ENTROPY-BASED VARIABILITY AND SUPPORT VECTOR REGRESSION-BASED FORECAST OF DROUGHT INDEX

Özlem Baydaroğlu , Kasım Koçak , Levent Şaylan

Istanbul Technical University, Faculty of Aeronautics and Astronautics, Department of Meteorological Engineering

Contacts: **baydaroglu@itu.edu.tr**, kkocak@itu.edu.tr, saylan@itu.edu.tr

Introduction

In this study, US Palmer Drought Severity Index (PDSI), Palmer Hydrological Drought Index (PHDI), Palmer Z-Index (ZNDX), Modified Palmer Drought Severity Index (PMDI), Standard Precipitation Index (SPI) monthly data between the years June 1929 and September 2014 which they belongs to NOAA are used for 3-year forecast.

PDSI is an meteorological drought index (Palmer, 1965) which it is a function of precipitation, temperature and available water content (AWC) of the soil (Palmer, 1965, Alley, 1984). PMDI is different from the PDSI during transition periods between dry and wet spells. PMDI bases on the weighting factor, which is a weighting average of the wet and dry index terms by probability (Heddingshause and Sabol, 1991). PHDI is a hydrological drought index and indicates the severity of a wet or dry spell like PDSI. It relies on precipitation (as inflow), outflow, and storage. But it does not take the long-term trend in consideration (Karl and Knight, 1985). ZNDX measures moisture anomaly. It indicates deviation from normal monthly PDSI. SPI, which is an meteorological drought index, is a normalized index. It bases on the precipitation data (observed and long term measured data (precipitation climatology)). Mainly relies on the standard deviation, precipitation at the time considered, and precipitation climatology. Precipitation data are required to be converted into a normal distribution.

Study

This study aims to determine the entropy-based variability of drought index and forecast these index by using Support Vector Regression (SVR (Cortes&Vapnik(1995)). SVR transforms input space which is formed from the original observations into high dimensional feature space by way of a kernel function and performs a linear regression in this space. Variability can be expressed with Disorder Index (DI). The higher the disorder index, the higher the variability.

Methods

- *Disorder Index*

$$H = -\sum_{i=1}^{n} p_i \log_2 p_i \quad \Rightarrow \quad DI = \log_2 n - H$$

- *Chaotic Approach (CA)* ⇒ Phase Space Reconstruction (Embedding Theorem-Takens,1981)

» Embedding dimension via False Nearest Neighbour(FNN) (Kennel et al., 1992)

» Time delay via Mutual Information Function(MIF)

- *Support Vector Regression (SVR)*

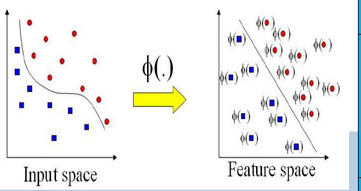

Fig. 1. Observation vector transforms to vectors in multidimensional space with Φ.

	Target	Input Data: 1st Dimension	2nd Dimension	3rd Dimension	...	mth Dimension
Training Data	X_{t+1}	X_t	$X_{t-\tau}$	$X_{t-2\tau}$	⋯	$X_{t-(m-1)\tau}$
	X_{t+2}	X_{t+1}	.	.	⋯	⋯
	.	.	.	.	.	.	.
	.	.	.	.	.	.	.
	.	.	.	.	.	.	.
Test Data	.	.	.	.	.	.	.

Table 1. The input matrix for SVR based on the phase space reconstruction.

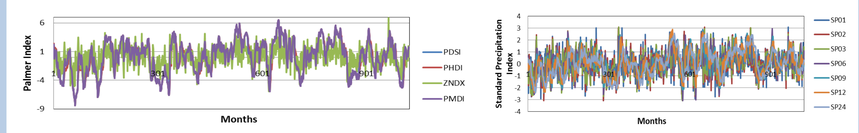

Fig. 2. NOAA monthly Palmer Index(PDSI, PHDI, ZNDX, PMDI) and Standard Precipitation Index (SPI) between the years June 1929 and September 2014.

Results

Variability

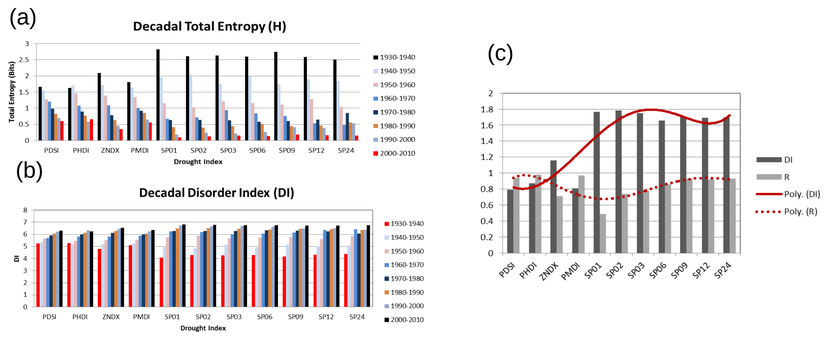

Fig. 3. (a) Decadal total entropy (b) Decadal DI (c) DI versus determination coefficients (R)

Phase Space Reconstructions

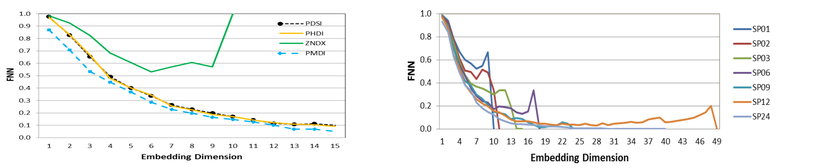

Fig. 4. False nearest neighbours for the Palmer Index and the Standard Precipitation Index

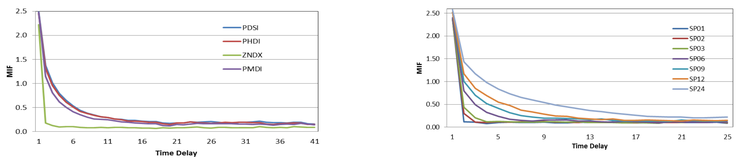

Fig. 5. Mutual information function for Palmer Index and the Standard Precipitation Index

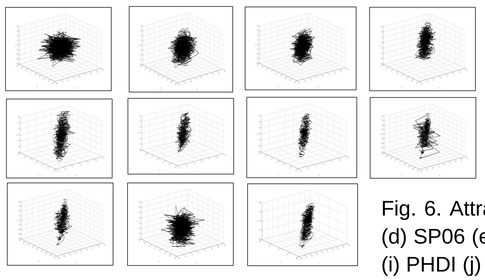

Fig. 6. Attractors (a) SP01 (b) SP02 (c) SP03 (d) SP06 (e) SP09 (f) SP12 (g) SP24 (h) PDSI (i) PHDI (j) ZNDX (k) PMDI

Table 2. Information about the prediction made by phase space reconstruction

	PDSI	PHDI	ZNDX	PMDI	SP01	SP02	SP03	SP06	SP09	SP12	SP24
Total Data Length	1024	1024	1024	1024	1024	1024	1024	1024	1024	1024	1024
Phase Space Length	853	872	1016	1016	997	1016	972	961	988	720	870
Training Data Length	817	836	980	980	961	980	936	925	952	684	834
Test Data Length	36	36	36	36	36	36	36	36	36	36	36
Embedding Dimension	10	9	3	3	10	3	14	10	10	20	8
Time Delay	19	19	4	4	3	4	4	7	4	16	22

Prediction

Table 3. SVR parameters of prediction of all index

SVR Parameters	g	C	ϵ
PDSI	0.114626	430.538971	0.000106
PHDI	0.297302	34.896248	0.012473
ZNDX	29.344130	0.162105	0.000232
PMDI	64	0.034078	0.000164
SP01	3.363586	3.084422	0.282241
SP02	128	0.125	0.083911
SP03	0.176777	58.688259	0.064704
SP06	4.362031	0.25	0.099787
SP09	0.028656	512	0.049894
SP12	0.420448	0.917004	0.108819
SP24	0.210224	98.701492	0.016176

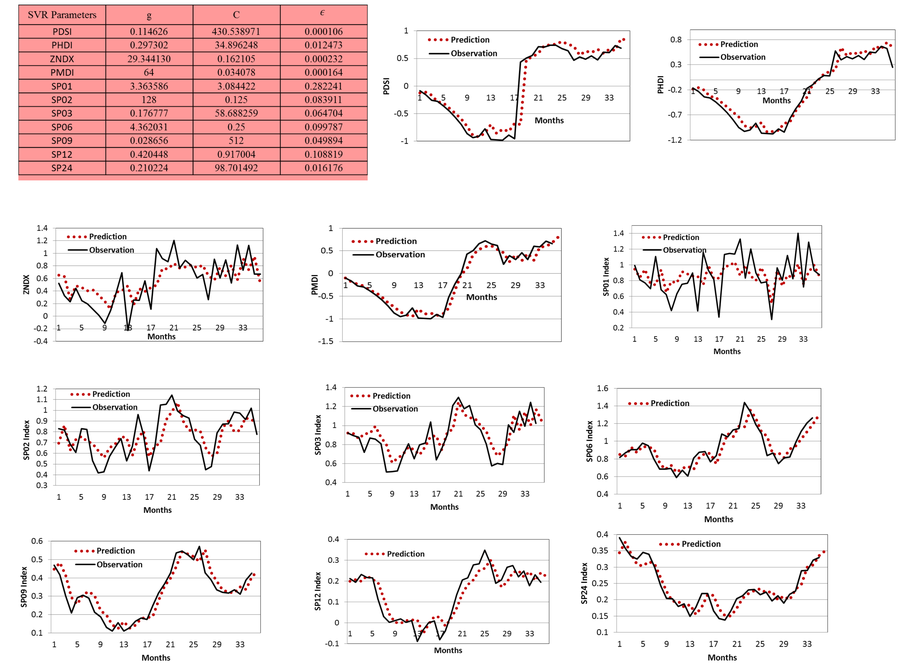

Fig. 7. Prediction vs. Observations for all index

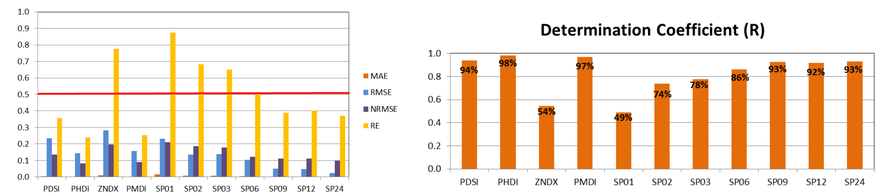

Fig. 8. Errors and determination coefficients of the predictions

References

Takens, F., 1981. Detecting strange attractors in turbulence. In: Rand, D.A., Young, L.S. (Eds.), Lecture Notes in Math. Springer-verlag, pp. 366-381.

Kennel, M.B., Brown, R., Abarbanel, H.D.I., 1992. Determining embedding dimension for phase-space reconstruction using a geometrical construction. Phys. Rev. A 45, 3403-3411.

Cortes, C.; Vapnik, V. (1995). "Support-vector networks". *Machine Learning* **20**(3): 273.

Palmer, W. C, 1965: Meteorological Drought. Res. Paper No.45, 58pp., Dept. of Commerce, Washington, D.C.

Alley, W.M. 1984. The Palmer Drought Severity Index: Limitations and assumptions.Journal of Climate and Applied Meteorology 23:1100–1109.

Heddinghaus, T.R.; and P. Sabol. 1991. A review of the Palmer Drought Severity Index and where do we go from here? In Proc. 7th Conf. on Applied Climatology, pp. 242–246. American Meteorological Society, Boston.

Karl, T.R.; and R.W. Knight. 1985. Atlas of Monthly Palmer Hydrological Drought Indices (1931–1983) for the Contiguous United States. Historical Climatology Series 3–7, National Climatic Data Center, Asheville, North Carolina.

Hegger, R., Kantz, H., Schreiber, T., 1999. Practical implementation of nonlinear time series methods: the TISEAN package. Chaos 9, 413-435.

Shannon, C. E., 1948. A mathematical theory of communication.

Itoh, K., 1995. A method for predicting cahotic time-series with outliers.