GEOPHYSICAL RESEARCH LETTERS, VOL. 28, NO. 17, PAGES 3305-3307, SEPTEMBER 1, 2001

Quantifying the Complexity in Mapping Energy Inputs and Hydrologic State Variables into Land-Surface Fluxes

Gabriel G. Katul [1], Chun-Ta Lai

School of the Environment, Duke University

John D. Albertson

Department of Environmental Sciences, University of Virginia

Brani Vidakovic

School of Industrial and Systems Engineering, Georgia Institute of Technology

Karina V. R. Schäfer

School of the Environment, Duke University

Cheng-I Hsieh

Humanities and General Sciences, National Taiwan University of Science and Technology

Ram Oren

School of the Environment, Duke University

Abstract. This study explores the complexity (or disorder) in mapping energy (R_n) forcing to land surface fluxes of sensible heat (H_s), water vapor (LE), and carbon dioxide (or net ecosystem exchange,NEE) for different soil water states (θ). Specifically, we ask, does the vegetation act to increase or dissipate statistical entropy injected from R_n? We address this question using novel scalar complexity measures applied to a long-term time series record of R_n, θ, H_s, LE, and NEE collected over a uniform pine forest. This analysis is the first to demonstrate that vegetation dissipates scalar flux entropy injected through R_n. We also find that the entropy or disorder in scalar fluxes increases with increasing R_n and that the complexity in mapping R_n to scalar fluxes is reduced with increasing θ.

Introduction

How ecosystems map net radiation (R_n) and precipitation to land surface fluxes of sensible heat (H_s), latent heat (LE), and net ecosystem exchange (NEE) of carbon dioxide continues to be a fundamental and practical research problem in surface hydrology, ecology, and meteorology. One approach to exploring the complexity of this mapping is through the use of scalar complexity measures originally derived in information theory. Information theory, developed to deal with compression and transmission of data, has significantly contributed to fields such as statistical mechanics, computer science (algorithmic complexity), probability theory, and nonlinear dynamics. Central to this theory is the concept of entropy, the fundamental basis of the second law of thermodynamics, now used to measure randomness or disorder of a process, mapping, or transformation [*Kapur and Kesavan*, 1992; *Baddi and Politi*, 1999; *Kantz and Schreiber*, 1999]. Within this context, we explore the following fundamental question: **Does the vegetation act to increase or dissipate entropy injected from radiative and hydrologic forcing?** We investigate this question using entropy measures applied to a long time series of soil moisture (θ), R_n, H_s, LE, and NEE collected in a uniform pine forest.

Method of Analysis

Formally, entropy is a measure on the space of probability distributions. For a measured time series x, it is possible to define an empirical entropy measure of the Shannon type as

$$H(x) = -\sum_i p_x(i)\ \ln(p_x(i)) \tag{1}$$

where, $p_x(i)$ is the magnitude of the empirical probability of x at bin i [*Kapur and Kesavan*, 1992; *Baddi and Politi*, 1999; *Kantz and Schreiber*, 1999]. For context, magnitudes of entropy estimates, $H(x)$ of sample stochastic processes are shown in Table 1. Similarly, the joint Shannon entropy $H(x, y)$ between two sequences (x, y) is defined as

$$H(x,y) = -\sum_j \sum_i p_{x,y}(i,j)\ \ln(p_{x,y}(i,j)) \tag{2}$$

[1] Also at the Center for Nonlinear and Complex Systems, Duke University

Paper number 2000GL012154.
0094-8276/01/2000GL012154$05.00

Table 1. Sample one and two-dimensional Entropy and information measures for several stochastic processes and for measured wind speed U, surface fluxes, and forcing time series. Entropy is estimated from p_x, where p_x is the empirical probability of sequence x computed from a histogram of 50 bins ($= N$). For joint variables, $p_{x,y}$ is the joint probability density function estimated from 50×50 bins, and r^2 is the coefficient of determination between sequences x and y. The maximum theoretical $H(x)$ is $\log N$.

Process	Entropy or Information Content
Fractional Brownian Motion (Hurst Exponent $= \frac{2}{3}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}$)	3.75, 3.71, 3.66, 3.62
Random Normal, Uniform Deviate	3.31, 3.91
$H(x)$, $x = H_s, LE, NEE, \theta, R_n, U$	2.88, 2.29, 2.57, 3.65, 3.26, 3.25
I(x,y), x, y independent normal variables ($r^2 = 0.01$)	0.08
I(x,y), x, y correlated normal variables ($r^2 = 0.98$)	3.26
H(θ, H_s), H(θ, LE), H(θ, NEE)	0.028, 0.065, 0.075
H(R_n, H_s), H(R_n, LE), H(R_n, NEE)	0.73, 0.52, 0.54

where $p_{x,y}(i,j)$ is the joint empirical probability defined at bin (i,j). The single and joint entropy measures can be used to estimate the mutual information content $I(x,y)$, or redundancy between two variables (x,y), which is given by:

$$I(x,y) = H(x) + H(y) - H(x,y) \tag{3}$$

Small values of $I(x,y)$ imply low levels of redundancy or limited mutual information content (see Table 1). Note that $H(x,y) \leq H(x) + H(y)$ so that $I(x,y) \geq 0$. The conditional information redundancy $I(x:y)$, defined as $H(y) - H(y|x)$, measures the information content that x carries about y. Using Bayes' theorem, it can also be shown that $I(x,y) = I(x:y) = I(y:x)$. These measures are applied to the flux time series described next.

Data

Data were collected at a site (35°9°N, 80.0°W, 163 m above mean sea level) in a managed loblolly pine (*Pinus Taeda L.*) forest on $< 2\%$ slope near Durham, North Carolina. Tree height was approximately 15 m in 1999. The mean annual precipitation is 1150 *mm* evenly distributed throughout the year, thus this ecosystem is typically energy limited rather than water limited (except for few drought episodes in the summer of 1998). An eddy covariance system of conventional configuration, positioned above the canopy was used to measure H_s, LE, and NEE over 30 minutes averaging periods [*Katul et al.*, 1999; *Lai et al.*, 2000]. The flux measurements, initiated in mid August 1997, are part of a long-term flux monitoring network (AmeriFlux) in the United States [*Kaiser*, 1998]. The study period considered here is from mid August 1997 to July 2000. After elimination of periods during rain events and eddy-covariance system downtime, 65% of the entire 3 *year* period was sampled. Synchronous with the turbulent flux measurements are R_n time series (measured by a *Q7* Fritschen type net radiometer) and θ time series (measured through the top 30 *cm* soil horizon using a Campbell Scientific *CSI-615* frequency domain reflectometry system comprising of 24 vertical probes randomly located within the forest). A hard clay pan restricts the root zone to the top 30*cm* of the soil profile; hence, the θ measurements reflect root-zone soil moisture content dynamics [*Oren et al.*, 1998]. The data comprised 25334 half-hour periods of simultaneous R_n, θ, H_s, LE, and NEE measurements. The entropy measures were calculated by computing the empirical probability density function (histogram) of each of the five variables over $N = 50$ bins. Our selection of the bin number is in agreement with current practice in related non-parametric statistical procedures [*Tapia and Thomson*, 1978]. With $N = 50$, the maximum theoretical bounds on $H(x)$ are $[0, 3.91]$.

Results and Discussion

This analysis explores whether vegetation reduces the entropy arriving from R_n. The analysis then progresses to explore whether the flux entropy state is dependent on the magnitude of R_n. Particularly, we seek to understand whether the disorder in the flux time series increases with increasing incident energy. Finally, we explore whether the soil moisture state of the ecosystem has a measurable impact on the complexity of mapping R_n to land-surface fluxes.

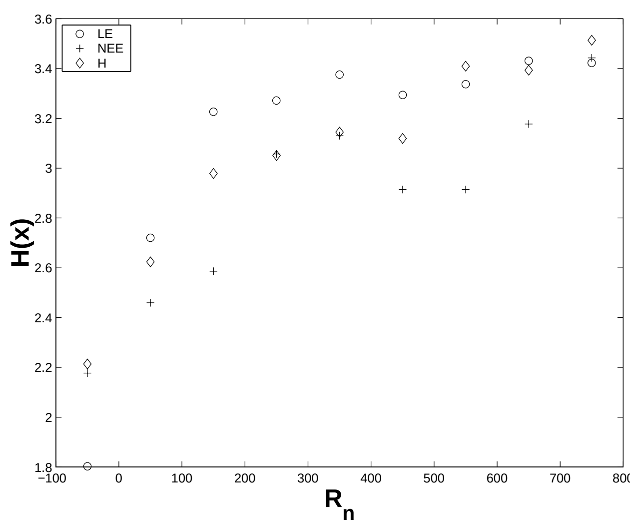

Figure 1. Variation of H(x) with $R_n(Wm^{-2})$ for $x = H_s, LE, NEE$. For reference, the maximum theoretical entropy is $H = 3.93$

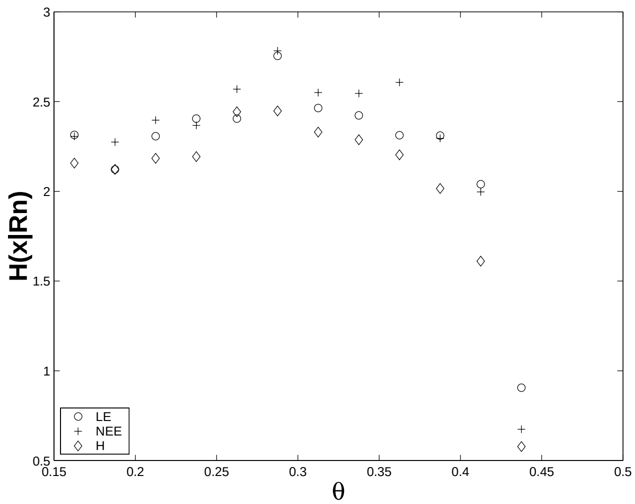

Figure 2. The variation of $H(x|R_n)$ with θ.

Single Entropy Measures

From Table 1, we find that the entropy in the individual H_s, LE, and NEE series is smaller than the entropy in incident R_n; thus the vegetation dissipates entropy introduced via R_n. To further explore the role of incident net radiation on the magnitude of entropy in land surface fluxes, $H(x)$ patterns for $x = H_s$, LE, and NEE are shown in figure 1 for varying R_n. These results indicate that complexity (or disorder) in all three land surface fluxes increase with increasing R_n (increased excitation energy). We emphasize here that for each R_n bin, $H(x)$ is computed from a zero-mean and unit variance x. That is, the effect of small night-time fluxes on entropy is already accounted for in such normalization.

Joint Entropy Measures

To investigate whether the complexity in the relationship between x ($=H_s$, LE , and NEE) and R_n is dependent on θ, we compute the conditional entropy $H(x|R_n)$ based on the state of θ as shown in figure 2. We find that with increasing θ, the information R_n carries about x is increased (or the disorder in such a relationship is reduced). Interestingly, the $H(x|R_n)$ does not exceed $H(R_n)$ for all moisture states, clearly suggesting disorder reduction by vegetation when compared to the incident energy forcing.

Conclusions

This study demonstrated through a case study that vegetation can serve to dissipate statistical entropy from the received energy forcing variable through the processes of water, energy, and carbon exchange with the atmosphere.

Acknowledgments. This project was supported by the National Science Foundation's WEAVE and Hydrologic Science programs (*EAR-9902957* and *EAR-9903471*), *DMS-9626159*, and *NSF-BIR951233*. The long term flux measurements are supported by the National Institute of Global Environmental Change (*NIGEC*) through the NIGEC South-East Regional Center at the University of Alabama, Tuscaloosa, under the cooperative agreement *DE-FC03-90-ER-61010*, and the FACE-FACTS project (*DE-FG05-95ER6283*). The MATLAB files and functions used in the entropy calculations are available upon request from the corresponding author.

References

Baddi, R. and A. Politi, *Complexity: Hierarchical structures and scaling in Physics*, Cambridge Nonlinear Science Series *6,* 318 pp., 1999.

Kaiser, J., Climate change - new network aims to take the world's CO_2 pulse source, *Science*, *281,* 506-507, 1998.

Kantz, H., and T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge Nonlinear Science Series *7,* 303 pp., 1999.

Kapur, J.N., and H. K. Kesavan, *Entropy Optimizations: Principles with Applications*, Academic Press, 408 pp., 1992

Katul, G.G., C.I Hsieh, D. Bowling, K. Clark, N. Shurpali, A. Turnipseed, J. Albertson, K. Tu, D. Hollinger, B. Evans, B. Offerle, D. Anderson, D. Ellsworth, C. Vogel, R. Oren, Spatial variability of turbulent fluxes in the roughness sublayer of an even-aged pine forest, *Boundary Layer Meteorology, 93,* 1-28, 1999.

Lai, C.T., G.G. Katul, D. Ellsworth, R. Oren, Modeling vegetation-atmosphere CO_2 Exchange by a coupled Eulerian-Lagrangian Approach, *Boundary Layer Meteorology, 95,* 91-12, 2000.

Oren, R., B. Ewers, P. Todd, N. Phillips, G.G. Katul, Water balance delineates the soil layer in which moisture affects canopy conductance, *Ecological Applications, 8,* 990-1002, 1998.

Tapia, R. A. and J.R. Thompson, *Nonparametric Probability Density Estimation.* Johns Hopkins University Press, 176 pp., 1978.

G. G. Katul, C. T. Lai, K. Schäfer, and R. Oren, School of the Environment, Box 90328 Duke University, Durham, NC 27708-0328, USA. (e-mail: gaby@duke.edu; cl9@duke.edu; Karina@duke.edu; ramoren@duke.edu)

J. D. Albertson, Department of Environmental Sciences, University of Virginia, Charlottesville, VA 22903, USA (e-mail: jdalbertson@virginia.edu)

B. Vidakovic, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0205, USA (e-mail: brani@isye.gatech.edu)

C. I. Hsieh, Environment and Ecology, Department of Humanities and General sciences, National Taiwan University of Science and Technology, Taipei, 10660, Taiwan, ROC. (e-mail: hsieh@duke.edu)

(Received August, 2000; revised August, 2000;
accepted August, 2000.)