Untangling complex shallow groundwater dynamics in the floodplain wetlands of a southeastern U.S. coastal river

D. Kaplan,[1] R. Muñoz-Carpena,[1] and A. Ritter[2,3]

Received 22 December 2009; revised 20 March 2010; accepted 30 March 2010; published 14 August 2010.

[1] Understanding the hydrological functioning of tidally influenced floodplain forests is essential for advancing ecosystem protection and restoration goals in impacted systems. However, finding direct relationships between basic hydrological inputs and floodplain hydrology is hindered by complex interactions between surface water, groundwater, and atmospheric fluxes in a variably saturated matrix with heterogeneous soils, vegetation, and topography. Thus, an explanatory method for identifying common trends and causal factors is required. Dynamic factor analysis (DFA), a time series dimension reduction technique, models temporal variation in observed data as linear combinations of common trends, which represent unidentified common factors, and explanatory variables. In this work, DFA was applied to model water table elevation (WTE) in the floodplain of the Loxahatchee River (Florida, USA), where altered watershed hydrology has led to changing hydroperiod and salinity regimes and undesired vegetative changes in the floodplain forest. The technique proved to be a powerful tool for the study of interactions among 29 long-term, nonstationary hydrological time series (12 WTE series and 17 candidate explanatory variables). Regional groundwater circulation, surface water elevations, and spatially variable net local recharge (cumulative rainfall – cumulative evapotranspiration) were found to be the main factors explaining groundwater profiles. The relative importance of these factors was spatially related to floodplain elevation, distance from the river channel, and distance upstream from the river mouth. The resulting dynamic factor model (DFM) simulated the WTE time series well (overall coefficient of efficiency, $C_{\text{eff}} = 0.91$) and is useful for assessing management scenarios for ecosystem restoration and predicted sea level rise.

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

1. Introduction

[2] Description and modeling of hydroperiod, surface water salinity, groundwater elevation and salinity, soil moisture, and soil porewater salinity are essential to understanding the hydrological and ecological functioning of coastal floodplain wetlands [e.g., *Glamore and Indraratna*, 2009; *Nyman et al.*, 2009; *Benke et al.*, 2000]. However, finding direct relationships between basic hydrological inputs (rainfall, evapotranspiration, surface water elevation and salinity, groundwater elevation and salinity, etc.) is often difficult because of complex interactions between surface water, groundwater, and porewater in a variably saturated matrix with heterogeneous soils, vegetation, and topography. For example, depth, duration, frequency, and salinity of floodplain inundation are functions of tidal range; distance from the ocean; distance from the river channel; local elevation (microtopography); volume of freshwater flow; and the direction, volume, and salinity of groundwater fluxes [e.g., *Wang*, 1988; *Liu et al.*, 2001; *Melloul and Goldenberg*, 1997], as well as soil hydraulic characteristics and floodplain vegetation properties.

[3] In these complex systems, long-term monitoring can characterize the ranges and temporal variation of hydrological and water quality variables [e.g., *Muñoz-Carpena et al.*, 2008] and support the development of initial relationships between measured variables [e.g., *Kaplan et al.*, 2010]. However, investigating relationships between multivariate time series to improve understanding of system dynamics using visual inspection and comparative statistics is difficult, subjective, and may not appropriately characterize the system [*Ritter et al.*, 2007]. Nevertheless, a better understanding of hydrological dynamics is vital to the development of management scenarios to protect valued ecosystems, especially in modified wetland systems. Thus, an alternative method for identifying common trends and causal factors is required. This study applies dynamic factor analysis (DFA), a times series dimension reduction technique, to untangle complex groundwater dynamics in the floodplain wetlands of a managed coastal river in the southeastern United States.

[1] Agricultural and Biological Engineering Department, University of Florida, Gainesville, Florida, USA.
[2] Departamento de Suelos y Riegos, Instituto Canario de Investigaciones Agrarias (ICIA), Tenerife, Spain.
[3] Now at Departamento de Ingeniería, Producción y Economía Agraria, Universidad de La Laguna, La Laguna, Spain.

Copyright 2010 by the American Geophysical Union.
0043-1397/10/2009WR009038

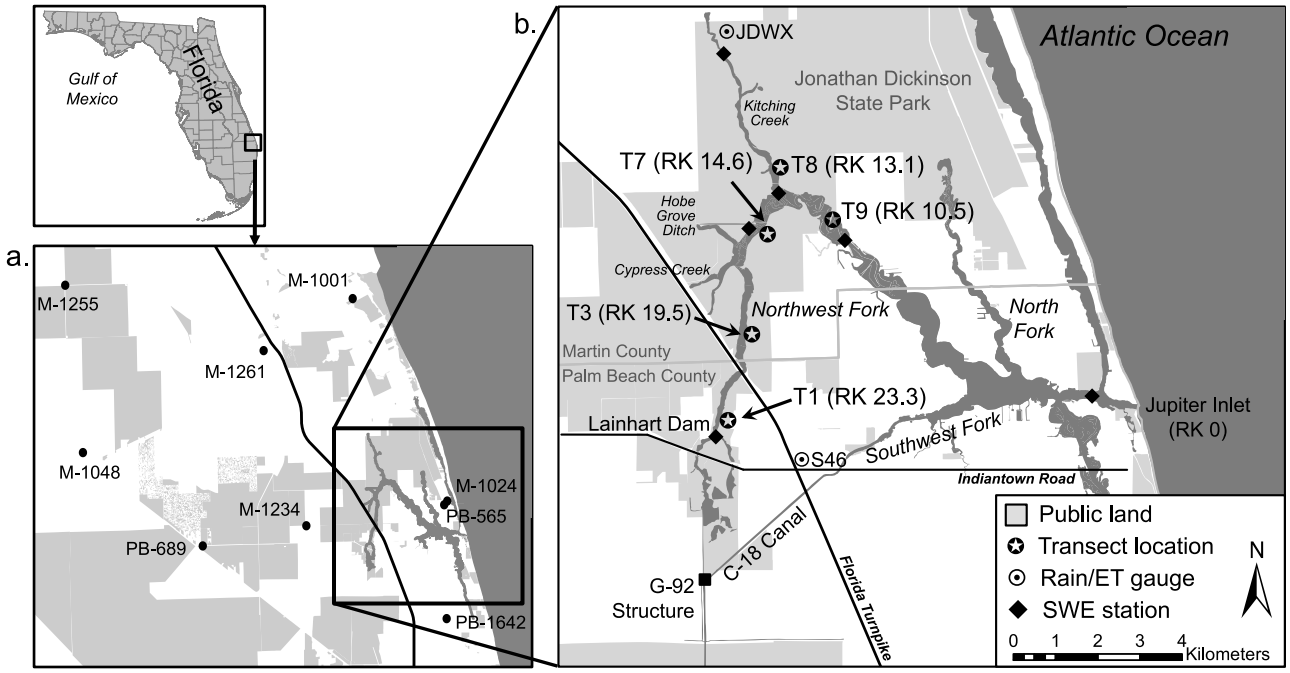

Figure 1. The Loxahatchee River and surrounding area, showing (a) the location of the nine regional USGS wells (WTE_R) used in this study and (b) transect locations (T1, T3, T7, T8, and T9), surface water elevation (SWE) and meteorological measurement locations, and major hydraulic infrastructure. Transect notation is followed by distance from river mouth (river kilometer, RK).

[4] DFA is a multivariate application of classic time series analysis originally developed for the interpretation of economic time series [*Geweke*, 1977] and can be a powerful tool for the modeling of short, incomplete, nonstationary time series in terms of common trends and explanatory variables [*Zuur et al.*, 2003a]. With DFA, underlying temporal variation in observed data (response variables) is modeled as linear combinations of common trends (unexplained variability), a constant level parameter, zero or more explanatory variables (additional observed time series), and noise [*Zuur et al.*, 2003b]. Like other time series models, DFA aims to maintain a good fit while minimizing the number of common trends. Using DFA, different formulations of dynamic factor models (DFMs) are possible, and thus, the best model selection is often made using goodness of fit indicators. The Nash and Sutcliffe coefficient of efficiency ($-\infty \leq C_{\text{eff}} \leq 1$; *Nash and Sutcliffe* [1970]) can be used to judge model performance. Additionally, Akaike's information criterion, AIC [*Akaike*, 1974] is often used as a decision tool for choosing between competing models [*Zuur et al.*, 2003b].

[5] The ability to model time series as a combination of common trends and explanatory variables is especially useful for analyzing complex environmental systems, where DFA can help assess what explanatory variables (if any) affect the time series of interest, and thus may be worthy of closer attention. DFA has been successfully applied in hydrology to identify common trends in groundwater levels [*Kovács et al.*, 2004; *Ritter and Muñoz-Carpena*, 2006], soil moisture dynamics [*Ritter et al.*, 2009], and interactions between hydrological variables and groundwater quality trends [*Muñoz-Carpena et al.*, 2005; *Ritter et al.*, 2007]. *Regalado and Ritter* [2009a; 2009b] used DFA for identifying common patterns of unexplained variability in soil water repellency measurements. It has also been used to identify trends and environmental variables affecting squid populations [*Zuur and Pierce*, 2004] and commercial fisheries [*Erzini*, 2005; *Tulp et al.*, 2008].

[6] Here, DFA is applied to study the interactions between floodplain groundwater elevations and other hydrological variables in the floodplain wetlands of the Loxahatchee River (Florida, USA), where reduced freshwater flow has led to saltwater intrusion and a transition to salt-tolerant, mangrove-dominated communities [*South Florida Water Management District* (*SFWMD*), 2006]. Groundwater plays a vital role in the water balance of many wetlands and is increasingly recognized as an important driver of ecological processes [*Hancock et al.*, 2009] and the development of particular ecological communities in wetlands from Florida, USA [*Harvey and McCormick*, 2009] to Australia [*Hatton and Evans*, 1998]. This is particularly true in coastal wetlands, where the effect of groundwater on salinity gradients can largely dictate habitat conditions in the estuary [*Jassby et al.*, 1995]. Except for one study of biogeochemical transport and submarine groundwater discharge [*Swarzenski et al.*, 2006], the role of groundwater in the Loxahatchee River and its floodplain is largely uninvestigated. The specific objectives of this research are to apply DFA to identify (1) important common trends among the time series (unexplained variability) and (2) the external local and/or regional hydrological factors (explained variability) that drive the observed shallow water table variation.

2. Study Site

[7] Historically, part of the greater Everglades watershed, the Loxahatchee River is located on the southeastern coast of Florida, USA (26°59′N, 80°9′W; Figure 1) and is often

referred to as the "last free-flowing river in southeast Florida" [*SFWMD*, 2006]. The river has three main branches (the North, Southwest, and Northwest Forks), which join in a central embayment that connects to the Atlantic Ocean via Jupiter Inlet. The watershed drains approximately 550 km^2 in Palm Beach and Martin counties and includes several large, publicly owned areas including Jonathan Dickinson State Park (JDSP), the Loxahatchee Slough Preserve, and the J.W. Corbett Wildlife Management Area. In 1985, a 15.3 km stretch of the Northwest Fork became Florida's first National Wild and Scenic River [*National Park Service* (*NPS*), 2004] and intensive data collection and modeling efforts in support of management and restoration planning have been underway for several years [e.g., *SFWMD*, 2006, 2006, 2009; *VanArman et al.*, 2005; *Mortl*, 2006; *Muñoz-Carpena et al.*, 2008, *Kaplan et al.*, 2010].

[8] The Northwest Fork of the Loxahatchee River (NW Fork) and its watershed contain a diverse array of terrestrial and aquatic ecosystems including sandhill, scrub, hydric hammock (a plant community characterized by 30–60 days of inundation yearly and mixed, facultative hardwood species), wet prairie, floodplain swamp, estuarine (mangrove) swamps, seagrass beds, tidal flats, oyster beds, and coastal dunes [*Roberts et al.*, 2006; *Treasure Coast Regional Planning Council* (*TCRPC*), 1999]. Many of these ecosystems remain relatively intact [*VanArman et al.*, 2005] and support a diversity of protected animal and plant species, including the endangered Florida manatee (*Trichechus manatus latirostris*) and four-petal pawpaw (*Asimina tetramera* Small) [*SFWMD*, 2006]. The upper watershed of the NW Fork is also home to one of the last remnants of bald cypress (*Taxodium distichum* [L.] Rich.) floodplain swamp in southeast Florida. However, changing hydrology and salinity regimes in the river and its floodplain have been linked to vegetative changes in the floodplain forest [*SFWMD*, 2002]. Of primary concern is the transition from bald cypress floodplain swamp to mangrove-dominated communities in the tidal floodplain as salinity increased and inadequate hydroperiod in the upstream riverine floodplain, which has shifted the system towards drier plant communities [*SFWMD*, 2009].

[9] Altered hydroperiods and encroaching salinity in the NW Fork have been linked to four major factors: (1) construction of major and minor canals that direct water away from the historic watershed; (2) the permanent opening of Jupiter Inlet in 1947 (Figure 1b); (3) construction of the C-18 canal in 1958, which transferred a majority of flow from the NW Fork to the Southwest Fork (Figure 1b); and (4) lowering of the regional groundwater table by community consumption [*SFWMD*, 2002]. These hydrologic changes have been linked to changes in the vegetative composition of the floodplain, where studies have documented the upriver retreat of bald cypress since at least the turn of the 20th century [*General Land Office*, 1855, Surveyor field notes from 1855 survey of the Jupiter/Loxahatchee River area, available at http://www.labins.org (verified 20 Aug. 2009), Labins, Tallahassee, Fla.; and more recently *Alexander and Crook*, 1975; *McPherson*, 1981; *Ward and Roberts*, 1996; *Roberts et al.*, 2008].

[10] The health of the Loxahatchee River and its adjacent ecosystems is a priority for many residents, visitors, agencies, and political leaders. As such, a number of planning efforts have been initiated over the past 20 years, including the North Palm Beach County Comprehensive Everglades Restoration Plan project (part of the 30 year, $8–$15 billion Everglades restoration project). Additionally, as in many other U.S. states [e.g., *Johnson*, 2008], Florida law requires its water management agencies to establish minimum flows and levels (MFLs) to protect water resources (section 373.042[1], Florida Statutes). MFL criteria are also designed to protect valued ecosystem components (VECs) from "significant harm." The MFL for the NW Fork [*SFWMD*, 2002] was adopted in 2003, and a restoration plan [*SFWMD*, 2006] was completed in 2006 with the goal of protecting the river's remaining cypress swamp and hydric hammock communities, as well as estuarine resources including oysters (*Crassostrea virginica*), fish larvae, and sea grasses, all identified as VECs. These MFL and restoration scenarios rely primarily on increased freshwater flow over Lainhart Dam (Figure 1b), which was found to be the most important driver of upstream hydroperiod and downstream surface water salinity. In spite of the Loxahatchee's "free-flowing" appellation, flow over Lainhart dam (calculated from headwater surface water elevation) is controlled by managing conveyance through the G-92 water control structure (Figure 1b).

3. Materials and Methods

3.1. Experimental Site and Setup

[11] In cooperation with the Florida Park Service, the SFWMD developed a network of 12 groundwater wells along five previously established vegetation survey transects perpendicular to the NW Fork (T1, T3, T7, T8, and T9; Figure 1b). Water table elevation (WTE) data were collected using TROLL 9000/9500 multiparameter water quality probes (In-Situ Inc., Ft. Collins, CO, USA) from September 2004 through January 2009. WTE data were measured every 30 min and converted to daily averages for this study. Upriver transects T1 and T3 each had one well, while transitional and tidal transects (T7, T8, and T9) had multiple wells to document differences in WTE across the floodplain (Figure 2). Table 1 summarizes well attributes. A full description of the groundwater dataset and quality assurance/quality control (QA/QC) procedure is available in *Muñoz-Carpena et al.* [2008].

[12] Transects 1 and 3 are upriver locations, not directly impacted by daily tides. T1 is located 23.3 km upstream of the river mouth (indicated as river kilometer, RK, 23.3) and has elevations ranging from 4.19 m (referenced to the National Geodetic Vertical Datum of 1929 [NGVD29]) on the top of a hydric hammock to 1.66 m in the river channel (Figure 2a). This freshwater transect is dominated by upland forest and hydric hammock at higher elevations and mature bald cypress swamp (average diameter at breast height, DBH = 49 cm) in the lower floodplain [*SFWMD*, 2006]. T3, located at RK 19.5, has several shallow braided streams in the floodplain and elevations ranging from 1.69 m in the floodplain to −3.00 m in the river channel (Figure 2b). This transect contains freshwater riverine swamp but is dominated by pop ash (*Fraxinus caroliniana* Mill.) with only four (very large) bald cypress in the canopy (average DBH = 92 cm). Intrusion of less flood-tolerant species into the riverine floodplain in these and other riverine transects has been documented, indicating the ecological impact of shortened hydroperiod in this area [*SFWMD*, 2006, 2009].

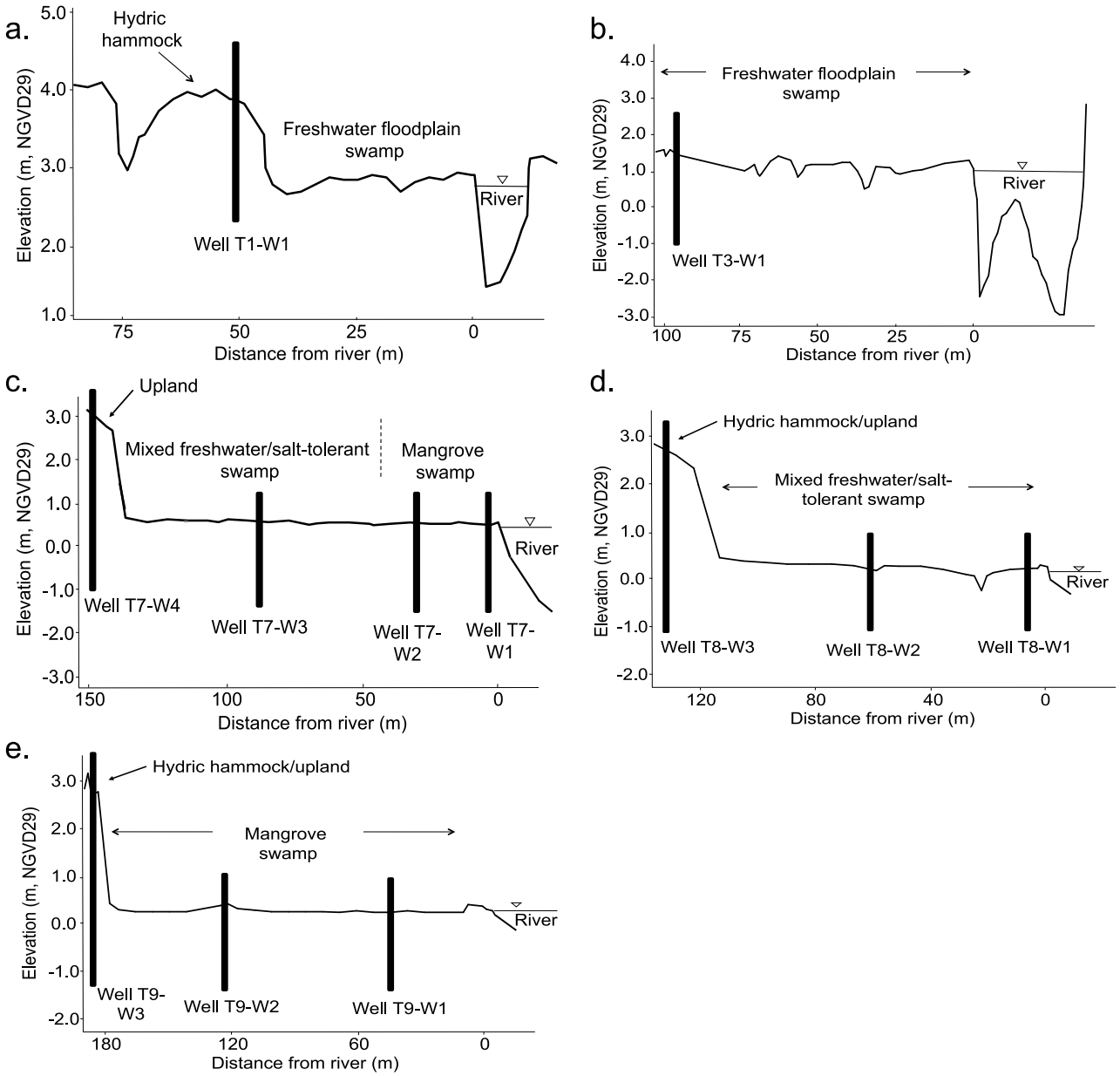

Figure 2. Transect topographic cross sections, detailing well installation locations and elevations and predominant vegetation types.

[13] Moving downriver, transects 7, 8, and 9 all receive daily tidal flooding of varying salinity over most or all of their length. T7 is in a transitionally tidal area (RK 14.6) and has elevations ranging from 3.07 m in the upland to 0.40 m in the floodplain (Figure 2c). Vegetation studies indicate that this transect has been impacted by saltwater intrusion, logging, and invasion by exotic plants [*SFWMD*, 2006]. T7 presently contains upper tidal swamp (dominated by red mangrove, *Rhizophora mangle* L.), which transitions to mixed riverine swamp approximately 30 m from the river channel. Transect 8 is located approximately 150 m upstream of the confluence of the NW Fork and Kitching Creek at RK 13.1. This transect has elevations ranging from 2.76 m in the upland to 0.23 m at the creek edge and transitions from hydric hammock in the uplands to upper tidal swamp in the floodplain (Figure 2d). The canopy is dominated by pond apple (*Annona glabra* L.), wax myrtle (*Myrica cerifiera* L.), and bald cypress, though red and white mangrove (*Laguncularia racemosa* [L.] C. F. Gaertn.) seedlings and subcanopy are present, especially within a braided channel with direct connection to the creek [*SFWMD*, 2009]. T9 is located at RK 10.5 on a small peninsula in the NW Fork and has elevations ranging from 2.89 m in the upland to 0.40 m at the river's edge (Figure 2e). This transect consists of lower tidal swamp, dominated by red and white mangrove except on an elevated trail, which supports sabal palm (*Sabal palmetto* [Walter] Lodd. Ex Schult. & Schult. f.). *Roberts et al.* [2008] documented intense vegetation changes on this transect, with a transition from freshwater to saltwater swamp species in less than 50 years.

Table 1. Locations and Attributes of the 12 Groundwater Wells in the Study[a]

Well	Transect Type	River Kilometer	Distance to River (m)	Well Elevation (m, NGVD29)	Well Depth (m, bgs)
T1-W1	Riverine	23.3	50	3.28	1.77
T3-W1	Riverine	19.5	95	1.60	1.76
T7-W1	Transitional	14.6	2	0.36	1.84
T7-W2			30	0.43	1.82
T7-W3			90	0.56	1.69
T7-W4			130	2.94	3.67
T8-W1	Transitional	13.1	5	0.12	1.62
T8-W2			65	0.36	1.60
T8-W3			105	2.28	2.64
T9-W1	Tidal	10.5	70[b]	0.41	1.86
T9-W2			50[b]	0.62	1.86
T9-W3			30[b]	2.94	4.24

[a]Wells are distributed across five transects (T1, T3, T7, T8, and T9). River kilometer indicates distance from the river mouth. Well depth is given in depth below ground surface (bgs).
[b]Shortest distance from well to river (T9 is on a peninsula).

3.2. Dynamic Factor Analysis

[14] DFA is based on structural time series models [*Harvey*, 1989] and aims to describe a set of N time series (termed response variables) using a dynamic factor model (DFM) that includes M common trends ($M < N$), K explanatory variables, a level or intercept parameter, and noise [*Lütkepohl*, 1991; *Zuur et al.*, 2003b]:

$$N \text{ time series} = M \text{ common trends} + \text{level parameter} + K \text{ explanatory variables} + \text{noise}. \quad (1)$$

In contrast to physically based or mechanistic models, DFMs are not built upon the underlying mechanisms of a given system but upon the common patterns among, and interactions between, response variables and explanatory factors. Thus, it requires no detailed information about the interactions between response and explanatory time series [*Ritter et al.*, 2009]. In the case presented here, this means that a complete a priori understanding of how groundwater (i.e., the response time series), surface water, and other hydrological variables interact in the floodplain is not necessary.

[15] By performing DFA, one or more common trends in the response time series are identified, which represent latent (unidentified) variation. The goal of DFA is to minimize the number of common trends (keeping M as small as possible) while still achieving a good fit. The use of explanatory variables can help improve the model fit and identify which environmental factors most affect the response variables. Equation (1) may be written in mathematical form as follows:

$$s_n(t) = \sum_{m=1}^{M} \gamma_{m,n}\alpha_m(t) + \mu_n + \sum_{k=1}^{K} \beta_{k,n}\nu_k(t) + \varepsilon_n(t) \quad (2)$$

$$\alpha_m(t) = \alpha_m(t-1) + \eta_m(t) \quad (3)$$

where $s_n(t)$ is a vector containing the set of N time series being modeled (response variables); $\alpha_m(t)$ [same units as the response variables] is a vector containing the mth common trend; $\gamma_{m,n}$ [dimensionless] are factor loadings or weighting coefficients, which indicate the importance of each of the common trends within the DFM; μ_n [same units as the response variables] is a constant level parameter; $\nu_k(t)$ [units vary] is a vector containing $0 - K$ explanatory variables; and $\beta_{k,n}$ [inverse units to convert explanatory variables to response variable units] are regression parameters, which indicate the importance of each of the explanatory variables within the DFM. In this study, N represents the 12 WTE time series. The terms $\varepsilon_n(t)$ and $\eta_m(t)$ [same units as the response variables] are independent, Gaussian noise with zero mean and unknown diagonal or symmetric/nondiagonal covariance matrix. Using a symmetric, nondiagonal matrix can lead to adequate model fits using fewer common trends than with a diagonal matrix but causes the number of parameters in the DFM to increase considerably [*Zuur et al.*, 2003a].

[16] Common trends, α_m(t), are modeled as a random walk [*Harvey*, 1989] and are predicted using the Kalman filter/smoothing algorithm and the expectation maximization (EM) technique [*Dempster et al.*, 1977; *Shumway and Stoffer*, 1982; *Wu et al.*, 1996]. Factor loadings ($\gamma_{m,n}$) and level parameters (μ_n) are also calculated using the EM technique. Regression parameters ($\beta_{k,n}$) are modeled using linear regression [*Zuur and Pierce*, 2004]. DFA deals with missing data in the response series by using a "design matrix" to identify missing observations and modify the factor loading, regression, and error matrices. The Kalman filter and smoothing algorithm then skips these missing observations [*Zuur et al.*, 2003b].

[17] Factor loadings ($\gamma_{m,n}$) and regression parameters ($\beta_{k,n}$) accompanying the common trends and explanatory variables allow for identification of the most relevant common trends and explanatory variables for each response variable. The magnitude of the $\beta_{k,n}$ and their associated standard errors were used to assess whether response and explanatory variables were significantly related (t value > 2). Additionally, cross-correlation between response variables and common trends was quantified using canonical correlation coefficients ($\rho_{m,n}$). Values of $\rho_{m,n}$ close to unity indicate that the common trend is highly associated with that response variable. In the following sections, "minor" correlations will refer to those with $|\rho_{m,n}| < 0.25$, "low" will refer to those with $0.25 \le |\rho_{m,n}| < 0.50$, "moderate" will refer to those with $0.50 \le |\rho_{m,n}| \le 0.75$, and "high" will refer to those with $|\rho_{m,n}| > 0.75$.

[18] One possible limitation of DFA, which has not previously been identified in the literature, is that common trends and explanatory variables are fit simultaneously. Thus, the method may identify one or more common trends that closely resemble candidate explanatory variables. If a common trend produced by DFA improves the model more than a similar explanatory variable, the resulting DFM will rely on the trend (which will have relatively high factor loadings) while overlooking the effect of the explanatory variable (which will have relatively low regression coefficients), potentially leading to spurious interpretation of results (i.e., deeming that an explanatory variable is unimportant). To ensure against this possibility, we calculated correlations between common trends and explanatory variables in DFMs that include both. High correlations between the series may indicate that an explanatory variable has been inappropriately disregarded in the model.

Table 2. Hydrological Time Series Used in the DFA[a]

Variable	Series Type	No. of Series	Description
WTE	Response	12	Groundwater table elevation (m, NGVD29) from wells in the Loxahatchee River floodplain
SWE	Explanatory	6	Surface water elevation (m, NGVD29) from stations in the Loxahatchee River at RK 23.3 (near T1), RK 14.6 (near T7), RK 13.1 (near T8), RK 9.5 (near T9), RK 1.1 (near Jupiter Inlet), and on Kitching Creek
R_{net}	Explanatory	3	Cumulative net local recharge (cumulative rainfall − cumulative ET_0, mm) calculated from weather stations at the S-46 structure and in Jonathan Dickinson State Park in the Loxahatchee River watershed (JDWX)
WTE_R	Explanatory	9	Groundwater table elevation (m, NGVD29) from USGS wells near the Loxahatchee River watershed

[a]See Figure 1 for locations.

3.3. Explanatory Variables

[19] Additional meteorological and hydrological variables were measured across the watershed, and a total of 29 daily time series (12 response variables and 17 candidate explanatory variables, each with 1589 daily values) were investigated for use in this analysis (Table 2). Not all candidate explanatory variables were used in the final DFMs since multicolinearity may exist between explanatory variables measured at nearby locations. The severity of multicolinearity was quantified using the variance inflation factor (VIF) of each set of explanatory variables [*Zuur et al.*, 2007]. Combinations of explanatory variables that resulted in VIF > 5 were avoided in these analyses [*Ritter et al.*, 2009]. In cases where this criterion was exceeded, the single candidate explanatory variable that best minimized AIC and maximized C_{eff} was selected. For example, if using two surface water elevation time series as explanatory variables resulted in VIF > 5, the series that yielded the poorer modeling results was discarded.

[20] Breakpoint rainfall data were recorded at the SFWMD S-46 structure on the Southwest Fork of the Loxahatchee River and at the weather station in Jonathan Dickinson State Park (station code IDWX), where daily reference evapotranspiration (ET_0) values were also measured (Figure 1b). These data are publicly available and were downloaded from the SFWMD's online database DBHYDRO (Stations S46_R and JDWX; accessed at http://www.sfwmd.gov/org/ema/dbhydro/index.html). Note that WTE data are autocorrelated (i.e., WTE at time t is dependent on WTE at $t - 1$), while this is not true for rainfall and ET, which contain no consistent information about previous data. In order to make rainfall and ET_0 data useful to the DFA [*Ritter et al.*, 2009], the difference between cumulative rainfall and cumulative ET_0 was used to create two net local recharge (R_{net}) time series. Rainfall was measured at the S-46 and JDWX gauging stations (Figure 1b), but ET_0 was only computed from measurements at JDWX, so cumulative ET from this station was used to calculate both R_{net} series ($R_{net,S46}$ and $R_{net,JDWX}$). Though only 11.2 km apart, the two rain stations exhibited large differences in cumulative rainfall over the 4 year study period. The effect of this spatial variability on model results was explored by developing DFMs using each of the R_{net} series, both series, and their average and comparing model results.

[21] Surface water elevation (SWE) data were recorded at five locations in the NW Fork and one station upstream on Kitching Creek (Figure 1b). A SFWMD monitoring station on the headwater side of Lainhart Dam (0.45 km upstream of T1) measured average daily SWE and is available on DBHYDRO (station LNHRT_H; Figure 1b). Cooperatively monitored United States Geological Survey (USGS)/SFWMD stations located at RK 14.6 (adjacent to T7), RK 13.1 (at confluence with Kitching Creek, near T8), RK 9.5 (~0.8 km downstream of T9), RK 1.1 (near the Jupiter inlet), and 2.8 km upstream of the confluence of the NW Fork with Kitching Creek each measured SWE every 15 min (Figure 1b). These data were acquired from USGS staff and converted to daily averages.

[22] Daily average WTE from nine USGS wells (Figure 1a) in and around the Loxahatchee River watershed (denoted as WTE_R) are publicly available and were downloaded from the USGS National Water Information System (accessed at http://waterdata.usgs.gov/nwis/). An initial cross-correlation analysis identified possible lead/lag relationships between WTE and WTE_R series, and candidate WTE_R series were lagged from +3 to −3 days to determine if lagged series improved the final DFM.

3.4. Analysis Procedure

[23] DFA was implemented using the Brodgar v. 2.5.7 statistical package (Highland Statistics Ltd., Newburgh, UK), which is based on the statistical software language "R" version 2.9.1 [*R Core Development Team*, 2009]. Response and explanatory variables were normalized (mean subtracted, divided by standard deviation) in Brodgar. This allowed us to compare the relative importance of explanatory variables across the set of response variables [*Zuur et al.*, 2003b; *Zuur and Pierce*, 2004]. The DFA was carried out sequentially and resulted in three models (Table 3). First, DFMs were built with an increasing number of common trends until satisfactory model performance was achieved according to goodness-of-fit indicators [*Zuur et al.*, 2003a]. This DFM is referred to as Model I. Once the minimum number of common trends (*M*) was identified, different combinations of explanatory variables were incorporated until a satisfactory combination of common trends and explanatory variables was identified without exceeding the VIF criterion (Model II). This reduced the unexplained variability and improved description of WTE in the floodplain. Finally, a reduced model was explored by removing common trends and using only the best subset of explanatory variables identified in the DFA to create a multi-linear model (Model III) using a multiple regression code run in Matlab [2009b, The MathWorks, Inc., Natick, MA, USA].

Table 3. Dynamic Factor Models Tested in This Study[a]

DFM	No. of Trends	Explanatory Variables	Regression Parameters	No. of Parameters	C_{eff}	AIC
Model I	6	None	—	81	0.94	4840
Model II	3	$SWE_{RK23.3}$, $SWE_{RK14.6}$, WTE_R_{M1001}, $R_{net,S46}$, $R_{net,JDWX}$	From DFA	117	0.91	2998
Model III	0	$SWE_{RK23.3}$, $SWE_{RK14.6}$, WTE_R_{M1001}, $R_{net,S46}$, $R_{net,JDWX}$	Multiple regression	60	0.81	22,063

[a]See explanation in text.

[24] Goodness of fit was assessed by visual inspection of the observed versus predicted WTE and quantified with the Nash Sutcliffe coefficient of efficiency ($-\infty \leq C_{eff} \leq 1$, *Nash and Sutcliffe* [1970]) and Akaike's information criterion [*Akaike*, 1974]. C_{eff} compares the variance about the 1:1 line to variance of the observed data, with $C_{eff} = 1$ 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. For two different DFMs, the DFM with largest C_{eff} and smallest AIC was preferred.

4. Results and Discussion

4.1. Experimental Time Series

[25] The hydrological data collected during this study represent a wide range of climatic conditions, including four wet/dry season cycles, two wet years with hurricane-induced flooding (2004–2005), and the driest 2 year period (2006–2007) recorded in south Florida since 1932 (C. Neidrauer, 2009, Water Conditions Summary, available at http://www.sfwmd.gov/portal/page/portal/pg_grp_sfwmd_governingboard/porlet_gb_subtab_presentations_page/tab20092120/3%20%20water%20conditions.pdf (verified 21 September 09), South Florida Water Management District, Operations Control Dept., West Palm Beach, Fla.). Daily time series of hydrological variables are presented in Figure 3. Rainfall (Figure 3a) followed a seasonal pattern, with wet season (May–October) rain accounting for 73%–80% of yearly totals over the 4 years (mean, 77%; Figure 3a). This is in agreement with previous seasonal rainfall observations in the Loxahatchee River Basin, which have shown that approximately two thirds of yearly rain falls in the wet season between May and October [*Dent*, 1997]. Significant spatial variation between rainfall data collected at the S-46 and JDWX stations was also found. Though the rain gauges were only 11.2 km apart and roughly equidistant from the shore in flat terrain, cumulative rainfall at the JDWX gauge in JDSP was 2151 mm greater than that at the S-46 structure over the 4 year study period, yielding divergent R_{net} series (Figure 3b). Correlation between the rainfall time series was also low ($r^2 = 0.18$), further justifying the use of both series in the DFMs developed (see following sections). Both rainfall series passed QA/QC procedures by the SFWMD and were deemed reliable.

[26] WTE was variable across wells and transects, as well as over seasons and years. WTE ranged from a maximum of 3.80 m upstream at well T1-W1 to a minimum of −0.88 m in the tidal floodplain of T8 (T8-W1). In general, WTE was highest in upriver wells (T1-W1, T3-W1) and downriver in higher elevation wells (T7-W4, T8-W3) (Figure 3c). Visual inspection of WTE in these higher elevation wells suggested common trends associated with wet and dry season rainfall patterns. For example, the impact of late season rains in 2004 and 2005 and dry summers in 2006 and 2007 on the WTE are apparent across these wells. WTE in lower elevation wells closer to the river appeared to be more influenced by daily tidal flooding (Figure 3d). Some seasonal wet/dry patterns were still apparent, but less so, as the signal was damped by daily and monthly tidal fluctuations. Note high water events in September 2004 during hurricanes Frances and Jeanne. WTE generally decreased from upstream (T1) to downstream (T9). One exception to this is well T7-W4, which maintained higher WTE than well T3-W1 (which is further upstream) throughout most of the period of record (Figure 3c) due to its high elevation (Table 1).

[27] Figures 3e–3f show temporal variation in the six SWE and nine WTE_R series explored as candidate explanatory variables. In the upriver SWE series (RK 23.3 and Kitching Creek), large rainfall events coincided with peaks in SWE, and distinct drawdowns during each of the four dry seasons were observed (most drastically in 2006 and 2007). Daily average SWE measured near T7, T8, T9, and the Jupiter Inlet were nearly identical and overlap in the figure ($0.94 \leq r^2 \leq 0.99$ for these four series; Figure 3e). WTE_R series measured in and around the Loxahatchee River exhibited a large range (from close to sea level to over 10 m) but consistently mirrored the wet and dry season variations observed in the two upriver SWE series (Figure 3f).

[28] Other WTE trends in the floodplain of the Loxahatchee River became apparent when looking closely at a single transect. For example, WTE increased with distance from the river on T7, with the highest elevation well (T7-W4) showing the maintenance of much higher WTE (Figure 4). During the dry seasons of 2006 to 2007, this freshwater [*Muñoz-Carpena et al.*, 2008] head fell, approaching but not reaching WTE in the floodplain. This indicates a variable but consistently positive flow of freshwater from the upland toward the river even in extremely dry seasons, which likely plays a role in mitigating the severity of saltwater intrusion on T7. This highlights the importance of understanding the dynamics of this hydrological flux. The same pattern was apparent on T8 (not shown), with higher WTE maintained in well T8-W3, except for the dry seasons of 2006 and 2007, when the groundwater levels in T8-W2 and T8-W3 met during an extreme WTE drawdown. At T9, which is on a peninsula with the river on two sides, these patterns were not as apparent, with higher elevation and lower elevation wells sharing similar WTE (not shown).

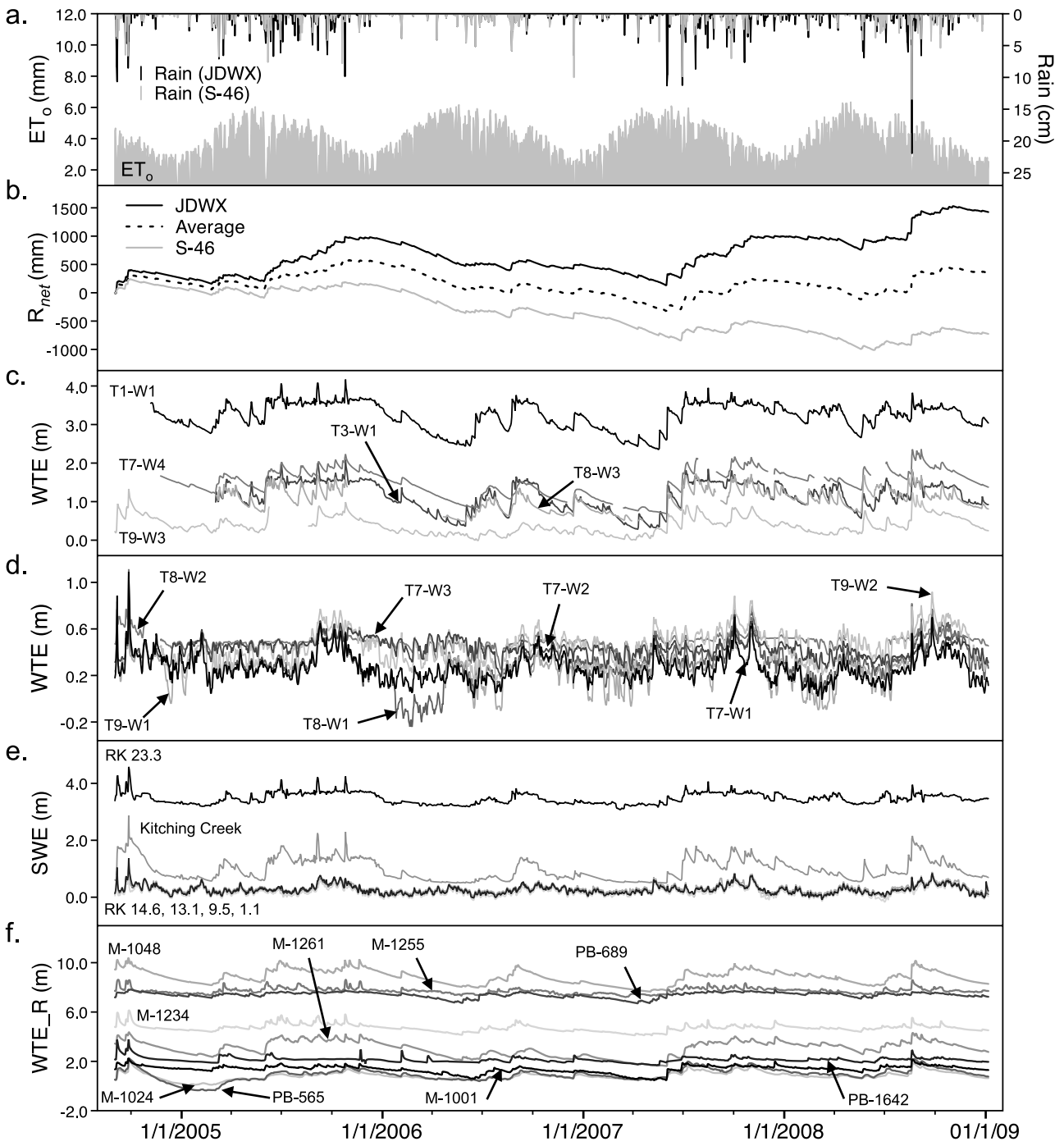

Figure 3. Precipitation, reference evapotranspiration (ET_0), calculated net local recharge (R_{net}), water table elevation (WTE), surface water elevation (SWE), and regional water table elevation (WTE_R) measured in and around the Loxahatchee River watershed. Gaps in times series in Figure 3c represent missing data.

4.2. Dynamic Factor Analysis

4.2.1. Baseline DFA (No Explanatory Variables)

[29] The DFA was advanced in three discrete steps. First, different DFMs were obtained by increasing the number of common trends until a maximum C_{eff} and minimum AIC were achieved. With a diagonal matrix, AIC was minimized and C_{eff} maximized with six trends (M = 6; Table 4). The minimized AIC of 4840 and maximized C_{eff} of 0.94 using six common trends (Model I) were then used as targets for subsequent DFMs. That six common trends (representing unexplained, but shared, information) were necessary to achieve the best DFM with no explanatory variables sug-

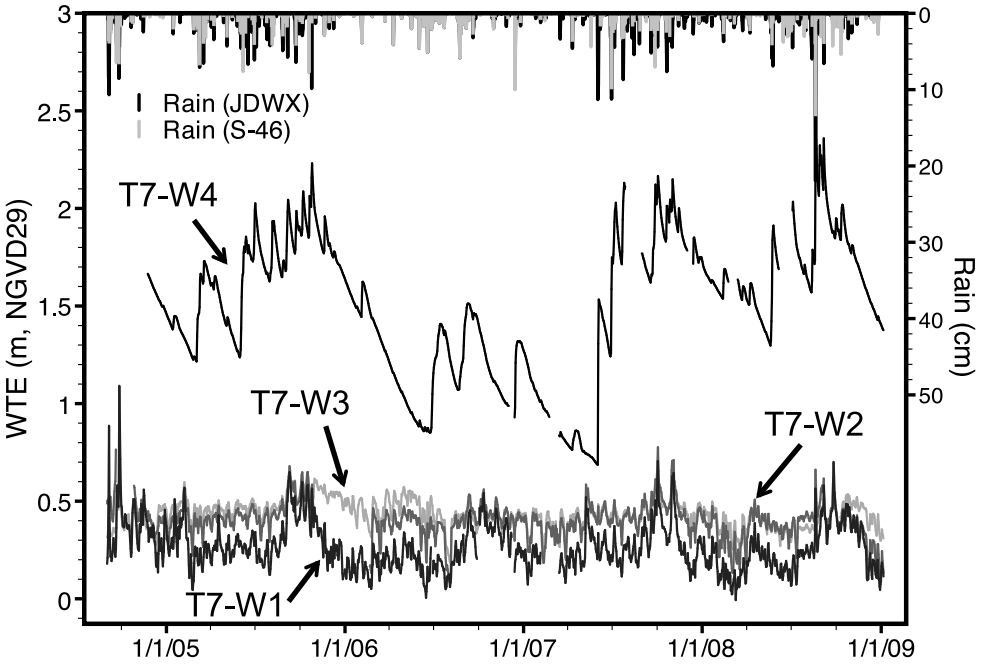

Figure 4. Average daily water table elevation (WTE) in the four wells on transect 7 (T7). Gaps in time series represent missing data.

gests that several latent effects influence the variability of WTE across the watershed. It is instructive to examine these common trends and their associated canonical correlation coefficients ($\rho_{m,n}$) since high $\rho_{m,n}$ values indicate high correlation between two latent variables.

[30] Three example trends from Model I with high $\rho_{m,n}$ values are shown in Figure 5. Though only describing latent (unknown) variability at this stage in the DFA, these trends and their patterns of correlation are useful for developing ideas about how WTE varies in the Loxahatchee River floodplain and where to look for the most useful explanatory variables. For example, the trend in Figure 5a was highly to moderately correlated (positively) with all five higher elevation wells (T1-W1, T3-W1, T7-W4, T8-W3, and T9-W3) but relatively unimportant (minor and low correlations) for the seven lower floodplain wells (T7-W1, T7-W2, T7-W3, T8-W1, T8-W2, T9-W1, and T9-W2). On the other hand, the trend in Figure 5b was negatively correlated with the upland and upriver wells (low to minor correlations) but positively and more strongly correlated with floodplain wells (moderate to low correlations). This geographic and topographic distribution of $\rho_{m,n}$ values across the twelve wells suggests that the use of explanatory variables that represent distinct parts of the river may help reduce the unexplained variability represented by these trends. The trend in Figure 5c had low to moderate correlation with two of the twelve wells, both on T8, and the correlations were in opposite directions. The rest of the correlations are minor. This indicates a latent effect specific to these wells and could be an indicator of anomalous data or an unidentified environmental factor (or factors) that only affects these wells (i.e., pumping in the area). Model I required this common trend to achieve the best match of WTE data in these wells. From the remaining three common trends (not shown), no clear spatial or physical interpretations could be drawn.

4.2.2. DFA With Explanatory Variables

[31] Next, explanatory variables were added in an attempt to reduce the number of common trends required to achieve an adequate fit of WTE (and to reduce the canonical correlation coefficients and factor loadings of remaining trends). Candidate explanatory variables included surface water elevations (SWE) at six locations in the NW Fork, regional groundwater elevations (WTE_R) in nine groundwater wells in and around the Loxahatchee River watershed, and net local recharge (R_{net}) calculated from two rain gauges and one ET monitoring station, for a total of 17 possible explanatory time series. When two or more candidate explanatory variables were colinear or multicolinear (resulting in VIFs > 5), the explanatory variable resulting in the best overall model fit (highest C_{eff} and lowest AIC) was selected.

[32] For the SWE time series, upstream river stage at Lainhart Dam (RK 23.3) and tidal river stage at RK 14.6 provided the best benefit to the model and were not col-

Table 4. Akaike's Information Criteria (AIC) and Nash-Sutcliffe Coefficients of Efficiency (C_{eff}) for Dynamic Factor Models With no Explanatory Variables and 1–7 Common Trends (M)[a]

M	C_{eff}	AIC
1	0.44	32,204
2	0.80	19,860
3	0.85	15,390
4	0.89	11,211
5	0.90	7337
6	**0.94**	**4880**
7	0.93	6875

[a]Best model is rendered in bold.

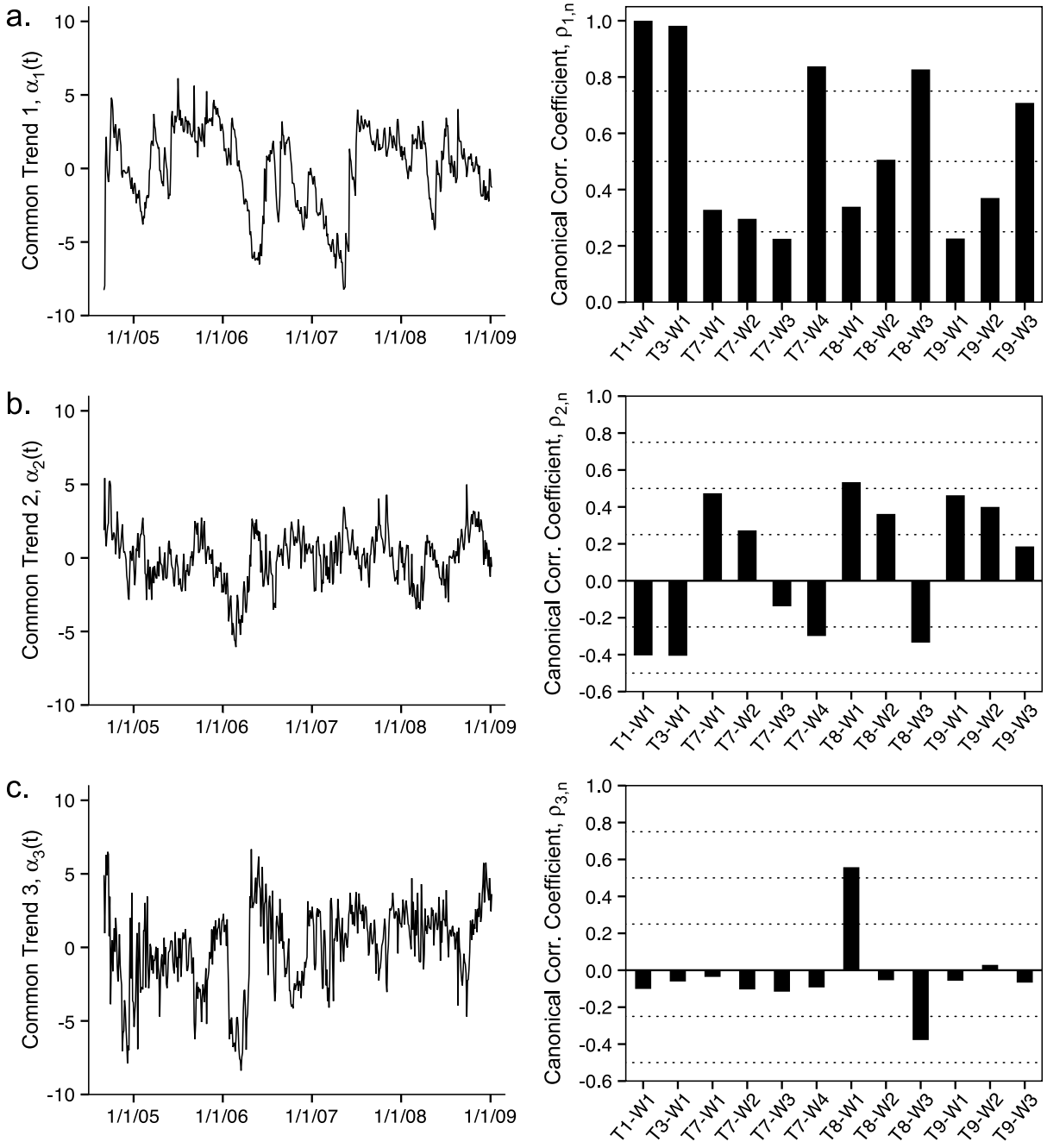

Figure 5. Three example normalized trends from Model I (left) and their associated canonical correlation coefficients (right). (a) Trend 1 shows high correlation to higher elevation and upstream wells; (b) trend 2 is most associated with lower elevation floodplain wells; and (c) trend 3 has low correlations except for wells T8-W1 and T8-W3.

linear. That both upriver and tidal SWE series were required for the best DFM follows from Model I, whose trends were split across high and low elevation floodplain wells. For WTE_R time series, USGS well M-1001 most improved the model. The DFM was not improved by lagging any of the WTE_R series by −3 to +3 days. The model was also improved by using both net local recharge series ($R_{net,S46}$ and $R_{net,JDWX}$) compared with either series alone or their average. Recorded rainfall at the S-46 and JDWX gauges were drastically different (Figure 3b), and thus, when used to calculate R_{net}, each series had distinct information that improved the DFM. The use of both R_{net} series also highlighted the effects of the high spatial variability of rainfall in the region. The VIFs for this set of five explanatory variables did not exceed the VIF threshold ($1.30 \leq \text{VIF} \leq 2.51$). Correlations between common trends and this set of explanatory variables were low ($0.001 \leq r^2 \leq 0.34$), indicating that the importance of these explanatory variables was not masked by common trends (though we did observe higher correlations between trends and explanatory variables

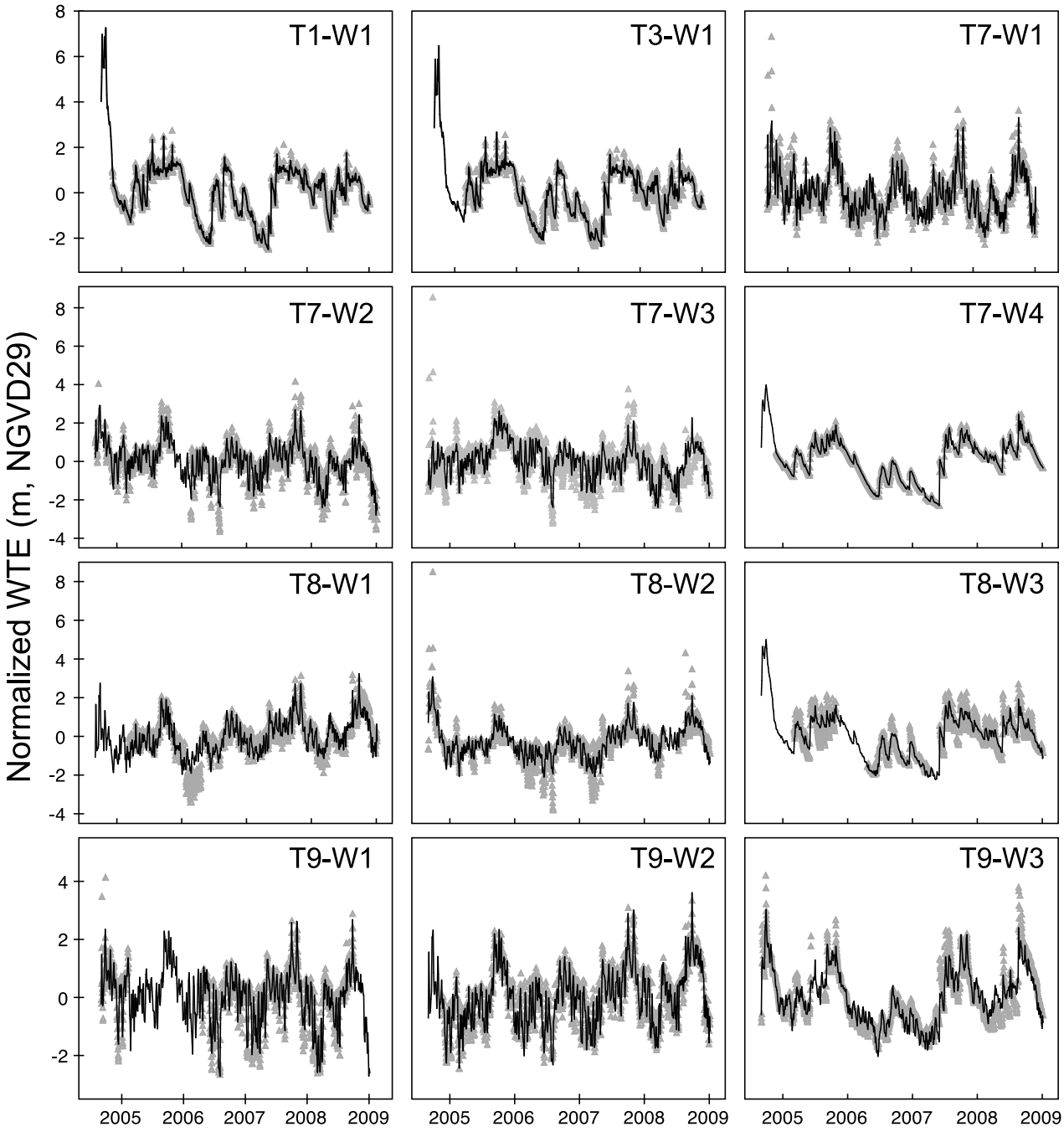

Figure 6. Observed (gray symbols) and modeled (black lines) normalized WTE for the 12 wells obtained from Model II using three common trends and five explanatory variables.

in other DFMs, confirming that it is important to check for this correlation before interpreting DFA results).

[33] In summary, the best DFM used five explanatory variables (K = 5): SWE at RK 23.3 and 14.6 ($SWE_{RK23.3}$ and $SWE_{RK14.6}$), WTE from USGS well M-1001 (WTE_R_{M1001}), and both net local recharge series ($R_{net,S46}$ and $R_{net,JDWX}$). With these explanatory variables, the number of required common trends was reduced from six 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 AIC value of 2998 (lower than the 4880 target from Model I) and a C_{eff} value of 0.91 across the 12 wells (compared with the target of 0.94). Model fits are illustrated in Figure 6. Model fits are good to excellent ($0.78 < C_{eff} < 1.0$). Some higher elevation wells lack data from the beginning of the time series, and model results help paint a more complete picture of WTE in these wells during the hurricanes of 2004 (e.g., frames 1 and 2 in Figure 6).

[34] Table 5 summarizes the results obtained from Model II ($M = 3, K = 5$). Significant regression parameters (t value > 2) are shown in bold. WTE in the 12 wells in the Loxahatchee

Table 5. Constant Level Parameters (μ_n), Canonical Correlation Coefficents ($\rho_{m,n}$), Factor Loadings ($\gamma_{m,n}$), Regression Coefficients ($\beta_{k,n}$), and Coefficients of Efficiency (C_{eff}) From Model II[a]

		Canonical Corr.			Factor Loadings			Regression Coefficients ($\beta_{k,n}$)					
s_n	μ_n	$\rho_{1,n}$	$\rho_{2,n}$	$\rho_{3,n}$	$\gamma_{1,n}$	$\gamma_{2,n}$	$\gamma_{3,n}$	$SWE_{RK23.3}$	$SWE_{RK14.6}$	$R_{net,S46}$	$R_{net,JDWX}$	WTE_R_{M1001}	$C_{eff,n}$
T1-W1	−0.44	0.61	0.10	−0.30	0.08	0.02	0.00	**0.53**	−0.01	**0.73**	0.10	**0.22**	1.00
T3-W1	−0.26	0.52	0.14	−0.31	0.05	0.02	0.00	**0.58**	0.01	**0.56**	0.13	**0.19**	0.97
T7-W1	−0.10	0.19	0.23	0.02	0.00	−0.01	0.04	**−0.07**	**0.92**	**0.23**	0.05	**0.16**	0.94
T7-W2	0.58	0.16	0.35	0.26	0.00	0.02	0.18	0.02	**0.62**	**0.71**	**0.74**	−0.05	0.90
T7-W3	0.22	−0.08	0.09	−0.04	−0.01	−0.02	0.20	**0.12**	**0.47**	**1.05**	**1.00**	−0.12	0.83
T7-W4	1.20	0.27	0.43	−0.38	0.00	0.10	−0.01	**0.10**	−0.01	0.09	**0.77**	**0.24**	1.00
T8-W1	−0.14	0.05	0.30	0.27	0.00	−0.01	0.03	**0.19**	**0.74**	**−0.26**	0.00	**0.12**	0.78
T8-W2	0.01	0.45	0.36	0.04	0.03	0.01	0.12	**−0.08**	**0.43**	**0.54**	**0.72**	**0.19**	0.80
T8-W3	0.90	0.35	0.55	−0.14	0.02	0.10	0.00	**0.21**	**−0.09**	0.3	**0.65**	**0.2**	0.88
T9-W1	0.50	0.16	0.31	0.29	0.00	0.00	0.23	**−0.12**	**0.58**	**0.92**	**1.03**	0.06	0.97
T9-W2	0.09	0.02	0.30	0.24	0.01	0.00	0.16	**0.06**	**0.70**	**0.49**	**0.84**	0.01	0.98
T9-W3	1.06	0.35	0.54	−0.16	−0.01	0.07	0.03	**0.12**	**0.27**	0.03	**0.70**	**0.16**	0.86
												Overall	0.91

[a]Model II includes three trends and five explanatory variables. Significant regression parameters are rendered in bold.

River had variable relationships to the common trends from Model II, but canonical correlations were reduced from Model I, indicating a reduced dependence of the DFM on these latent series. The trends in Model II had zero "high" and four "moderate" correlations with response variables, compared to four "high" and seven "moderate" correlations in Model I.

[35] The spatially distributed effects of the explanatory variables and common trends on Model II are compared in Figure 7. Regression parameters ($\beta_{k,n}$; Figures 7a–7e) represent the relative importance of each explanatory variable to each response time series, with black bars indicating significant regression parameters by t test. In general, inclusion of explanatory variables in Model II reduced factor loadings (Figure 7f) over those in Model I (overall average $|\gamma_n|$ for the six trends in Model I was 0.13 ± 0.16 compared to 0.05 ± 0.04 in Model II), suggesting that the patterns observed in the Loxahatchee River floodplain wells may be adequately described using only the selected explanatory variables (see following section).

[36] Visualizing the spatial distribution of the importance of each explanatory variable in the floodplain can be useful when assessing river management options. For example, Figure 7a shows that the Lainhart Dam surface water time series ($SWE_{RK23.3}$) was most important in describing variability in wells T1-W1 and T3-W1 but had reduced impact downriver. As the major management tool in the NW Fork, river stage (i.e., flow) at Lainhart Dam had only limited impact in maintaining WTE downstream of T3. Similarly, Figure 7b demonstrates the strong importance of tidal surface water ($SWE_{RK14.6}$) in lower elevation wells further downstream. This variable was most important for explaining WTE variability on downstream transects (T7, T8, and T9) and was strongest for those wells closest to the river, decreasing with distance from the river, for example, from T7-W1 (strongly significant, with $\beta = 0.92$) to T7-W4 (insignificant, with $\beta = 0.01$). This explanatory variable, and by extension the response variables that it influences most, is most susceptible to sea level rise.

[37] Figures 7c–7d show regression parameters for the two net local recharge series ($R_{net,S46}$ and $R_{net,JDWX}$). Though the importance of these two series was distributed across the 12 wells in the floodplain, a geographic pattern is apparent. Wells T1-W1 and T3-W1 are closer to the rainfall gauging station at the S-46 structure (3.2 and 3.9 km, respectively) than the JDWX gauging station (9.7 and 7.2 km, respectively). These wells were more strongly affected by $R_{net,S46}$ (significant, with β values of 0.73 and 0.56, respectively) than by $R_{net,JDWX}$ (insignificant β values of 0.10 and 0.13, respectively). The importance of the two net local recharge series were split fairly equally over the remainder of the wells (average $\beta_{Rnet,S46}$: 0.46 ± 0.34; average $\beta_{Rnet,JDWX}$: 0.65 ± 0.35), with $\beta_{Rnet,JDWX}$ being slightly more important in describing the downstream wells. The importance of capturing this spatially distributed rainfall is reinforced when building a DFM using just one of the R_{net} series or the average of the two, which yielded poorer results ($4859 \leq AIC \leq 6998$; $0.88 \leq C_{eff} \leq 0.89$).

[38] Figure 7e shows that highest β values for WTE_R were associated with upstream wells (T1-W1, T3-W1) and downstream, high elevation wells (T7-W4, T8-W3, and T9-W3), whose time series closely resembled regional groundwater circulation. The importance of regional groundwater elevation (WTE_R_{M1001}) increased with well elevation and was significant for 9 of the 12 wells. A lowered regional groundwater table has been identified as a cause of reduced hydroperiod and increased saltwater intrusion in the Loxahatchee River [*SFWMD*, 2002], and the dependence of floodplain WTE on regional groundwater is substantiated by these results. It is interesting to note that, although the regional groundwater trend and SWE at Lainhart Dam were correlated ($r^2 = 0.71$), including both explanatory variables in Model II allowed us to decompose the general effect of the regional groundwater circulation from the more local effect of SWE at Lainhart Dam shown in Figure 7a.

[39] The remaining three trends in Model II and their associated $\rho_{m,n}$ values are given in Figure 8. These common trends represent the remaining unexplained (latent) variability among the WTE series. Common trend 1 has a high starting value, likely associated with high water events during the hurricanes of 2004, which may not be sufficiently described by explanatory variables, especially if measurement errors occurred during these extreme events. This trend was most important to wells T1-W1 and T3-W1, which were also most strongly affected by SWE at Lainhart Dam. WTE in all wells were generally positively correlated with

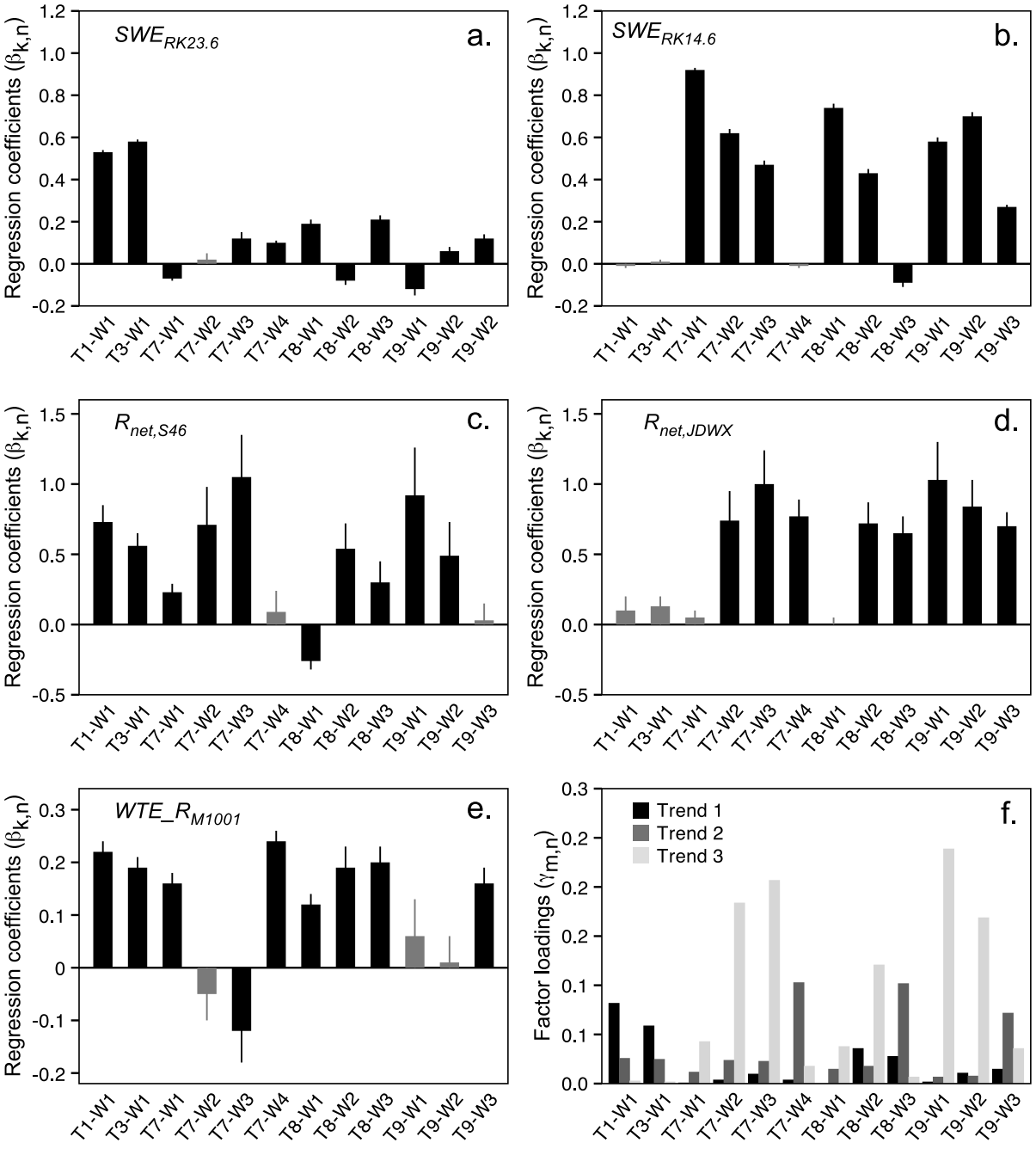

Figure 7. (a–e) Regression parameters and (f) factor loadings for Model II ($M = 3$, $K = 5$). Regression parameters are shown with their standard errors, with black bars indicating significance.

both common trends 1 and 2 but have low correlations (average $\rho_{1,n}$ value: 0.25 ± 0.21; average $\rho_{2,n}$ value: 0.31 ± 0.15). Common trend 3 is weaker and less consistent, with positive correlations for most floodplain wells and negative correlation for most upland wells, all of which were either "minor" or "low." Correlations between these remaining trends and the explanatory variables in Model II were also low, confirming that the importance of explanatory variables was not masked by common trends. Though not pursued here, in cases where common trends do mask explanatory variables, it may be possible to use a two-step process wherein the "ignored" explanatory variable(s) are fit to the data first using linear regression. DFA could then be run on the residuals of this process as previously described. This method would essentially be a time series-based partial regression technique [e.g., *Zuur et al.*, 2007].

4.2.3. Multilinear Regression Model (DFA With No Common Trends)

[40] Finally, common trends were removed from the model to assess the validity of a DFM using only explanatory variables. In this model (Model III), the five explanatory variables identified in the DFA were used to create a multilinear model of the response variables. As expected, C_{eff} values for Model III were somewhat reduced from

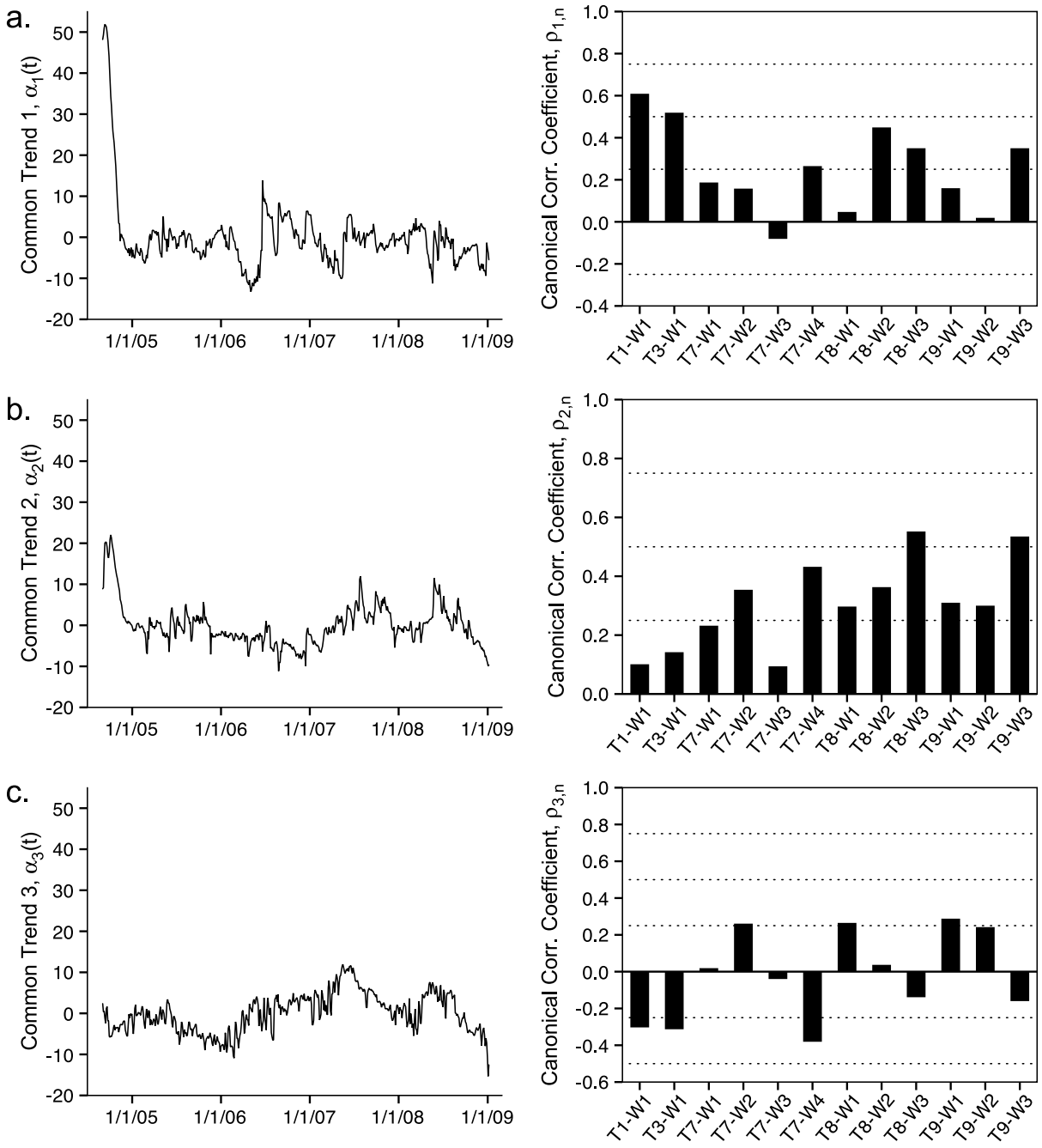

Figure 8. (left) Common normalized trends and (right) their associated canonical correlation coefficients for Model II.

Table 6. Model Parameters and Coefficients of Efficiency (C_{eff}) From Model III[a]

	Model Parameters					
s_n	$SWE_{RK23.3}$	$SWE_{RK14.6}$	WTE_R_{M1001}	$R_{net,S46}$	$R_{net,JDWX}$	C_{eff}
T1-W1	**0.69**	**−0.09**	**0.41**	**0.07**	−0.02	0.91
T3-W1	**0.70**	**−0.06**	**0.35**	**0.08**	0.00	0.94
T7-W1	**−0.07**	**0.95**	**0.09**	**0.08**	**−0.05**	0.93
T7-W2	**0.07**	**0.86**	**0.09**	**−0.05**	**−0.31**	0.76
T7-W3	**0.13**	**0.65**	**−0.32**	**0.42**	**0.30**	0.59
T7-W4	**0.18**	**0.03**	**0.68**	**0.07**	**0.23**	0.91
T8-W1	**0.18**	**0.78**	**0.07**	**−0.38**	**−0.09**	0.80
T8-W2	**0.06**	**0.55**	**0.35**	**−0.12**	0.01	0.68
T8-W3	**0.34**	**−0.04**	**0.69**	**−0.05**	**−0.04**	0.81
T9-W1	**−0.12**	**0.87**	**0.10**	**−0.06**	**−0.11**	0.81
T9-W2	**0.12**	**0.87**	**−0.04**	**−0.17**	**0.15**	0.86
T9-W3	**0.14**	**0.38**	**0.50**	−0.01	**0.05**	0.77
					Overall	0.81

[a]No trends, five explanatory variables. Significant model parameters are rendered in bold.

Model II (overall $C_{eff} = 0.81$, $0.59 < C_{eff} < 0.94$; compared to $C_{eff} = 0.91$, $0.78 < C_{eff} < 1.0$ for Model II) but were still adequate for most wells (Table 6). Model III accurately predicted WTE series in higher elevation wells farthest from the river (e.g., Figure 9, well T3-W1) and in lower elevation wells close to the river (e.g., Figure 9, well T7-W1) but performed worse for middle distance and elevation wells (e.g., Figure 9, well T7-W3). Closer to the edges of the system, explanatory variables act as boundary conditions (e.g., regional WTE at the farthest landward end of transects and SWE acting at the river), and their effects can be seen directly in the WTE series. In middle distance and middle elevation wells, the interaction of surface water and groundwater is most complex and nonlinear, which may not be as well captured by a linear combination model. Despite these limitations, overall performance of Model III was adequate to describe variations in WTE in the Loxahatchee River floodplain and may be useful for assessment of Loxahatchee River restoration scenarios [*SFWMD*, 2006], especially considering the wide range of climatic conditions captured in the study. In general, all restoration scenarios rely primarily on increased freshwater flow over Lainhart Dam, which was found to be the most important driver of upstream hydroperiod and downstream surface water salinity. The Restoration Plan for the Northwest Fork of the Loxahatchee River [*SFWMD*, 2006] identified a single preferred restoration flow scenario (PRFS) that incorporated seasonally and yearly variable flows to maintain healthy, functioning ecosystems. In addition to estimating water table elevation under the PRFS, the models developed in this study can be applied to any number of possible future scenarios, including increased groundwater withdrawals, sea level rise, and changes in rainfall and ET patterns associated with climate change.

5. Conclusions

[41] Detailed hydrological multivariate time series, obtained in and around the Loxahatchee River watershed in south Florida, were studied and modeled using dynamic factor analysis (DFA). The analysis was successfully applied to understand the hydrological processes in this area, which has been affected by reduced hydroperiod and increased saltwater intrusion. The technique proved to be a powerful tool for the study of interactions among 29 long-term, nonstationary hydrological time series (12 water table elevation [WTE] series and 17 candidate explanatory variables). Upstream and tidal surface water elevations (SWE), regional groundwater circulation (WTE_R), and cumulative net local recharge (R_{net}) were found to be the most important factors responsible for groundwater variation in the floodplain wetlands of the Loxahatchee River, and the analysis quantified the spatial distribution of the importance of each explanatory variable to WTE in the 12 monitoring wells.

[42] Upstream SWE at Lainhart Dam is the primary managed hydrological input in the Loxahatchee River and was important for describing variability in wells T1-W1 and T3-W1 but had limited impact on WTE on downstream transects. Although SWE at Lainhart Dam has been shown to largely dictate downstream surface water salinity [*SFWMD*, 2006], its role in explaining WTE variation is limited to the upstream, riverine river reaches. Tidal SWE at RK 14.6, which is susceptible to climate change-induced sea level rise, was important for explaining observed WTE variability for downstream, lower elevation wells.

[43] WTE_R was significant for nine of the 12 wells, corroborating the noted dependence of floodplain WTE on regional groundwater [*SFWMD*, 2002]. The best DFM used two R_{net} series, with wells T1-W1 and T3-W1 gaining the most benefit from the R_{net} series calculated using rain from the nearby (to these wells) S-46 structure. The importance of the R_{net} series from the JDWX gauging station were split fairly equally over the remainder of the wells. Using the average of the two series (a common technique in small watersheds) yielded inferior results. This highlights the importance of using the best available local rainfall data for hydrological modeling, whether it be empirical or mechanistic, and stresses the need to move to more advanced rainfall measurement techniques, including Next Generation Radar (NexRad).

[44] The DFM resulting from the DFA (Model II) had good results (overall $C_{eff} = 0.91$, $0.78 \leq C_{eff} \leq 1.0$, visual inspection) and is useful for filling in data gaps during the study period and identifying the relative importance and relationships between hydrological variables of interest. The reduced model with no common trends (Model III) did a fair to excellent job (overall $C_{eff} = 0.81$, $0.59 \leq C_{eff} \leq 0.94$) and is likely adequate for describing variations in WTE in the Loxahatchee River floodplain. This empirical model may be deemed useful for assessment of the effects of Loxahatchee River restoration and management scenarios on WTE dynamics.

[45] The study also provides a quantitative validation of our qualitative expectations that tidal effects propagate some distance inland (along a river or estuary), river effects propagate some distance inland from the banks (here to swamps/floodplains), hydraulic structure effects propagate some distance from the structure, and net recharge effects are highly localized. Results of the analysis presented here have practical implications, in addition to guiding climate change mitigation planning and ecohydrologic analysis of salinity in coastal river wetlands. For example, mechanistic modeling efforts that consider spatial variability of land covers and soils would likely benefit from knowledge of where tidal effects end and the degree to which local rainfall variability is an important determining factor. Mechanistic

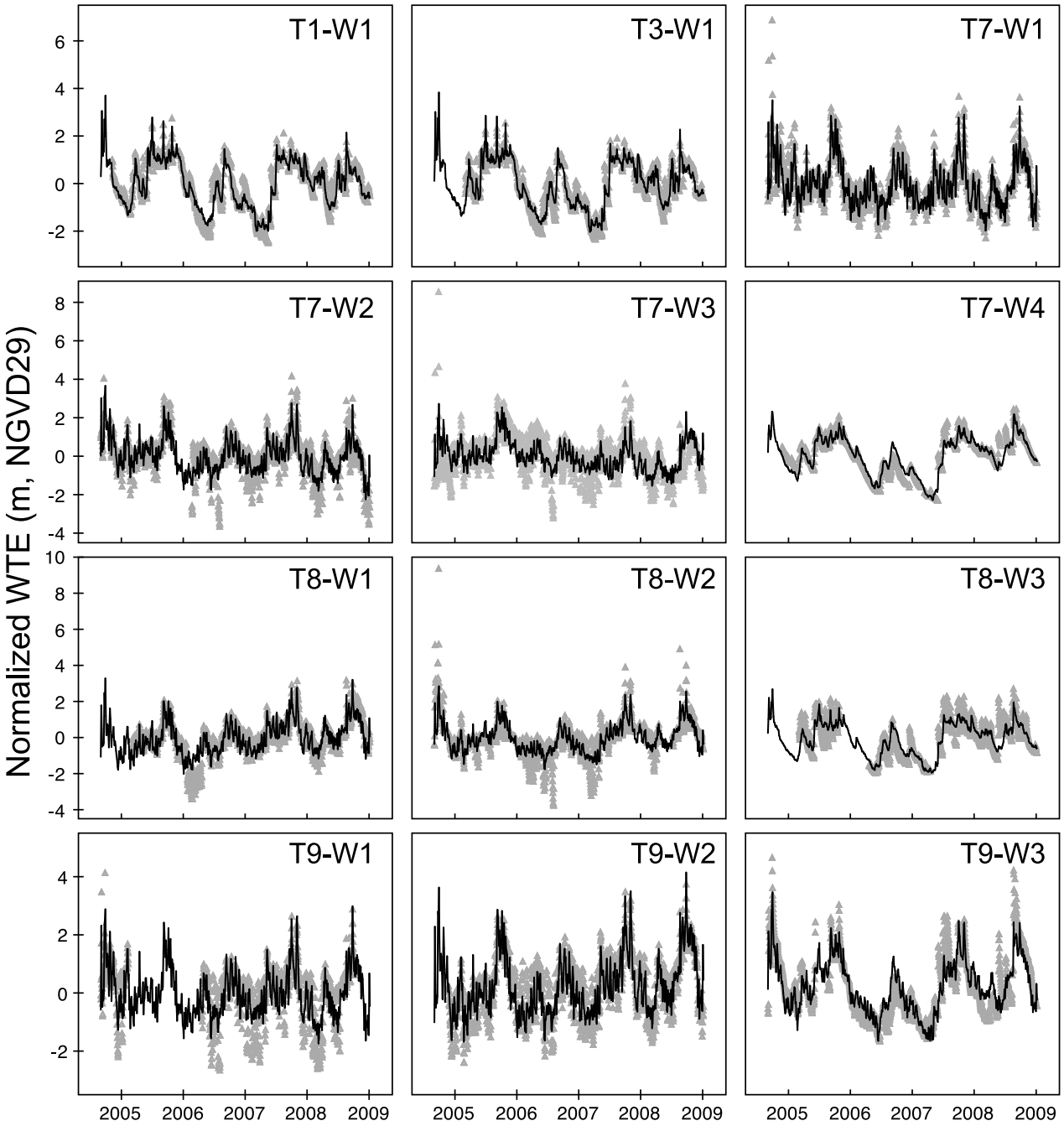

Figure 9. Observed (gray symbols) and modeled (black lines) normalized WTE for the 12 wells obtained from Model III using five explanatory variables and no trends.

frameworks using conditional modeling approaches can also benefit from knowing which explanatory variables are most important as a function of the relative location of a study area to the ocean, a tidal river, or hydraulic structures.

[46] **Acknowledgments.** This work was funded in part by the South Florida Water Management District (SFWMD). The authors would like to thank Fawen Zheng, Gordon Hu, Yongshan Wan, and Marion Hedgepeth (SFWMD) for their support and assistance interpreting groundwater data. Special thanks to Rob Rossmanith (Florida Park Service) for downloading data and maintaining field equipment under brutal field conditions. D. Kaplan acknowledges additional financial support from the University of Florida Graduate Fellowship Program.

References

Akaike, H. (1974), A new look at the Statistical Model Identification, *IEEE Trans. Automat. Control*, *19*, 716–723.

Alexander, T. R., and A. G. Crook (1975), Recent and long-term vegetation changes and patterns in South Florida, Appendix G, Part I, *South Florida Ecological Study*, Univ. of Miami, Coral Gables, Fla.

Benke, A. C., I. Chaubey, G. M. Ward, and E. L. Dunn (2000), Flood pulse dynamics of an unregulated river floddplain in the southeastern U.S. coastal plain, *Ecology*, *81*(10), 2730–2741, doi:10.1890/0012-9658(2000)081[2730:FPDOAU]2.0.CO;2.

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

Dent, R. C. (1997), Rainfall observations in the Loxahatchee River Watershed, report, Loxahatchee River Dist., Jupiter, Fla.

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.

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

Glamore, W. C., and B. Indraratna (2009), Tidal-forcing groundwater dynamics in a restored coastal wetland: implications of saline intrusion, *Aust. J. Earth Sci.*, *56*(1), 31–40.

Hancock, P. J., R. J. Hunt, and A. J. Boulton (2009), Preface: hydrogeoecology, the interdisciplinary study of groundwater dependent ecosystems, *Hydrogeol. J.*, *17*(1), 1–4, doi:10.1007/s10040-008-0409-8.

Harvey, A. C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge Univ. Press, New York.

Harvey, J. W., and P. V. McCormick (2009), Groundwater's significance to changing hydrology, water chemistry, and biological communities of a floodplain ecosystem, Everglades, South Florida, USA, *Hydrogeol. J.*, *17*(1), 185–201, doi:10.1007/s10040-008-0379-x.

Hatton, T., and R. Evans (1998), Dependence of ecosystems on groundwater and its significance to Australia, *Land Water Resour. Res. Develop. Corp. Occ. Paper*, LWRRDC, Canberra, Australia.

Jassby, A. D., W. J. Kimmerer, S. G. Monismith, C. Armor, J. E. Cloern, T. M. Powell, J. R. Schubel, and T. J. Vendlinski (1995), Isohaline position as a habitat indicator for estuarine populations, *Ecol. Appl.*, *5*(1), 272–289.

Johnson, J. W. (2008), *United States Water Law: An Introduction*, Taylor and Francis, New York.

Kaplan, D., R. Muñoz-Carpena, Y. Wan, M. Hedgepeth, F. Zheng, R. Roberts, and R. Rossmanith (2010), Linking river, floodplain, and vadose zone hydrology to improve restoration of a coastal river impacted by saltwater intrusion, *J. Environ. Qual.* (in press).

Kovács, J., L. Márkus, and G. Halupka (2004), Dynamic factor analysis for quantifying aquifer vulnerability, *Acta Geol. Hung.*, *47*(1), 1–17, doi:10.1556/AGeol.47.2004.1.1.

Liu, W., M. Hsu, A.Y. Kuo, and M. Li (2001), Influence of bathymetric changes on hydrodynamics and salt intrusion in estuarine system, *J. Am. Water Resour. Assoc.*, *37*(5), 1405–1419.

Lütkepohl, H. (1991), *Introduction to Multiple Time Series Analysis*, Springer, Berlin.

McPherson, B. F. (1981), The cypress forest community in the tidal Loxahatchee River Estuary: distribution, tree stress, and salinity, report, USGS, Ft. Lauderdale, Fla.

Melloul, A., and L. Goldenberg (1997), Monitoring of seawater intrusion in coastal aquifers: Basics and local concerns, *J. Environ. Manage.*, *51*(1), 73–86.

Mortl, A. (2006), Monitoring soil moisture and soil water salinity in the Loxahatchee floodplain, M.S. thesis, Univ. of Florida, Gainesville.

Muñoz-Carpena, R., A. Ritter, and Y. C. Li (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., D. Kaplan, and F. J. Gonzalez (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., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models, Part 1-A: Discussion of principles, *J. Hydrol.*, *10*, 282–290.

NPS (2004), The National Wild and Scenic Rivers Program, available at www.rivers.gov/wsr-loxahatchee.html (verified 20 August 2009), Nat. Park Serv., Burbank, Wash.

Nyman, J. A., M. K. La Peyre, A. Caldwell, S. Piazza, C. Thom, and C. Winslow (2009), Defining restoration targets for water depth and salinity in wind-dominated *Spartina patens* (Ait.) Muhl. coastal marshes, *J. Hydrol.*, *376*(3–4), 327–336, doi:10.1016/j.jhydrol.2009.06.001.

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

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

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

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

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

Roberts, R. E., R. O. Woodbury, and J. Popenoe (2006), Vascular plants of Jonathan Dickinson State Park, *Florida Scientist*, *69*(4), 288–327.

Roberts, R. E., M. Y. Hedgepeth, and T. R. Alexander (2008), Vegetational responses to saltwater intrusion along the Northwest Fork of the Loxahatchee River within Jonathan Dickinson State Park, *Florida Scientist*, *71*(4), 383–397.

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, Fla.

SFWMD (2006), Restoration plan for the Northwest Fork of the Loxahatchee River. Coastal ecosystems division, South Florida Water Management District, West Palm Beach, Fla.

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, Fla.

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

Swarzenski, P. W., W. H. Orem, B. F. McPherson, M. Baskaran, and Y. Wan (2006), Biogeochemical transport in the Loxahatchee River estuary, Florida: The role of submarine groundwater discharge, *Mar. Chem.*, *101*(3–4), 248–265, doi:10.1016/j.marchem.2006.03.007.

TCRPC (1999), Strategic regional policy plan for the Treasure Coast Region, Rule 29K-5.002, Florida Administrative Code, Treasure Coast Regional Planning Council, Stuart, Fla.

Tulp, I., L. J. Bolle, and A. D. Rijnsdorp (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.

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

Wang, F. C. (1988), Dynamics of saltwater intrusion in coastal channels, *J. Geophys. Res.*, *93*(C6), 6937–6946.

Ward, T. H., and R. E. Roberts (1996), Vegetation analysis of the Loxahatchee River corridor, report, Florida Park Service (5th District), Hobe Sound, Fla.

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

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

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

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

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

D. Kaplan and R. Muñoz-Carpena, Agricultural and Biological Engineering Department, University of Florida, 287 Frazier Rogers Hall, PO Box 110570 Gainesville, FL 32611-0570, USA. (carpena@ufl.edu)

A. Ritter, Departamento de Ingeniería, Produccíon y Economía Agraria, Universidad de La Laguna, Ctra. Geneto, 2, 38200 La Laguna, Spain.