Supplementary Materials:

Data-driven assessment of climate change and vegetative cover dynamics in traditional oases

Elisa Baioni[1,*]

Giulia Fiantanese[1]

Giovanni Michele Porta[1]

1. Politecnico di Milano, Department of Civil and Environmental Engineering, 20133 Milan (IT);

* Corresponding author; e-mail: giovanni.porta@polimi.it.

S1 Estimate of the NDVI threshold for M'Chouneche oasis

A common approach for identifying oasis and desert areas is the application of Otsu's algorithm to Landsat images (Xie et al., 2010). However, as shown in Figure S1(d)-(f) the application of Otsu's algorithm separately to the NDVI Landsat and Sentinel images leads to different results. When applied to Landsat data Otsu's method identifies a lower τ^* value, leading to an overestimation of the vegetative surface with respect to results obtained with Landsat data. Our approach assumes Sentinel-2 as a reference and through the proposed procedure yields consistent results across the two datasets. Figure S1(f) displays the mask obtained using the MATLAB function *otsuthresh*, which computes a global threshold from histogram counts based on Otsu's method.

This comparative approach yields a strong correspondence between the vegetative surfaces estimated from both Landsat 8 with τ^* value and Sentinel-2. The similarity quantified according to Equation (4) exhibits a mean value of 0.98 and a variance of 3.13×10^{-5}. As illustrated in Figure S2, τ^* demonstrates seasonal fluctuations attributable to climatic and vegetative variations. To mitigate this seasonality dependence, the annual mean optimal threshold ${\tau^*}_m$ is calculated. This metric demonstrates a more stable trend over the study period with a maximum variation of approximately 5%. The minimum and maximum values of ${\tau^*}_m$ across 2016-2023 were then employed to account for the inherent uncertainty associated with the oasis's vegetative cover estimation.

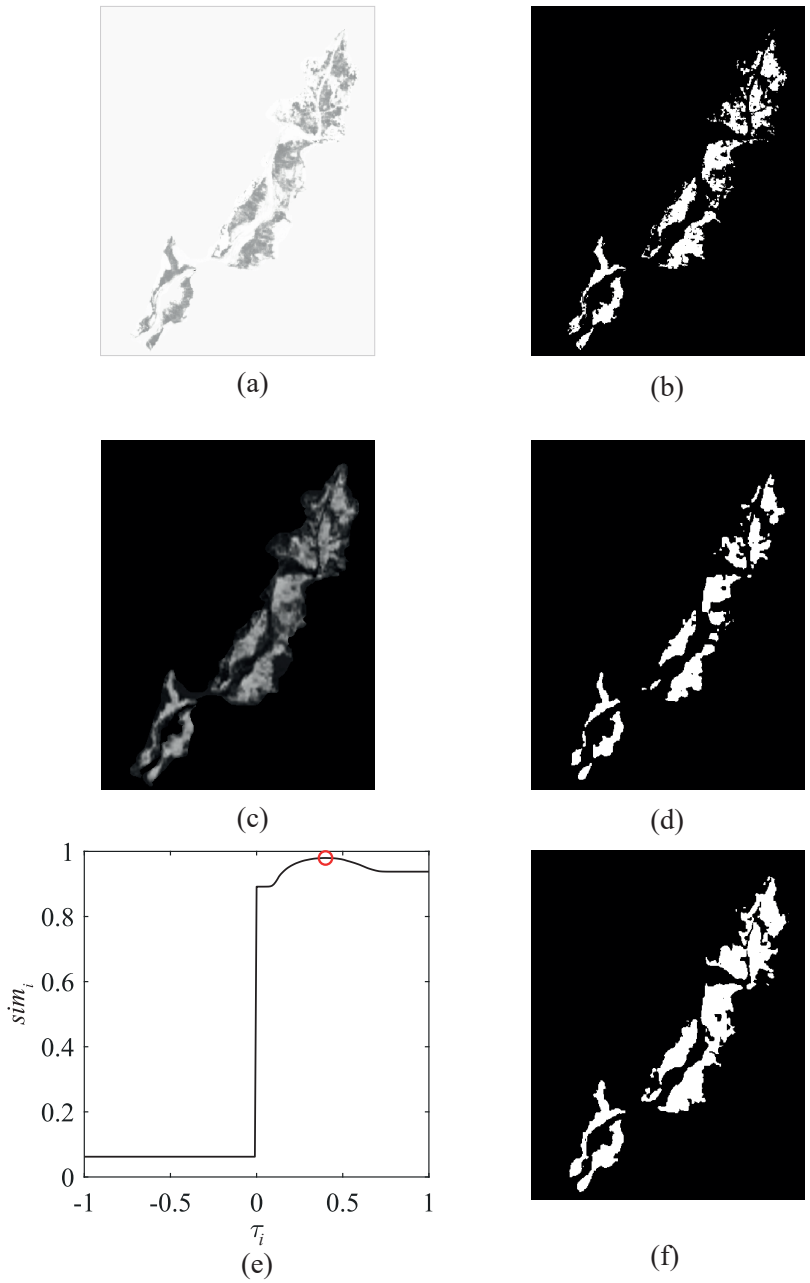

Figure S1: Output of the procedure for determining τ^*: (a) Sentinel-2 NDVI map, (b) Mask of the vegetative surface M_S obtained from Sentinel-2, (c) Landsat NDVI map ,(d) Mask of the vegetative surface M_S obtained from Landsat for $\tau_i = \tau^*$, $(\tau^* = 0.4)$, (e) Similarity for different τ_i values, the red dot corresponds to τ^*, (f) Mask of the vegetative surface M_S obtained from Landsat with Otsu's method $(\tau^* = 0.24)$

The set $[\tau^*_{min}, \tau^*_{max}]$ is used to quantify the maximum and minimum vegetative surface A_v, respectively, from Landsat 8 for the period 2013-2015 and from Landsat 5 for an earlier time frame (1985-2011). To bridge

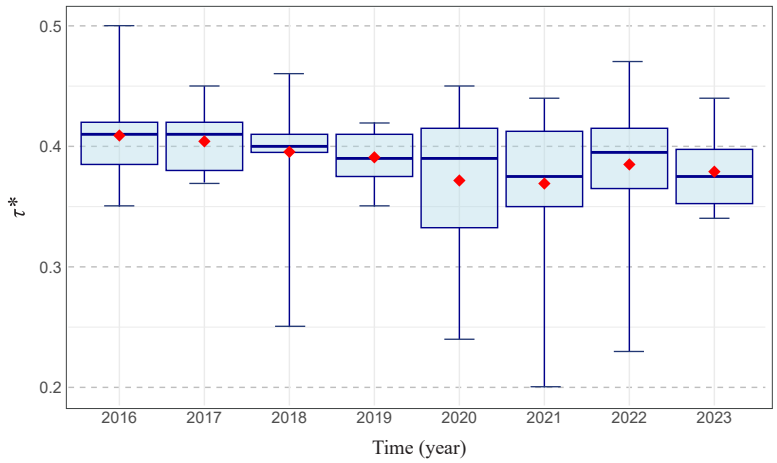

Figure S2: Variation of the monthly optimal threshold τ^* from 2016 to 2023. The box represents the first and third percentile, the blue thick line is the median, the whiskers correspond to ${\tau^*}_{min}$ and ${\tau^*}_{max}$. The red markers represent the monthly mean value of τ^*

the time gap between the two employed satellite imagery datasets, the vegetative area for 2014 was estimated through linear interpolation between the preceding and subsequent years. Prior to applying ${\tau^*}_{min}$,${\tau^*}_{max}$ to estimate vegetative cover for years preceding the calibration period (2016-2023), the similarity of the NDVI distribution between these earlier years (pre-2016) and the established calibration period (2016-2023) is evaluated. Here we rely on the cosine similarity metric (Leydesdorff, 2005)

$$cosine(\overrightarrow{H_i}, \overrightarrow{H_r}) = \frac{\overrightarrow{H_i} \cdot \overrightarrow{H_r}}{\left\|\overrightarrow{H_r}\right\| \left\|\overrightarrow{H_i}\right\|} \tag{S1}$$

where $\overrightarrow{H_i}$ and $\overrightarrow{H_r}$ are the vectors corresponding to the histogram of the NDVI values of the assessed period i and those of the corresponding reference time r, respectively. The analysis is performed monthly, with NDVI data from 2016 used as a reference point. The cosine similarity index measures the similarity between two data sets, ranging from -1 to 1. A value closer to 1 indicates stronger similarity. An average cosine similarity of 0.95 is observed between the NDVI distributions from 2013 to 2015, while the average *cosine* value between Landsat 5 and Landsat 8 NDVI from 1975 to 2011 is 0.81. This lower similarity is likely due to differences in the vegetative surface and data acquisition characteristics of the two Landsat sensors. This level of similarity is deemed sufficient to apply the procedure to Landsat 5 data as well.

S1.1 Estimate of the NDVI threshold for Baniane and Ghoufi oases

The optimal NDVI threshold for Baniane and Ghoufi oases is determined by applying the calibration procedure described in Section 2.2 to Sentinel-2 and Landsat 8 satellite imagery from 2016 to 2023. Similarly to Section S1, the analysis is performed with monthly frequency. Figure S3 (a) shows that for Baniane τ^* fluctuates between 0.32 and 0.40, with an average annual values spanning from 0.347 to 0.362. These values are lower than those observed for M'Chouneche and demonstrate less variability. This discrepancy is likely due to geographical and physical characteristics specific to each oasis system.

Ghoufi exhibits a wider range of variation in the monthly optimal NDVI threshold. It fluctuates between 0.31 and 0.50, with higher values typically observed during the summer months. The minimum and maximum average annual τ^* values for Ghoufi are 0.38 and 0.425, respectively, showing values closer to those found in M'Chouneche. Compared to M'Chouneche, both Baniane and Ghoufi oases display a slightly lower accuracy in calibrating the optimal NDVI threshold. The similarity quantified according to Equation (4) results in an average of 0.95 and a variance of 9.05×10^{-5} for Baniane and an average value of 0.94 and a variance of 7.22×10^{-4} for Ghoufi. This result could be explained by the relatively low spatial resolution of the satellite imagery compared to the smaller dimensions of these oases in relation to M'Chouneche.

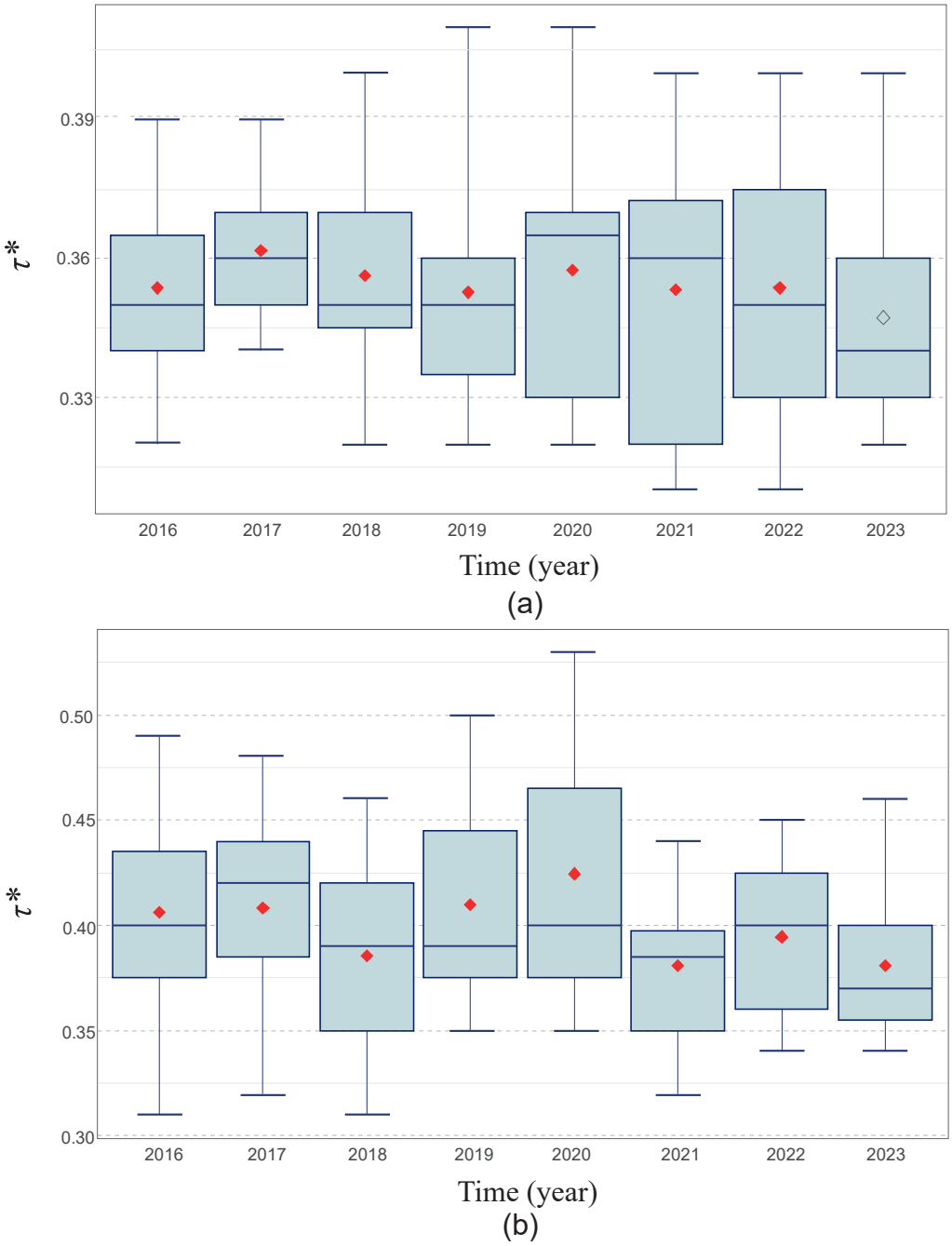

Figure S3: Variation of the monthly optimal threshold τ^* from 2016 to 2023 for Baniane (a) and Ghoufi (b) oases. The box represents the first and third percentile, the blue thick line is the median, the whiskers correspond to ${\tau^*}_{min}$ and ${\tau^*}_{max}$. The red markers indicate the monthly mean value of τ^*

S2 Augmented Dickey-Fuller test

Augmented Dickey-Fuller (ADF) test is an extension of Dickey-Fuller test by including extra lagged terms of the dependent variables in order to eliminate the problem of autocorrelation. The ADF is a unit root test used to determine whether a time series is stationary and operates under the assumption of an autoregressive time series model of order p as (Banerjee et al., 1993)

$$y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t \tag{S2}$$

Subtracting y_{t-1} from both sides Equation S2 writes

$$\Delta y_t = \mu + \delta y_{t-1} \sum_{i=1}^{p} \beta_i y_{t-i} + \varepsilon_t \tag{S3}$$

The test statistic is quantified as

$$t_{\hat{\beta}_i} = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \tag{S4}$$

The test operates under the null hypothesis of the existence of a unit root in the time series, i.e. the time series is non-stationary. The alternative hypothesis is that the time series has no unit root implying stationarity. If the test statistic is lower than the critical value or if the p-value is less than a pre-specified significance level (e.g., 0.05), the null hypothesis is rejected and the time series is considered stationary.

S2.1 ADF test for monthly data

The results of the ADF test performed on the monthly frequency data (cumulative precipitation P_{cm}, mean temperature T_m and evapotranspiration ET_{0_m}, and oasis'water-stressed areas R_{ws}) for an autoregressive model of order 4 with a constant and linear trend are listed in Table S1. For all assessed variables both statistic value and the p-value are smaller than the corresponding values at 5% significance level. This indicates that all the time series are stationary.

Table S1: Results of ADF test for the monthly input data

Variable	t-statistic	Critical value (0.05)	p-value
P_{cm}	-4.8	$-3.424\,48$	<0.01
T_m	-10.23	$-3.424\,48$	<0.01
ET_{0m}	-10.26	$-3.424\,48$	<0.01
R_{ws}	-5.6954	$-3.424\,48$	<0.01

S2.2 ADF test for annual data

The ADF test is performed on the yearly frequency data (cumulative precipitation P_{ca}, annual mean of the temperature T_a and evapotranspiration ET_{0a}, and the oasis's vegetative area A_v) by an autoregressive model of order 3 with a constant and linear trend as listed in Table S2. All the endogenous variables exhibit a t-statistic and p-value greater than the counterparts at the 5% significance level implying the non-stationarity of the time series.

Table S2: Results of ADF test for the annual input data

Variable	t-statistic	Critical value (0.05)	p-value
P_{ca}	-3.2747	$-3.540\,12$	$0.090\,68$
ΔP_{ca}	-5.31	$-3.540\,12$	<0.01
T_a	-2.008	$-3.540\,12$	0.5702
ΔT_a	-4.899	$-3.540\,12$	<0.01
ET_{0a}	-2.0605	$-3.540\,12$	0.5497
ΔET_{0a}	-4.764	$-3.540\,12$	<0.01
A_v	-2.8439	$-3.540\,12$	0.2435
ΔA_v	-3.867	$-3.540\,12$	0.0263

S3 Johansen test

The Johansen test is employed to determine the existence of cointegrations between the endogenous variables. The test checks for a situation of no cointegration occurring when the matrix Π in Equation (9) is equal to 0. The number of cointegrations r corresponds to the rank of the matrix Π which is determined by its eigenvalue decomposition. Johansen test sequentially evaluates whether this rank is zero (indicating no cointegration) up

to $K-1$, where K denotes the number of endogenous variables. The null hypothesis of $r=0$ indicates the absence of cointegration. Conversely, a non-zero rank suggests the presence of a cointegrating relationship between two or more time series. For each $r=i$, the t-statistic is computed and compared against the critical value at a specified significance level. If the t-statistic exceeds the critical value, the null hypothesis is rejected, and the hypothesis $r=i+1$ is subsequently tested.

S3.1 Johansen test for annual data

To understand the long-run relationships between the variables in the annual data discussed in Section 2.3.2, the Johansen test is employed to assess for cointegration. The test is performed in R considering 4 endogenous variables and 6 lags, chosen based on AIC criteria. The results reported in Table S3 reveal that the cointegration matrix has a rank of 2 at the 5% significance level, indicating the existence of 2 cointegrating relationships between the endogenous variables.

Table S3: Johansen test results for the annual input data

	Test Statistics	Critical Values		
$\mathcal{H}_0$		90 %	95 %	99 %
$r=0$	117.78	49.65	53.12	60.16
$r=1$	49.19	32.00	34.91	41.07
$r=2$	18.05	17.85	19.96	24.60
$r=3$	3.17	7.52	9.24	12.97

S4 Jarque-Bera (JB) test

The Jarque-Bera test helps determine if the model residuals follow a normal distribution. The univariate versions of the Jarque-Bera test are applied to the residuals of each equation. A multivariate version of this test can be computed by using the residuals that are standardized by a Choleski decomposition of the variance-covariance matrix for the centered residuals (Pfaff and Taunus, 2007). The test statistics for the

multivariate case are defined as (Pfaff and Taunus, 2007)

$$JB_{mv} = {s_3}^2 + {s_4}^2 \tag{S5}$$

with the multivariate skewness ${s_3}^2$ and kurtosis ${s_4}^2$ quantified as

$${s^2}_3 = \frac{N {b_1}' b_1}{6} \tag{S6}$$

$${s^2}_4 = \frac{N(b_2 - 3)'(b_2 - 3)}{24} \tag{S7}$$

where N is the number of observations, b_1 and b_2 are the third and fourth non-central moment vectors of the standardized residuals, respectively. The test statistic JB_{mv} is distributed as $\chi^2(2K)$ while ${s^2}_3$ and ${s^2}_4$ are distributed as $\chi^2(K)$, with K corresponding to the number of variables. The test is performed under the null hypothesis that the residuals are normally distributed. The null hypothesis is rejected if the p-value associated with the test is small (typically below a significance level like 0.05).

S5 Test for residual correlation

The residual autocorrelation in VAR and VEC models can be assessed by the residual Portmanteau and Breusch-Godfrey-LM tests.

The null hypothesis in Portmanteau test is that all residual autocovariances are zero, i.e. the residuals are not autocorrelated. The alternative hypothesis is that at least one autocovariance, and one autocorrelation, is nonzero. The test statistic up to the order h relies upon the residual autocovariances $\hat{C}_j$ as (Pfaff and Taunus, 2007)

$$Q_h = N \sum_{j=1}^{h} \mathrm{tr}(\hat{C}_j' \hat{C}_0^{-1} \hat{C}_j \hat{C}_0^{-1}) \tag{S8}$$

The Breusch-Godfrey LM test is suggested for low-order autocorrelation (Lütkepohl, 2013). The t-statistic

relies upon the following p-order regressions (Pfaff and Taunus, 2007)

$$\hat{\mathbf{u}}_t = A_1\mathbf{y}_{t-1} + \ldots + A_p\mathbf{y}_{t-p} + CD_t + B_1\hat{\mathbf{u}}_{t-1} + \ldots + B_h\hat{\mathbf{u}}_{t-h} + \varepsilon_t \tag{S9}$$

As for the Portmanteau test, the null hypothesis posits that the residuals are not autocorrelated, i.e. $B_1 = \cdots = B_h = 0$. The test statistic is defined as

$$LM_h = N(K - \mathrm{tr}(\tilde{\Sigma}_R^{-1}\tilde{\Sigma}_e)) \tag{S10}$$

where K is the number of variables, $\tilde{\Sigma}_R$ and $\tilde{\Sigma}_e$ assign the residual covariance matrix of the restricted and unrestricted model, respectively. In both tests, if the t-statistic or the p-value derived from the model exceeds the critical counterparts, the residuals are deemed not autocorrelated

S6 Heteroscedasticity (ARCH) test

The heteroscedasticity of the residuals is performed by the ARCH test (Lütkepohl, 2005). The test relies on the following regression (Pfaff and Taunus, 2007)

$$\mathrm{vech}(\hat{\mathbf{u}}_t\hat{\mathbf{u}}_t^\top) = \boldsymbol{\beta}_0 + B_1\,\mathrm{vech}(\hat{\mathbf{u}}_{t-1}\hat{\mathbf{u}}_{t-1}^\top) + \ldots + B_q\,\mathrm{vech}(\hat{\mathbf{u}}_{t-q}\hat{\mathbf{u}}_{t-q}^\top) + \mathbf{v}_t \tag{S11}$$

where $\mathbf{v}_t$ assigns a spherical error process and vech is the column-stacking operator for symmetric matrices that stacks the columns from the main diagonal on downward. The test statistic is defined as (Pfaff and Taunus, 2007)

$$\mathrm{VARCH}_{LM}(q) = \frac{1}{2}NK(K+1)R_m^2 \tag{S12}$$

with $R_m^2 = 1 - \frac{2}{K(K+1)}\,\mathrm{tr}(\hat{\Omega}\hat{\Omega}_0^{-1})$ and $\hat{\Omega}$ assigns the covariance matrix of the above defined regression model. The null hypothesis is that there is no ARCH effect (heteroscedasticity) in the model, i.e. $B_1 = B_2 = \ldots = B_q = 0$. As for all the mentioned checking tests, the null hypothesis is accepted if the calculated

t-statistic or p-value from the model is greater than the corresponding critical counterparts.

S7 Forecast Error Variance Decomposition (FEVD)

The Forecast Error Variance Decomposition (FEVD) is a tool to assess the dynamic interaction between the endogenous variables. These approaches are based upon the Wold moving average decomposition for stable VAR(p)-processes defined as (Pfaff and Taunus, 2007)

$$y_t = \Phi_0 u_t + \Phi_1 u_{t-1} + \Phi_2 u_{t-2} + \dots \tag{S13}$$

with $\Phi_0 = I_K$ and Φ_s can be computed recursively according to

$$\Phi_s = \sum_{j=1}^{s} \Phi_{s-j} A_j \text{ for } s = 1, 2, \dots \tag{S14}$$

where $A_j = 0$ for $j > p$. Each $(i, j)th$ coefficient of the matrix Φ_s describes the expected response of the variable $y_{i,t+s}$ to a unit change in variable $y_{j,t}$. These effects can be accumulated through time, $s = 1, 2, .., h$, and hence one would obtain the simulated impact of a unit change in variable j to the variable i at time s (Pfaff and Taunus, 2007). When contemporaneous correlations between the components of the error process u_t exist, i.e. the off-diagonal elements of Σ_u are non-zero, the orthogonal impulse response can be used. It is derived from the Choleski decomposition of the error variance-covariance matrix $\Sigma_u = PP^\top$, with P being a lower triangular. The moving average representation in Equation (S13) results

$$y_t = \Psi_0 \varepsilon_t + \Psi_1 \varepsilon_{t-1} + \dots \tag{S15}$$

with $\varepsilon_t = P^{-1} u_t$ and $\Psi_i = \Phi_i P$ for $i = 0, 1, 2, \dots$ and $\Psi_0 = P$. The forecast error variance decomposition is based on the orthogonal impulse response coefficient matrices Ψ_n. The FEVD allows to analyze the impact of the variable j on the h-step forecast error variance of the variable k. The result is expressed in percentage by dividing the element-wise squared orthogonal impulse responses by the variance of the forecast error variance, $\sigma_k^2(h)$ defined as

$$\sigma_k^2(h) = \sum_{n=0}^{h-1} (\psi_{k1,n}^2 + \ldots + \psi_{kK,n}^2) \tag{S16}$$

S8 Autocorrelation

In time series analysis, autocorrelation, also referred to as lagged correlation, quantifies the similarity between a series and a lagged version of itself across different time intervals. Autocorrelation can be estimated at different lags to evaluate the relationship between a time series and its past values. Figure S4 displays the autocorrelation function (ACF) plots for monthly data of cumulative precipitation P_{cm}, mean temperature T_m and evapotranspiration ET_{0m}, and water-stressed area R_{ws}.

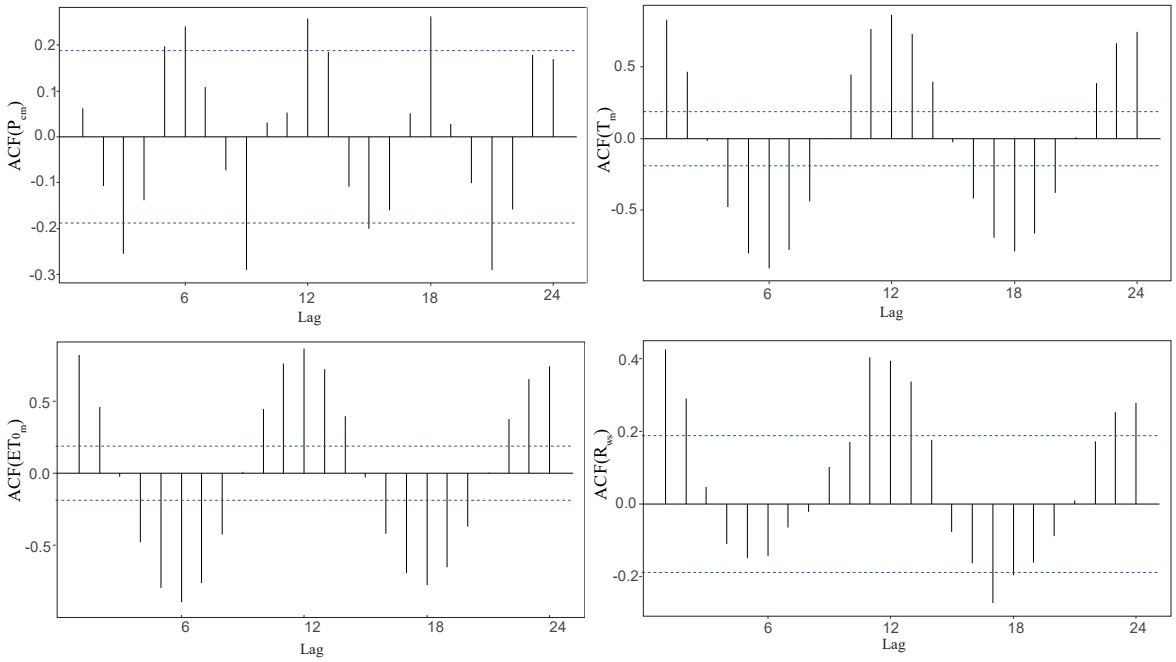

Figure S4: Autocorrelation of the endogenous variables: cumulative precipitation P_{cm}, mean temperature T_m, mean evapotranspiration ET_{0m}, and water-stressed areas R_{ws}

S9 Estimated VAR model

The VAR model coefficients estimated by the OLS are reported in Table S4. The term $.li$ refers to the ith lag of the variable. The coefficients $sd_1....sd_{11}$ are associated to the seasonal dummies. The standard error for each estimated coefficient is presented in brackets below the respective coefficient.

Table S4: Estimated model coefficients for the endogenous variables P_{cm}, T_m, ET_{0m}, R_{ws}

	P_{cm}	T_m	ET_{0m}	A_s
$P_{cm}.l1$	-0.011	0.013	0.007	0.001^{*}
	(0.139)	(0.023)	(0.008)	(0.001)
$P_{cm}.l1$	-0.011	0.013	0.007	0.001^{*}
	(0.139)	(0.023)	(0.008)	(0.001)
$ET_{0m}.l1$	1.094	0.693	0.299	0.028
	(3.769)	(0.610)	(0.217)	(0.021)
$R_{ws}.l1$	-21.583	0.420	0.910	0.115
	(20.362)	(3.298)	(1.170)	(0.111)
$P_{cm}.l2$	-0.171	0.023	0.012	0.002^{***}
	(0.141)	(0.023)	(0.008)	(0.001)
$T_m.l2$	0.396	0.081	-0.027	0.005
	(1.204)	(0.195)	(0.069)	(0.007)
$ET_{0m}.l2$	-4.895	0.406	0.279	-0.009
	(3.828)	(0.620)	(0.220)	(0.021)
$R_{ws}.l2$	-25.486	-1.224	1.091	0.244^{**}
	(19.071)	(3.089)	(1.096)	(0.104)
$P_{cm}.l3$	0.113	0.016	0.004	-0.001^{*}
	(0.142)	(0.023)	(0.008)	(0.001)

Continued on next page

Table S4 – continued from previous page

	P_{cm}	T_m	ET_{0m}	A_s
$T_m.l3$	-1.735	0.311	0.129^{*}	0.006
	(1.229)	(0.199)	(0.071)	(0.007)
$ET_{0m}.l3$	6.215	-0.507	-0.247	-0.029
	(3.848)	(0.623)	(0.221)	(0.021)
$R_{ws}.l3$	15.382	1.946	0.280	0.199^{*}
	(19.770)	(3.203)	(1.136)	(0.108)
$const$	20.916	10.658^{**}	3.110^{**}	0.061
	(26.840)	(4.348)	(1.542)	(0.146)
sd_1	-0.921	1.304	0.902^{*}	0.014
	(8.979)	(1.454)	(0.516)	(0.049)
sd_2	-1.034	5.318^{**}	2.223^{***}	0.076
	(14.148)	(2.292)	(0.813)	(0.077)
sd_3	10.230	8.673^{***}	3.559^{***}	0.023
	(17.165)	(2.780)	(0.986)	(0.094)
sd_4	6.694	12.872^{***}	4.969^{***}	0.005
	(17.267)	(2.797)	(0.992)	(0.094)
sd_5	-2.036	15.956^{***}	6.009^{***}	-0.036
	(16.623)	(2.693)	(0.955)	(0.091)
sd_6	-11.219	19.609^{***}	7.431^{***}	0.026
	(16.299)	(2.640)	(0.937)	(0.089)
sd_7	-15.901	20.090^{***}	7.150^{***}	0.035
	(17.661)	(2.861)	(1.015)	(0.096)
sd_8	-5.341	17.837^{***}	5.584^{***}	0.009
	(18.357)	(2.974)	(1.055)	(0.100)

Continued on next page

Table S4 – continued from previous page

	P_{cm}	T_m	ET_{0m}	A_s
sd_9	2.925	13.454***	3.816***	0.001
	(17.164)	(2.780)	(0.986)	(0.094)
sd_10	−0.314	7.400***	2.046**	−0.049
	(13.924)	(2.255)	(0.800)	(0.076)
sd_11	6.857	1.863	0.413	−0.074
	(8.405)	(1.362)	(0.483)	(0.046)

***Significant at the 1% level

**Significant at the 5% level

*Significant at the 10% level

S10 Estimated VEC model

The estimated coefficient of the VEC model (Model$_c$) discussed in Section 4.2.2 are listed in Table S5. The term $.li$ represents the ith lag of the variable, while ECT is the error correct term (α coefficient in Equation (9)) describing long-term relationships among endogenous variables. The estimated coefficients are referred to the VEC model expressed in first difference.

Table S5: Estimated VEC model coefficients for the endogenous variables P_{ca}, T_a, ET_{0a}, A_v

	P_{ca}	T_a	ET_{0a}	A_v
ECT_1	0.0784	0.0012	−0.0003	0.0009
	(0.185)	(0.003)	(0.001)	(0.0003)**
ECT_2	−58.1089	−0.5697	0.2061	−0.4758
	(98.608)	(1.394)	(0.502)	(0.142)**
$P_{ca}.l1$	−1.5653	0.0029	0.0036	−0.0018

Continued on next page

Table S5 – continued from previous page

	P_{cm}	T_m	ET_{0m}	A_s
	(0.527)*	(0.007)	(0.003)	(0.001)*
$T_a.l1$	23.7370	−0.7312	−0.1527	0.3784
	(76.899)	(1.086)	(0.391)	(0.110)**
$ET_{0a}.l1$	−553.6759	1.0752	1.8410	−1.4534
	(378.92)	(5.3552)	(1.9304)	(0.5451)*
$A_v.l1$	270.3014	1.0995	−0.5556	1.1378
	(244)	(3.44)	(1.243)	(0.3510)**
$P_{ca}.l2$	−1.9033	0.0054	0.0061	−0.0026
	(0.651)*	(0.009)	(0.003).	(0.0009)*
$T_a.l2$	25.8440	−0.4357	0.0682	0.2333
	(72.134)	(1.019)	(0.367)	(0.104)*
$ET_{0a}.l2$	−669.6430	2.5105	2.5105	−1.3709
	(313.521).	(4.431)	(1.5972)	(0.4510)*
$A_v.l2$	163.8843	1.0338	−0.9899	0.7150
	(223.847)	(3.163)	(1.140)	(0.322).
$P_{ca}.l3$	−1.6939	0.0073	0.0066	−0.0026
	(0.674)*	(0.01)	(0.003).	(0.001)*
$T_a.l3$	−21.0114	−0.2758	0.2067	0.1417
	(75.426)	(1.066)	(0.384)	(0.108)
$ET_{0a}.l3$	−471.7939	2.2769	1.5670	−0.9465
	(288.568)	(4.078)	(1.470)	(0.415)*
$A_v.l3$	-71.4781	1.8139	-0.1713	1.0648
	(207.865)	(2.938)	(1.059)	(0.299)**
$P_{ca}.l4$	-1.2573	0.0065	0.0045	-0.0024

Continued on next page

Table S5 – continued from previous page

	P_{cm}	T_m	ET_{0m}	A_s
	(0.560)*	(0.008)	(0.003)	(0.001)*
$T_a.l4$	−58.9040	0.5013	0.4146	−0.0068
	(61.350)	(0.867)	(0.312)	(0.090)
$ET_{0a}.l4$	−284.5224	0.5914	0.5624	−0.5394
	(221.565)	(3.131)	(1.129)	(0.319)
$A_v.l4$	−4.3986	2.2512	0.4110	0.3384
	(174.727)	(2.469)	(0.890)	(0.251)
$P_{ca}.l5$	−0.4313	0.0030	0.0016	−0.0017
	(0.375)	(0.005)	(0.002)	(0.001)**
$T_a.l5$	−27.6675	0.6170	0.2943	0.0460
	(39.849)	(0.563)	(0.203)	(0.0573)
$ET_{0a}.l5$	−86.9430	−0.2346	0.3019	−0.5421
	(170.505)	(2.410)	(0.869)	(0.245).
$A_v.l5$	−154.7819	−0.2060	−0.2805	0.6550
	(196.028)	(2.7704)	(0.998)	(0.282)*

***Significant at the 1% level

**Significant at the 5% level

*Significant at the 10% level

The normalized cointegration vectors related to the two cointegration relationships between the endogenous variables resulting from the Johansen test are reported in Table S6.

S10.1 Checking test

The results of the diagnosis tests for the VEC model Model$_c$, assessing the model's performance in terms of stability, serial correlation, heteroskedasticity, and normality, are reported in Table S7. Given the limited

Table S6: Estimated normalized cointegration vectors

	r_1	r_2
P_{ca}	1.00	0
T_a	0	1.00
ET_{0a}	-8683.60	-20.19
A_v	-7001.56	-10.90
const	59756.32	113.21

available data, the residual autocorrelation is evaluated based on the Breusch-Godfrey LM test. The test results are expressed in p-value.

Table S7: Test results of the VEC model

Stability (roots)	**Serial correlation**	**Heteroskedasticity (ARCH)**	**Normality (JB)**
< 1	0.14	1	0.60

References

Banerjee, A., Dolado, J. J., Galbraith, J. W., and Hendry, D. (1993). *Co-integration, error correction, and the econometric analysis of non-stationary data*. Oxford university press.

Leydesdorff, L. (2005). Similarity measures, author cocitation analysis, and information theory. *Journal of the American society for Information Science and Technology*, 56(7):769–772.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.

Lütkepohl, H. (2013). Vector autoregressive models. In *Handbook of research methods and applications in empirical macroeconomics*, pages 139–164. Edward Elgar Publishing.

Pfaff, B. and Taunus, K. (2007). Using the vars package. *Kronberg: im Taunus*, page 2007.

Xie, Y., Li, L., Wang, H., and Zhao, X. (2010). The application of threshold methods for image segmentation in oasis vegetation extraction. In *2010 18th International Conference on Geoinformatics*, pages 1–4. IEEE.