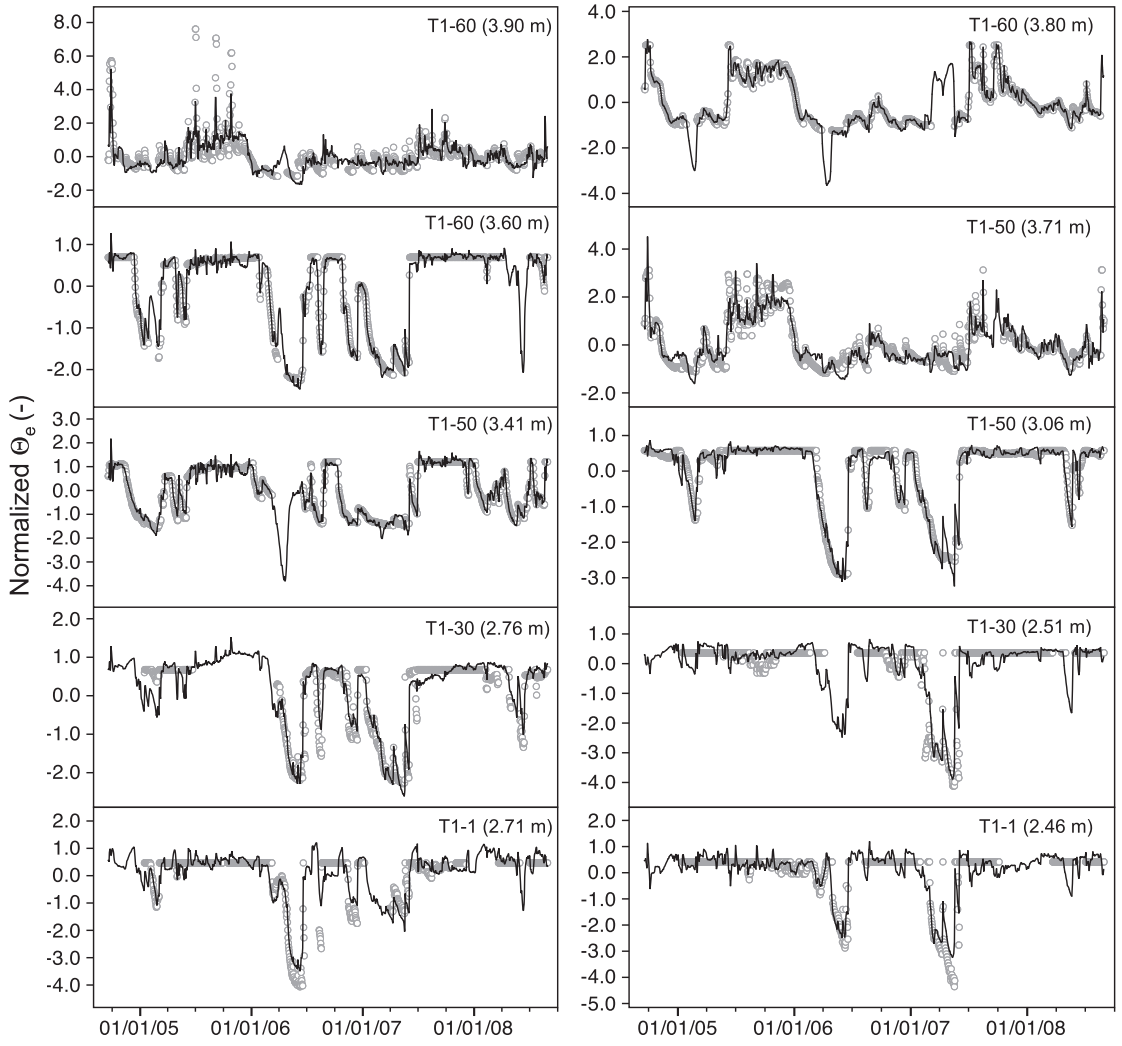

Fig. 5. Observed (gray symbols) and modeled (black lines) normalized Θ_e for the 10 response variables obtained from multi-linear Model II using three explanatory variables and three explanatory variables.

variability in this step, we also aimed to reduce canonical correlation coefficients and factor loadings of any remaining trends, indicating reduced dependency on unknown variation. A total of 19 candidate explanatory time series were explored, including: surface water elevation at Lainhart Dam (SWE); water table elevation in well T1-W1 (WTE); SWE_κ and WTE_κ series calculated with capped elevations (κ) ranging from 2.2 to 3.9 m; and net recharge calculated as the difference between cumulative rainfall and ET series (R_{net}) (Table 1). When two or more candidate explanatory variables were collinear or multi-collinear (resulting in VIFs > 5), the explanatory variable resulting in the best overall model fit (highest C_{eff} and lowest AIC) was selected.

Approximately 50 DFMs were developed with different combinations of common trends and explanatory variables. Results of selected DFMs are shown in Table 2. Finally, the best DFM used three explanatory variables ($K = 3$; $1.28 \leqslant VIF \leqslant 1.66$): SWE at Lainhart Dam, WTE_κ calculated with $\kappa = 3.10$ m, and net recharge (R_{net}). That the best DFM uses WTE_κ with $\kappa = 3.10$ m is likely due to the topography of the floodplain at T1 – when WTE is greater than about 3.10 m, the lower floodplain is inundated and the WTE series capped at 3.10 m does the best job of describing saturated conditions. Also note that the median estimated SWE adjacent to T1 is close to this elevation (see box-and-whisker plot inset in Fig. 2). On the other hand, DFMs developed using WTE_κ with different values of κ, capped SWE (SWE_κ), or only un-capped series did not perform as well (i.e., examples in Table 2). With these explanatory variables the number of required common trends was reduced from five to three ($M = 3$), reducing the unexplained variability in the model while achieving performance similar to that of Model I. This model (Model II) yielded an overall C_{eff} value of 0.90 across

Table 3
Constant level parameters (μ_n), canonical correlation coefficients ($\rho_{m,n}$), factor loadings ($\gamma_{m,n}$), regression coefficients ($\beta_{k,n}$), and coefficients of efficiency from Model II ($C_{eff,n}$) and empirical sigmoidal model ($C_{eff,sig}$) of Kaplan et al. (2010a). Significant regression parameters in bold.

s_n	μ_n	Canon. Corr. Coef.			Factor loadings			Regression coefficients			$C_{eff,n}$	$C_{eff,sig}$
		$\rho_{1,n}$	$\rho_{2,n}$	$\rho_{3,n}$	$\gamma_{1,n}$	$\gamma_{2,n}$	$\gamma_{3,n}$	βSWE_n	$\beta WTE_{(\kappa=3.1),n}$	$\beta R_{net,n}$		
T1–60 (3.90 m)	−0.03	0.38	0.27	−0.03	0.08	0.02	−0.02	**0.81**	**−0.16**	−0.03	0.67	0.64
T1–60 (3.80 m)	−0.06	0.31	0.86	−0.01	−0.02	0.21	0.03	**0.41**	**−0.13**	**0.18**	0.98	0.72
T1–60 (3.60 m)	−0.05	0.14	0.16	0.74	−0.04	0.02	0.15	**0.24**	**0.50**	**0.08**	0.99	0.78
T1–50 (3.71 m)	0.03	0.30	0.46	−0.11	0.03	0.07	−0.02	**0.71**	**−0.10**	**0.24**	0.87	0.77
T1–50 (3.41 m)	−0.09	−0.32	0.48	0.17	−0.14	0.04	0.03	**0.50**	**0.20**	**0.29**	0.97	0.82
T1–50 (3.06 m)	0.00	0.17	0.18	0.32	0.01	0.03	0.03	**−0.05**	**0.93**	**0.04**	0.94	0.65
T1–30 (2.76 m)	0.24	0.11	−0.07	0.57	0.01	−0.01	0.06	**0.10**	**0.63**	**0.20**	0.93	0.54
T1–30 (2.51 m)	−0.04	−0.13	−0.36	0.34	0.01	−0.07	−0.03	−0.03	**0.93**	**0.03**	0.86	0.33
T1–1 (2.71 m)	0.07	0.48	0.12	0.46	0.06	0.05	0.11	−0.02	**0.63**	**−0.07**	0.89	0.51
T1–1 (2.46 m)	0.09	0.24	−0.32	0.32	0.06	−0.08	−0.02	**0.11**	**0.76**	**−0.11**	0.89	0.34
									Overall:		0.90	0.68

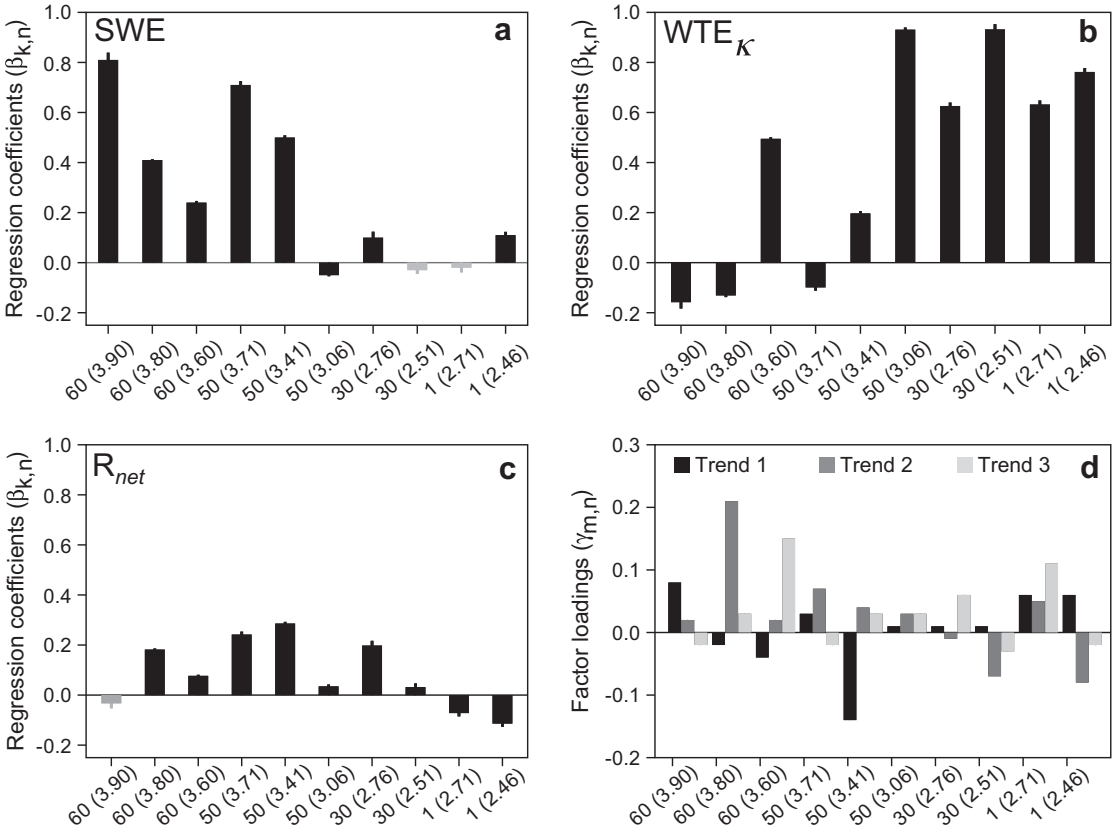

Fig. 6. Regression parameters (a–c) and factor loadings (d) for Model II (three common trends and three explanatory variables). Regression parameters (a–c) are shown with their standard errors, with black bars indicating significance. Note different *y*-axis scale on factor loadings panel (d).

the 10 Θ_e series (equal to the target of 0.90 from Model I) and an AIC value of 2408 (higher than the 1292 target from Model I due to an increase in the number of model parameters, but the lowest of any DFM with explanatory variables). Comparisons between Model II simulations and observed data are shown in Fig. 5.

Table 3 summarizes parameters obtained using Model II. Significant regression parameters (t-value > 2) are shown in bold. Canonical correlations were reduced from Model I, indicating a reduced dependence of the DFM on these latent series. While the trends in Model I had five "high" and 13 "moderate" correlations with response variables, trends in Model II had only two "high" and one "moderate" correlations. Model fits are fair to excellent ($0.67 \leqslant C_{\text{eff}} \leqslant 0.99$). Model II also out-performed the empirical sigmoidal model of Kaplan et al. (2010a) for all Θ_e series (Table 3), although at the cost of more parameters (122 for Model II versus 20 for the sigmoidal model). Model performance is particularly

Table 4
Constant level parameters (μ_n), model parameters, and coefficients of efficiency ($C_{\text{eff},n}$) from Model III (no trends, three explanatory variables) and empirical sigmoidal model ($C_{\text{eff,sig}}$) of Kaplan et al. (2010a). Significant model parameters in bold.

s_n	μ_n	Model parameters			$C_{\text{eff},n}$	$C_{\text{eff,sig}}$
		βSWE_n	$\beta WTE_{(\kappa=3.1),n}$	$\beta R_{\text{net},n}$		
T1–60 (3.90 m)	−0.04	**0.80**	**−0.15**	0.03	0.59	0.64
T1–60 (3.80 m)	−0.11	**0.78**	0.00	**0.20**	0.76	0.72
T1–60 (3.60 m)	−0.07	**0.27**	**0.72**	**0.00**	0.84	0.78
T1–50 (3.71 m)	0.03	**0.86**	**−0.17**	**0.23**	0.79	0.77
T1–50 (3.41 m)	−0.09	**0.69**	**0.21**	**0.13**	0.77	0.82
T1–50 (3.06 m)	0.00	0.01	**0.95**	**0.02**	0.93	0.65
T1–30 (2.76 m)	0.26	**0.05**	**0.72**	**0.27**	0.90	0.54
T1–30 (2.51 m)	−0.07	**−0.22**	**1.06**	**0.06**	0.82	0.33
T1–1 (2.71 m)	0.05	0.00	**0.75**	**−0.05**	0.73	0.51
T1–1 (2.46 m)	0.11	**−0.17**	**0.83**	0.02	0.78	0.34
		Overall:			0.79	0.68

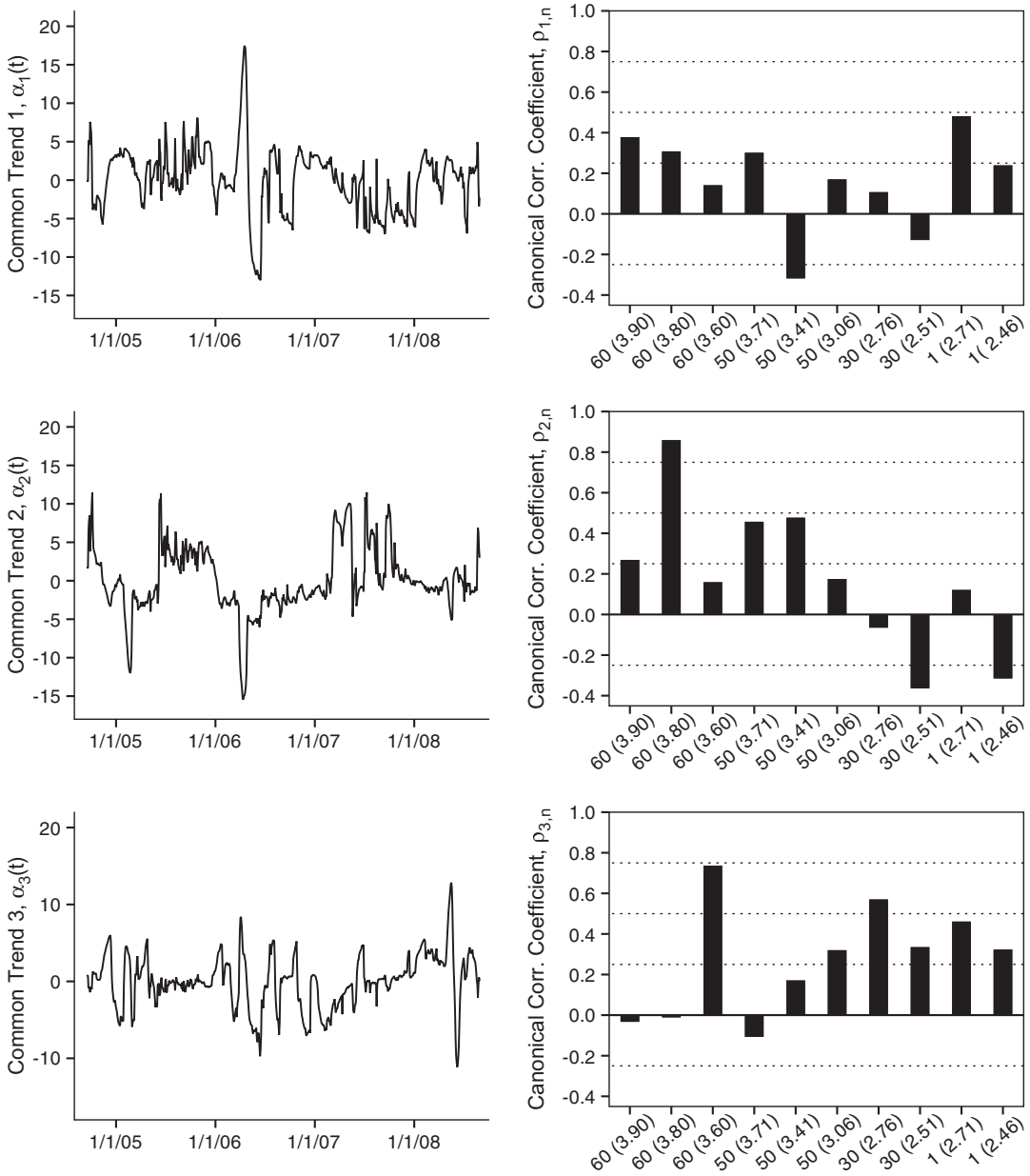

Fig. 7. Common trends from Model II (left) and their associated canonical correlation coefficients (right).

improved for Θ_e series in the lower floodplain (i.e., stations T1–30 and T1–1), where bald cypress seed germination will dictate the success of proposed restoration and management scenarios.

The spatially distributed effects of the explanatory variables and common trends on Model II are compared in Fig. 6. Fig. 6a shows that SWE was most important in describing variability in high and middle elevation soils on the hydric hammock, but had a reduced effect in lower elevation soils, particularly in the lower floodplain. On the other hand, WTE_κ had strong effects on Θ_e in lower elevation soils. This pattern follows from Model I, which identified different common trends grouped around elevation. Regression coefficients for the R_{net} series were weaker and spread across response variables, though generally positive and significant (Table 3). Inclusion of explanatory variables in Model II reduced factor loadings (Fig. 6d) slightly over those in Model I (overall average $|\gamma_n|$ for the five trends in Model I was 0.08 ± 0.08 compared to 0.05 ± 0.05 in Model II). While these trends are important for improving model fits for some response variables, this suggests that the Θ_e patterns observed in the Loxahatchee River floodplain may be adequately described using only the selected explanatory variables (see section 3.2.3).

The remaining three trends in Model II and their associated $\rho_{m,n}$ values are shown in Fig. 7. These common trends represent remaining unexplained (latent) variability among the Θ_e. Trend 1 (Fig. 7a) has low or minor correlations with all response series; trends 2 and 3 (Fig. 7b and c) are each highly associated with just one response variable, improving model fits for these series with little effect on other series. No additional spatial or physical interpretations were clear from these three remaining trends, suggesting that shared variation is being accounted for with explanatory variables.

3.2.3. Multi-linear regression model with no common trends (Model III)

Finally, common trends were removed from the model to assess model performance model using only explanatory variables. The three explanatory variables identified in Model II were used to create a multi-linear model of the response variables, Model III. As expected, C_{eff} values for Model III were somewhat reduced from Model II (overall $C_{eff} = 0.79$, $0.59 \leqslant C_{eff} \leqslant 0.93$; compared to $C_{eff} = 0.90$, $0.67 \leqslant C_{eff} \leqslant 0.99$ for Model II), but are adequate for most measurement locations (Table 4).

Comparisons between Model III simulations and observed data are presented in Fig. 8 and are fair to excellent, although for some series there are periods with reduced performance vis-à-vis Model II (e.g., compare model performance during extreme summer drawdowns in series T1–1 2.76 m and 2.41; bottom panels in Figs. 5 and 8). Despite this decrease in performance over Model II, Model III still out-performed the sigmoidal model for eight of the 10 Θ_e series and had only slightly inferior performance for the remaining two series, T1–60 (3.90 m) and T1–50 (3.41 m). Model improvements

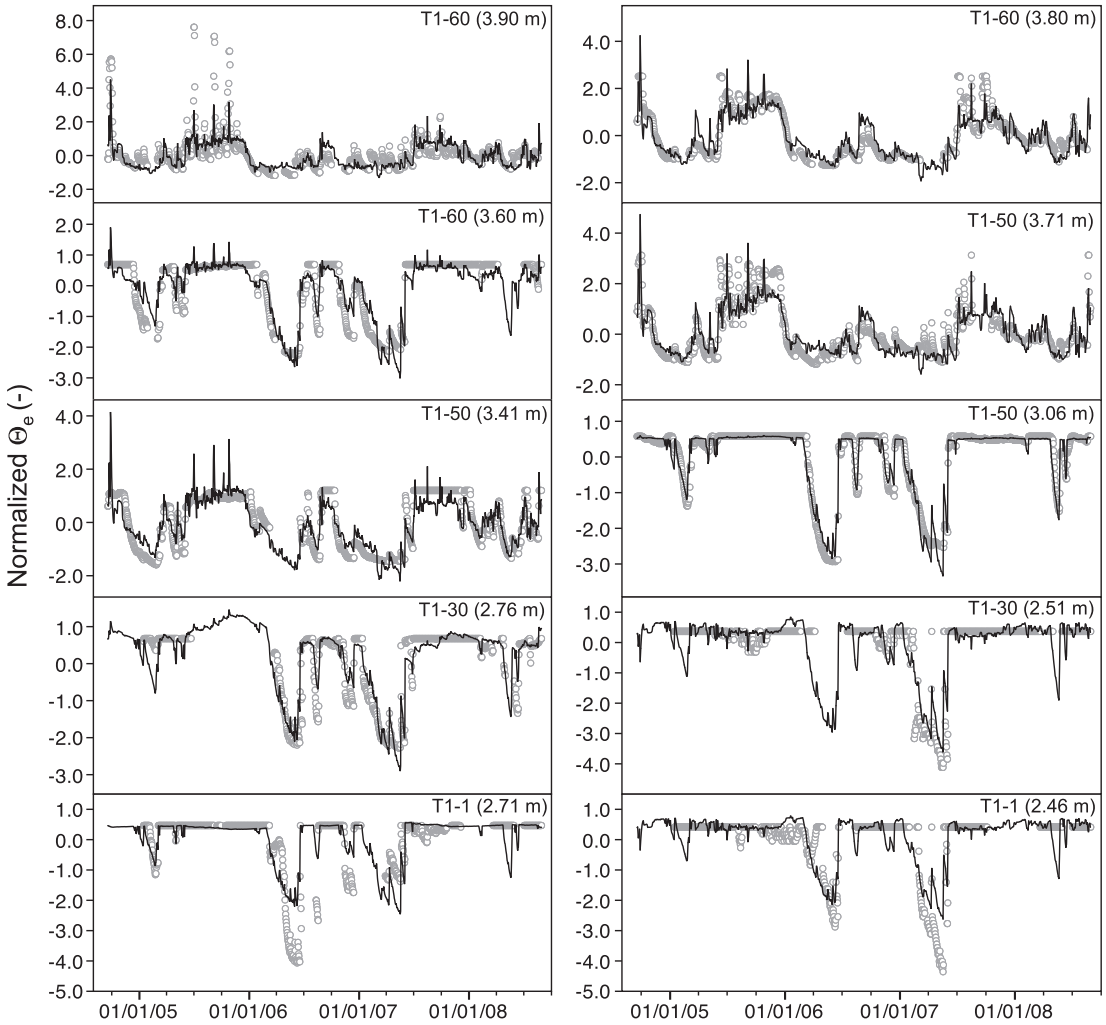

Fig. 8. Observed (gray symbols) and modeled (black lines) normalized Θ_e for the 10 response variables obtained from multi-linear Model III using three explanatory variables and no common trends.

for response variables in the lower floodplain realized in Model II were retained. Although empirical by nature, these results indicate that Model III may be useful for assessment of restoration scenarios for the floodplain wetlands of the Loxahatchee River, particularly in light of the wide range of climatic conditions captured in the experimental period.

3.2.4. Spatial complementarity of surface water and groundwater effects

Fig. 9 illustrates the spatial relationships between regression coefficients and floodplain elevation for the three explanatory variables from Model III. βSWE (Fig. 9a) and βWTE_{κ} (Fig. 9b) are strongly correlated with floodplain elevation (with equal but opposite slopes), highlighting the complementary effects of these variables on floodplain soil moisture. Noise in these relationships is likely due to soil heterogeneities with depth and distance from the river and increasing water table elevation with increasing distance from the river (Kaplan et al., 2010a). Correlation between βR_{net} and floodplain elevation (Fig. 9c), on the other hand, is not significant, suggesting that the effect of net recharge is relatively homogenous across the transect and is not strongly affected by different soils, vegetation, or duration of flooding. Incorporation of these elevation-based relationships for regression coefficients explicitly introduces spatial dependence in Model III. Additionally, replacing the 30 regression parameters from Model III with four parameters from the two linear equations proposed in Fig. 9a and b and a single parameter for the effect of R_{net} (estimated as the average value; Fig. 9c) allows a further model simplification. In this way, the total number of empirical parameters required is reduced from 40 (30 regression parameters and 10 level parameters) to just 15 (five elevation-based empirical parameters and 10 level parameters), while slightly reducing the model performance (overall $C_{eff} = 0.70$, $0.51 \leqslant C_{eff} \leqslant 0.85$).

4. Summary and conclusions

Bald cypress floodplain forests rely on a series of environmental sieves to maintain mature trees and achieve periodic seed germination and seedling recruitment. Accordingly, restoring natural hydrological and ecological dynamics to degraded ecosystems requires a thorough understanding of surface water–groundwater–vadose zone relationships. In particular, restoration flows must ensure an appropriate soil moisture regime when seeds are available for germination. In this study, long-term multivariate hydrological time series, measured in and around the Loxahatchee River in south Florida, were studied using dynamic factor analysis (DFA) in order to investigate soil moisture variation along a floodplain transect in a degraded coastal floodplain forest. The method proved to be a useful tool for the study of interactions among 29 long-term, non-stationary hydrological time series (10 effective soil moisture [Θ_e] series and 19 candidate explanatory variables). Using DFA, the factors underlying the complex variability observed in these multivariate hydrological datasets were identified. The resulting models are useful for assessing the effects of proposed ecological restoration and management scenarios on Θ_e dynamics in the floodplain of the Loxahatchee River.

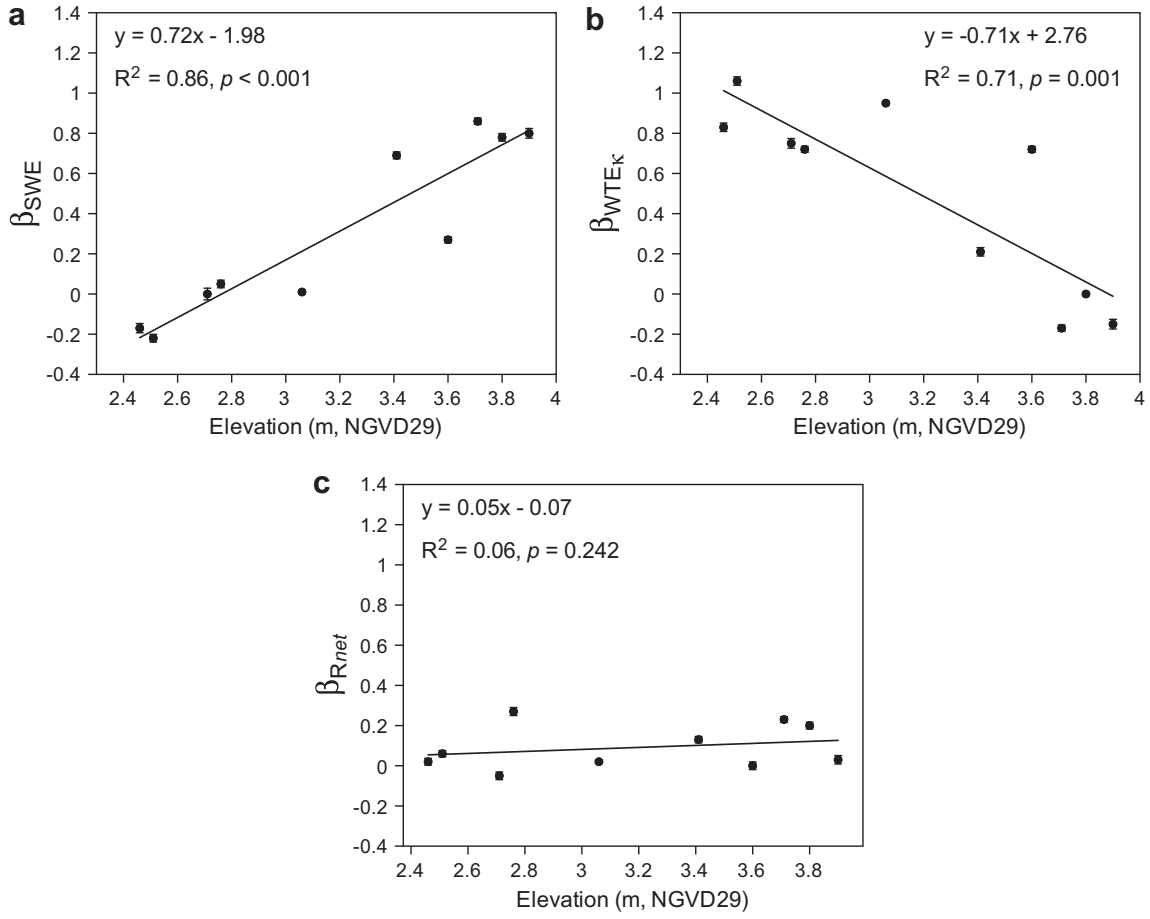

Fig. 9. Model III regression coefficients plotted against floodplain elevation for (a) surface water elevation (SWE); (b) water table elevation capped at 3.10 m (WTE_{κ}); and (c) net recharge (R_{net}). Error bars correspond to standard error calculated from multiple regression.

We found a minimum of five common trends (representing unexplained variability) were necessary to adequately describe observed Θ_e variation in the 10 response variables (Model I). Dependence on this unexplained variability was reduced by including appropriate explanatory variables selected from hydrological data measured in the area. The resulting model (Model II) required fewer trends, and those that remained were less important to the model (reduced canonical correlations and factor loadings). Model II also identified the most useful explanatory variables for describing Θ_e variation – surface water elevations (SWE), capped water table elevation (WTEκ), and cumulative net recharge (R_{net}) – and quantified the spatial distribution of their importance to Θ_e in each location. Finally, by removing common trends from Model II, we found variation in Θ_e series to be adequately described using just these three explanatory variables (Model III; overall $C_{eff} = 0.79$, $0.59 \leqslant C_{eff} \leqslant 0.93$).

A quantitative measure (strong linear correlation) of the differential and complementary effects of surface water and groundwater on floodplain Θ_e as a function of floodplain elevation was identified (with equal but opposite slopes), while the effects of R_{net} were weaker and homogenous across the experimental transect. These findings were used first to reduce model dependence on empirical factors, and second to introduce landscape effects into the model by explicitly considering floodplain elevation. The results from this study have practical implications, beyond guiding restoration planning and ecohydrological analysis for the Loxahatchee River. For instance, the effect of elevation on the interdependence of groundwater–surface water–vadose zone dynamics dictates the need for structured hydrological monitoring plans that account for elevation. Furthermore, restoration plans that address the vadose zone must consider the impacts of both surface water and groundwater management. For example, in floodplain systems like that studied here, increased consumptive use that draws down local water table elevation will most strongly impact the deeper vadose zone (and the mature, deeply-rooted plants). On the other hand, surface water withdrawals or management that reduce the frequency and duration of overbank flooding will mostly affect the shallower vadose zone and floodplain surface (and the shallow-rooted plants and seed germination). While the surface water–groundwater–vadose zone system is often highly connected in these variably flooded systems, understanding and quantifying the intricacies of these dynamics will provide a robust science-based water management plan to best restore degraded ecosystems.

Despite the empirical and data-driven foundation of DFA, we believe that the identification and quantification of complex time series shared variation and spatial interactions provided by DFA offers critical information for ensuing mechanistic or conceptual modeling efforts. Therefore we do not propose DFA as the only modeling alternative, but as a useful exploratory tool that can inform refined monitoring, restoration, and modeling efforts. In ongoing work, we aim to incorporate these and other hydrological relationships into an ecohydrological model to predict long-term effects of restoration scenarios on floodplain vegetation.

Acknowledgments

This work was funded in part by the South Florida Water Management District (SFWMD). The authors would like to thank Paul Lane (UF) for field and laboratory support; Yongshan Wan, Fawen Zheng, Gordon Hu, and Marion Hedgepeth (SFWMD) for their continued support; and Rob Rossmanith (Florida Park Service) for providing groundwater data. D. Kaplan acknowledges additional financial support from the University of Florida Graduate Fellowship Program.

References

Addis, P., Dean, J.M., Pesci, P., Locci, I., Cannas, R., Corrias, S., Cau, A., 2008. Effects of local scale perturbations in the Atlantic bluefin tuna (*Thunnus thynnus* L.) trap fishery of Sardinia (W. Mediterranean). Fish. Res. 92 (2–3), 242–254.

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

Bozdogan, H., 1987. Model selection and Akaike's information criterion (AIC): the general theory and its analytical extensions. Psychometrika 52, 345–370.

Brooks, R.H., Corey, A.T., 1964. Hydraulic Properties of Porous Media. Hydrology Paper No. 3, Civil Eng. Dept., Colorado State Univ., Fort Collins, 27p.

Burns, R.M., Honkala, B.H., 1990. Silvics of North America. USDA Forest Serv. Agric. Handbook 654, USDA Forest Serv., Washington, DC.

Campbell, J.E., 1990. Dielectric properties and influence of conductivity in soils at one to fifty Megahertz. Soil Sci. Soc. Am. J. 54, 332–341.

Conner, W.H., Toliver, J.R., Sklar, F.H., 1986. Natural regeneration of baldcypress [*Taxodium distichum* (L.) Rich.] in a Louisiana swamp. Forest Ecol. Manage. 14, 305–317.

Conner, W.H., 1988. Natural and artificial regeneration of baldcypress (*Taxodium distichum* [L.] Rich.) in the Barataria and Lake Verret basins of Louisiana. Ph.D. Diss. Louisiana State Univ., Baton Rouge.

Conner, W.H., Toliver, J.R., 1987. Vexar seedling protectors did not reduce nutria damage to planted baldcypress seedlings. USDA Forest Serv., Tree Planter's Notes 38 (3), 26–29.

Darst, M.R., Light, H.M., 2008. Drier Forest Composition Associated With Hydrologic Change in the Apalachicola River Floodplain, Florida. USGS Scientific Investigations Report 2008-5062, USGS, Reston, VA.

DeLaune, R.D., Nyman, J.A., Patrick Jr., W.H., 1994. Peat collapse, ponding and wetland loss in a rapidly submerging coastal marsh. J. Coastal Res. 10 (4), 1021–1030.

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

Dent, R.C., 1997. Rainfall Observations in the Loxahatchee River Watershed. Report, Loxahatchee River District, Jupiter, FL.

Erzini, K., 2005. Trends in NE Atlantic landings (southern Portugal): identifying the relative importance of fisheries and environmental variables. Fish. Oceanogr. 14 (3), 195–209.

Gardner, L.R., Reeves, H.W., Thibodeau, P.M., 2002. Groundwater dynamics along forest-marsh transects in a southeastern salt marsh, USA: description, interpretation and challenges for numerical modeling. Wetlands Ecol. Manage. 10, 145–159.

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.

Harper, J.L., 1977. Population Biology of Plants. Academic Press, New York.

Harvey, A.C., 1989. Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press, New York.

Highland Statistics, 2000. Software package for Multivariate Analysis and Multivariate Time Series Analysis, Version 2. Highland Statistics, Ltd., Newburgh, UK.

Jung, M., Burt, T.P., Bates, P.D., 2004. Toward a conceptual model of floodplain water table response. Water Resour. Res. 40 (12), 9.1–9.13.

Junk, W.J., Bayley, P.B., Sparks, R.W., 1989. The flood pulse concept in river–floodplain systems. In: Dodge, D.P (Ed.), Proc. of the International Large River Symposium, Canadian Special Publication of Fisheries and Aquatic Sciences, pp. 110–127.

Kampf, S., Burges, S.J., 2010. Quantifying the water balance in a planar hillslope plot: effects of measurement errors on flow prediction. J. Hydrol. 380 (1–2), 191–202.

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.

Kovács, J., Márkus, L., Halupka, G., 2004. Dynamic factor analysis for quantifying aquifer vulnerability. Acta Geol. Hung. 47 (1), 1–17.

Langevin, C., Swain, E., Wolfert, M., 2005. Simulation of integrated surface-water/ground-water flow and salinity for a coastal wetland and adjacent estuary. J. Hydrol. 314, 212–234.

Leyer, I., 2005. Predicting plant species' responses to river regulation: the role of water level fluctuations. J. Appl. Ecol. 42, 239–250.

Middleton, B.A., 1999. Wetland Restoration, Flood Pulsing and Disturbance Dynamics. John Wiley and Sons, Inc.,, New York.

Middleton, B.A., 2002. Flood Pulsing in Wetlands: Restoring the Natural Hydrological Balance. John Wiley and Sons, Inc., New York.

Mortl, A., 2006. Monitoring Soil Moisture and Soil Water Salinity in the Loxahatchee Floodplain. M.S. Thesis, Univ. of Florida, Gainesville.

Mortl, A., Muñoz-Carpena, R., Kaplan, D., Li, Y., 2011. Calibration of a combined dielectric probe for soil moisture and porewater salinity measurement in organic and mineral coastal wetland soils. Geoderma. doi:10.1016/j.geoderma.2010.12.007.

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., Kaplan, D., Gonzalez, F.J., 2008. Groundwater Data Processing and Analysis for the Loxahatchee River Basin. Final Project Report to the South Florida Water Management District, Univ. of Florida, Gainesville.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models, Part 1 – a discussion of principles. J. Hydrol. 10, 282–290.

Neidrauer, C., 2009. Water Conditions Summary. SFWMD, Operations Control Dept., West Palm Beach, FL. <http://www.sfwmd.gov/paa_dad/docs/F755378308/W%20Item8A_Water%20Conditions%20-%20C%20Neidrauer.pdf> (verified 12.05.10).

Pottier, J., Bédécarrats, A., Marrs, R.H., 2009. Analysing the spatial heterogeneity of emergent groups to assess ecological restoration. J. Appl. Ecol. 46 (6), 1248–1257.

R Development Core Team, 2009. R: A Language and Environment for Statistical Computing. Statistical Computing, Vienna, Austria.

Regalado, C.M., Ritter, A., 2009a. A bimodal four-parameter lognormal linear model of soil water repellency persistence. Hydrol. Process. 23, 881–892.

Regalado, C.M., Ritter, A., 2009b. A soil water repellency empirical model. Vadose Zone J. 8, 136–141.

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.

Ritter, A., Muñoz-Carpena, R., Bosch, D.D., Schaffer, B., Potter, T.L., 2007. Agricultural land use and hydrology affect variability of shallow groundwater nitrate concentration in South Florida. Hydrol. Process. 21, 2464–2473.

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.

Rocha, D., Abbasi, F., Feyen, J., 2006. Sensitivity analysis of soil hydraulic properties on subsurface water flow in furrows. J. Irrig. Drain E – ASCE 132 (4), 418–424.

Rodriguez-Iturbe, I., D'Odorico, P., Laio, L., Ridolfi, L., Tamea, S., 2007. Challenges in humid land ecohydrology: interactions of water table and unsaturated zone with climate, soil, and vegetation. Water Resour. Res. 43, W09301.

Schwarz, G.E., 1978. Estimating the dimension of a model. Ann. Stat. 6 (2), 461–464.

SFWMD, 2002. Technical Criteria to Support Development of Minimum Flow and Levels for the Loxahatchee River and Estuary. Water Supply Department, Water Resources Management, South Florida Water Management District, West Palm Beach, FL.

SFWMD, 2006. Restoration Plan for the Northwest Fork of the Loxahatchee River, Coastal Ecosystems Division. South Florida Water Management District, West Palm Beach, FL.

SFWMD, 2009. Riverine and Tidal Floodplain Vegetation of the Loxahatchee River and its Major Tributaries. South Florida Water Management District (Coastal Ecosystems Division) and Florida Park Service (5th District), West Palm Beach, FL, 2009.

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., Jarvis, N.J., van Genuchten, M.T., Gärdenäs, A., 2003. Review and comparison of models for describing nonequilibrium and preferential flow and transport in the vadose zone. J. Hydrol. 272, 14–35.

Skaggs, W., 1991. Theory of Drainage – Saturated Flow. BAE 671 Course Notes. Biological and Agricultural Engineering Dept., North Carolina State University, Raleigh.

Soil Survey Staff, 1981. Soil Survey of Martin County Area, Florida. USDA Soil Conserv. Serv. Washington, DC.

Tulp, I., Bolle, L.J., Rijnsdorp, A.D., 2008. Signals from the shallows: In search of common patterns in long-term trends in Dutch estuarine and coastal fish. J. Sea Res. 60 (1–2), 54–73.

Vachaud, G., Vauclin, M., Addiscott, T.M., 1993. Solute transport in the vadose zone: a review of models. In: Avogadro, A., Ragaini, R.C. (Eds.), Technologies for Environmental Cleanup: Soil and Groundwater, vol. 1. Kluwer Academic Publishers, Dordrecht, pp. 157–185.

van der Valk, A.G., 1981. Succession in wetlands: a Gleasonian approach. Ecology 62, 688–696.

van Genuchten, M., 1980. A closed form equation for predicting the hydraulic conductivity of unsaturated soil. Soil Sci. Soc. Am. J. 44, 892–898.

VanArman, J., Graves, G.A., Fike, D.L., 2005. Loxahatchee watershed conceptual ecological model. Wetlands 25 (4), 926–942.

Wang, F.C., 1987. Dynamics of intertidal marshes near shallow estuaries in Louisiana. Wetlands Ecol. Manage. 5 (2), 131–143.

Wu, L.S.-Y., Pai, J.S., Hosking, J.R.M., 1996. An algorithm for estimating parameters of state-space models. Stat. Prob. Lett. 28, 99–106.

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

Zuur, A.F., Ieno, E.N., Smith, G.M., 2007. Analysing Ecological Data. Springer, New York.

Zuur, A.F., Tuck, I.D., Bailey, N., 2003a. Dynamic factor analysis to estimate common trends in fisheries time series. Can. J. Fish. Aquat. Sci. 60, 542–552.

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