UNIVERSITÄT BAYREUTH

Bayceer
Bayreuth Center of Ecology and Environmental Research

Model–data comparison with permutation entropy: Moving beyond summary statistics

Christina Bogner[1], Britta Aufgebauer[1] and Bernd Huwe[2]

[1]Ecological Modelling, BayCEER, University of Bayreuth DE
[2]Soil Physics Group BayCEER, University of Bayreuth DE

2017-04-27

Background

- Soil water participates in terrestrial water and energy cycles.
- Soil matric potential is an important driving force for soil water flow in unsaturated soil.
- Movement of soil water is often modelled by the Richards equation.
- Model quality assessed by summary statistics (root mean square error, Nash–Sutcliffe coefficient)

Does the model capture the complexity of the underlying soil processes?

Background

- Soil water participates in terrestrial water and energy cycles.
- Soil matric potential is an important driving force for soil water flow in unsaturated soil.
- Movement of soil water is often modelled by the Richards equation.
- Model quality assessed by summary statistics (root mean square error, Nash–Sutcliffe coefficient)

Does the model capture the complexity of the underlying soil processes?

Complexity of a time series

- wide range of complexity: regular (low complexity) to random (high complexity)
- caused in natural time series by nonlinearity of underlying processes and their interactions
- soil matric potential
  - signal propagation from precipitation to throughfall to infiltration
  - influenced by soil hydraulic properties, evapotranspiration and possible measurement errors

Permutation entropy (PE)

Definition

$$H(n) = -\sum p(\pi) \log p(\pi)$$

$$PE = H(n)/\log(n!)$$

- works on ranks (order of values) (Bandt and Pompe, 2002)
- suitable for an arbitrary series of observations

Example

$$x = (3.1; 5.2; 7.9; -3.1; 4.0; 11.9; 0.3) \qquad n = 2$$

$$A = \{(x_t, x_{t+1}) | x_t < x_{t+1}\} \qquad p(A) = 4/6$$

$$B = \{(x_t, x_{t+1}) | x_t > x_{t+1}\} \qquad p(B) = 2/6$$

$$H(2) = -4/6 \log(4/6) - 2/6 \log(2/6) \approx 0.92$$

$$PE = H(2)/\log(2!) \approx 0.92$$

Permutation entropy (PE)

Definition

$$H(n) = -\sum p(\pi) \log p(\pi)$$

$$PE = H(n)/\log(n!)$$

- works on ranks (order of values) (Bandt and Pompe, 2002)
- suitable for an arbitrary series of observations

Example

$$x = (3.1; 5.2; 7.9; -3.1; 4.0; 11.9; 0.3) \qquad n = 2$$

$$A = \{(x_t, x_{t+1}) | x_t < x_{t+1}\} \qquad p(A) = 4/6$$

$$B = \{(x_t, x_{t+1}) | x_t > x_{t+1}\} \qquad p(B) = 2/6$$

$$H(2) = -4/6 \log(4/6) - 2/6 \log(2/6) \approx 0.92$$

$$PE = H(2)/\log(2!) \approx 0.92$$

Weighted permutation entropy (WPE)

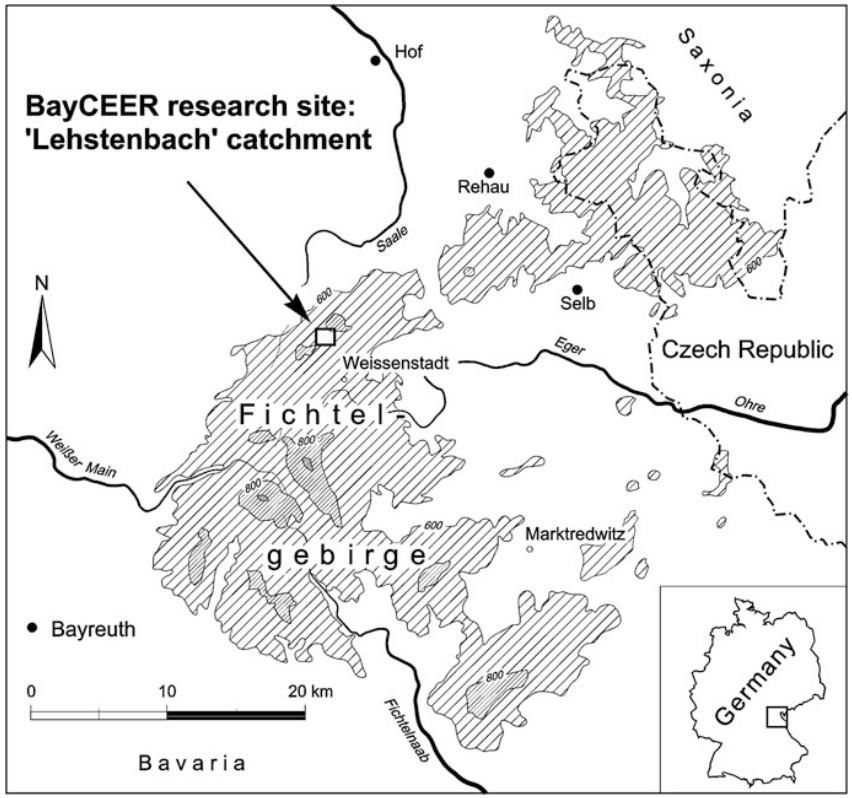
- for tracking abrupt changes
- weights the frequency of permutations by variance of values (Fadlallah et al., 2013)
- better suited for time series with large amplitudes in the signal compared to noise
- small WPE: small number of different patterns (regular/monotonic signal)
- large WPE: large number of different patterns (typical for noise)

Weighted permutation entropy (WPE)

- for tracking abrupt changes
- weights the frequency of permutations by variance of values (Fadlallah et al., 2013)
- better suited for time series with large amplitudes in the signal compared to noise

- small WPE: small number of different patterns (regular/monotonic signal)
- large WPE: large number of different patterns (typical for noise)

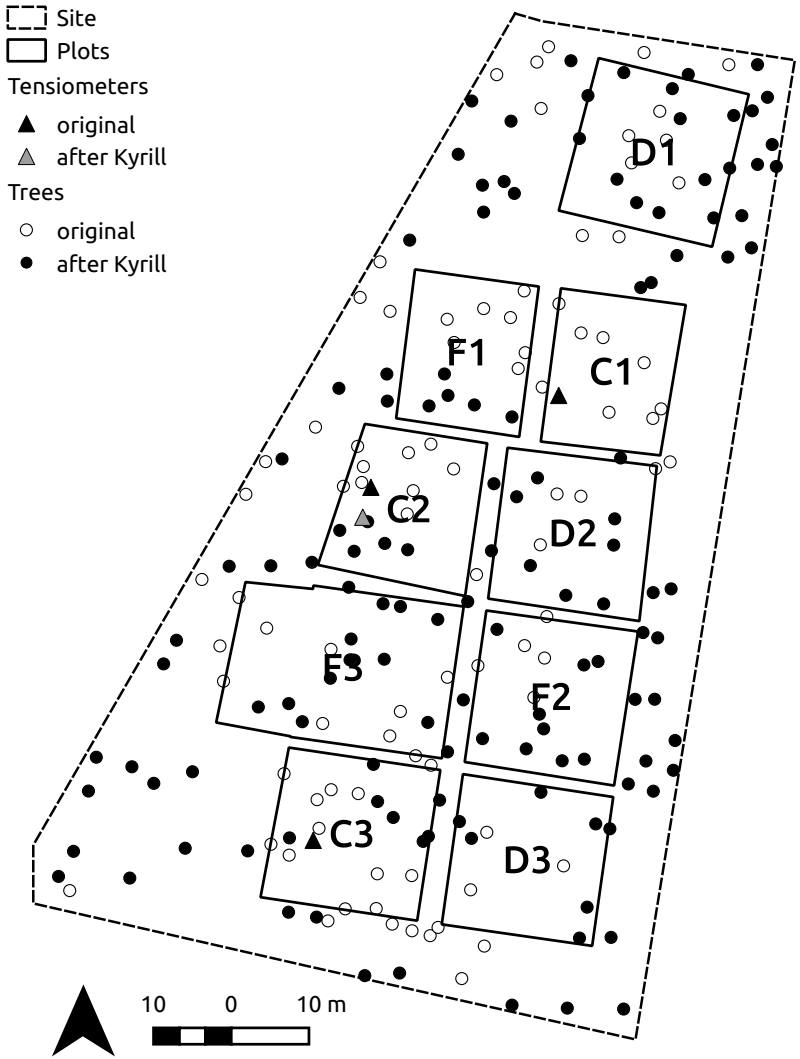
Study area – Lehstenbach catchment



Gerstberger et al., 2004

- 4.5 km^2, 695–877 m a.s.l
- annual precip 1 162 mm, mean temp 5.3 °C
- Norway spruce (*Picea abies*), age 65 years

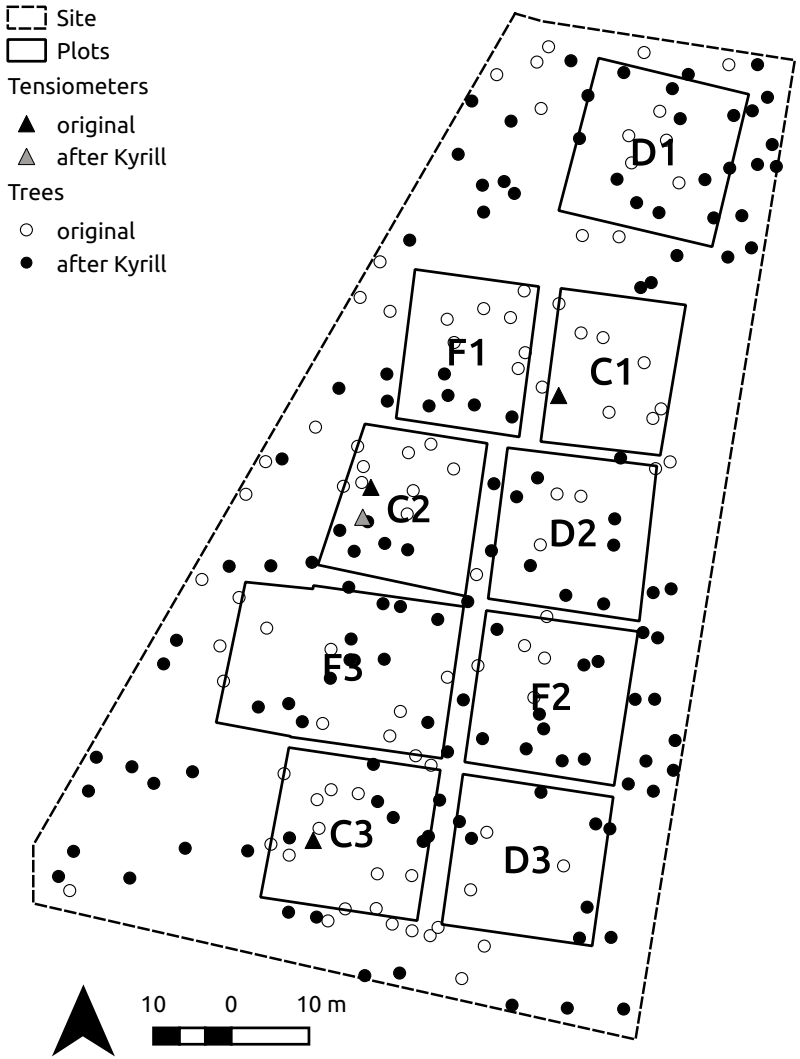
Plots



- Cambisols and Podzols
- sandy to loamy texture
- up to 15-cm thick mor-type organic layer

Bogner et al., 2017

Plots



Bogner et al., 2017

Storm Kyrill

Modelling the soil matric potential

WHNSIM

- Water Heat and Nitrogen Simulation Model (Huwe and Totsche, 1995)
- solves one-dimensional Richards equation
- K_{sat} measured with constant head method
- soil-water retention curve measured in the lab on undisturbed soil cores
- atmospheric conditions measured at nearby sites (Foken et al., 2017)
- model run in forward mode on the daily basis

Calculating the entropy

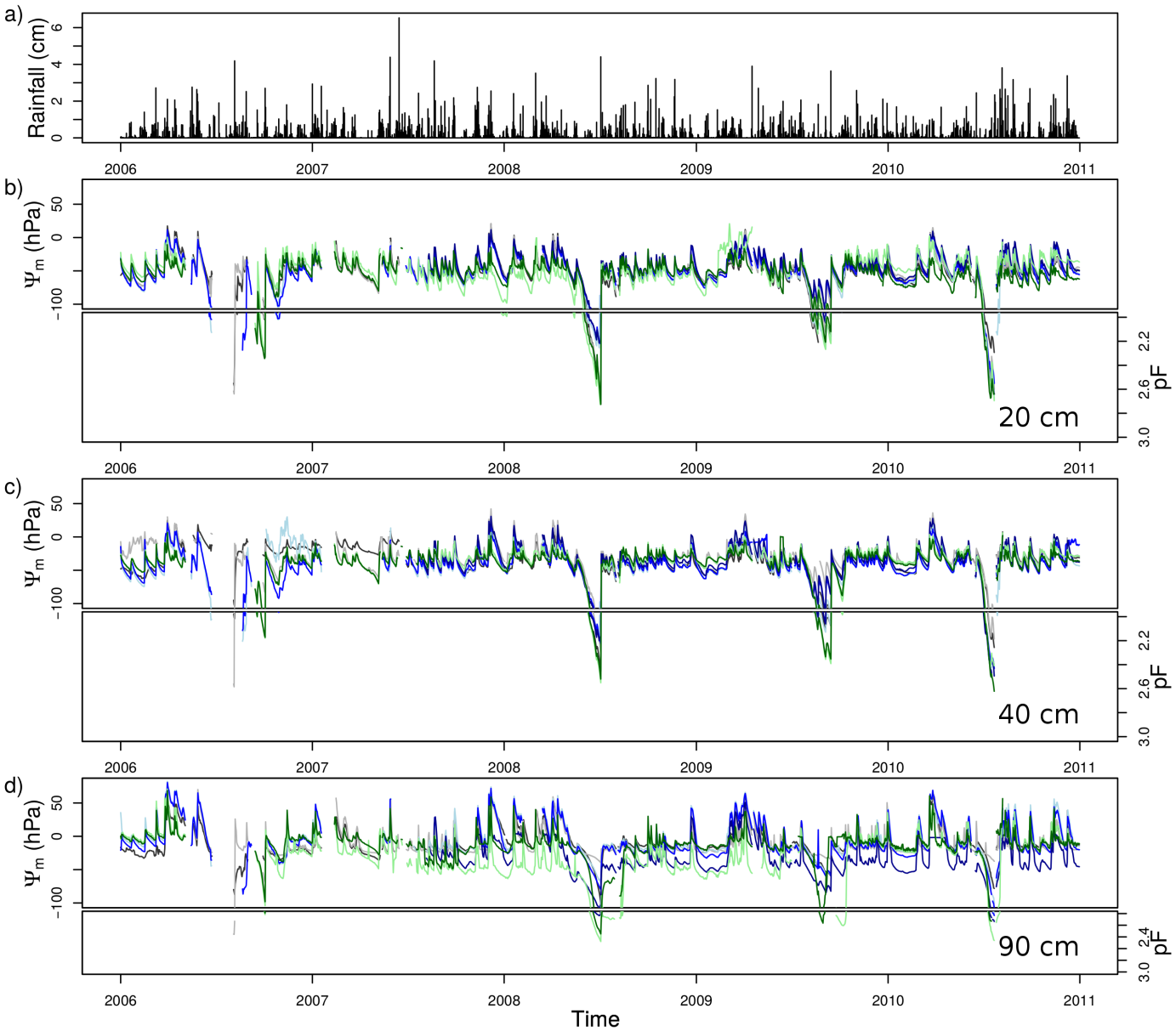
Data pretreatment

- measured matric potential at C1 (the completest series)
- daily median values (from hourly data)
- gap filling with Singular Spectrum Analysis (Golyandina and Osipov, 2007, e.g.)

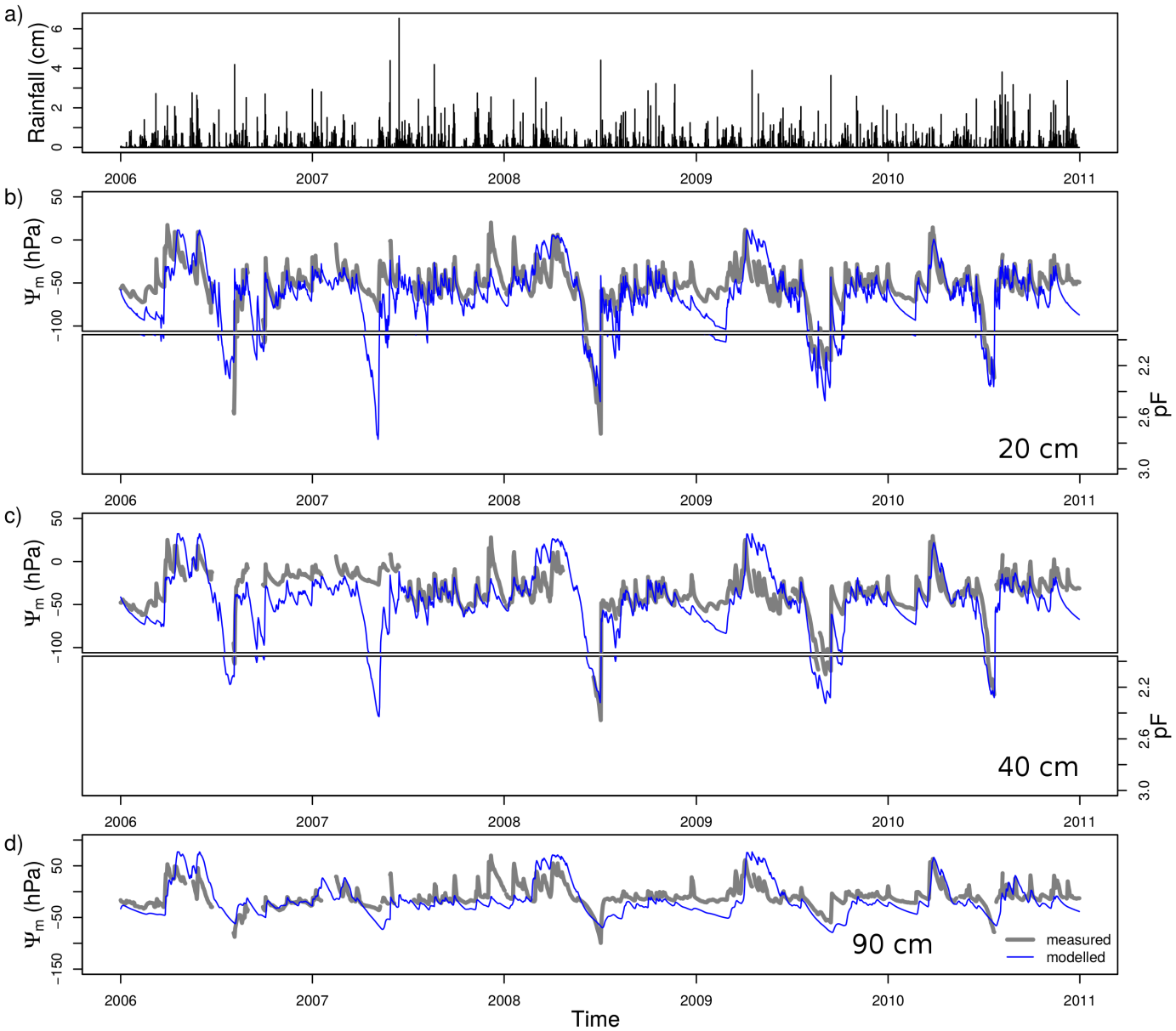
Weighted permutation entropy

- pretreated measured and modelled matric potential
- window length 90 days, slid by 1 day
- order $n = 4$

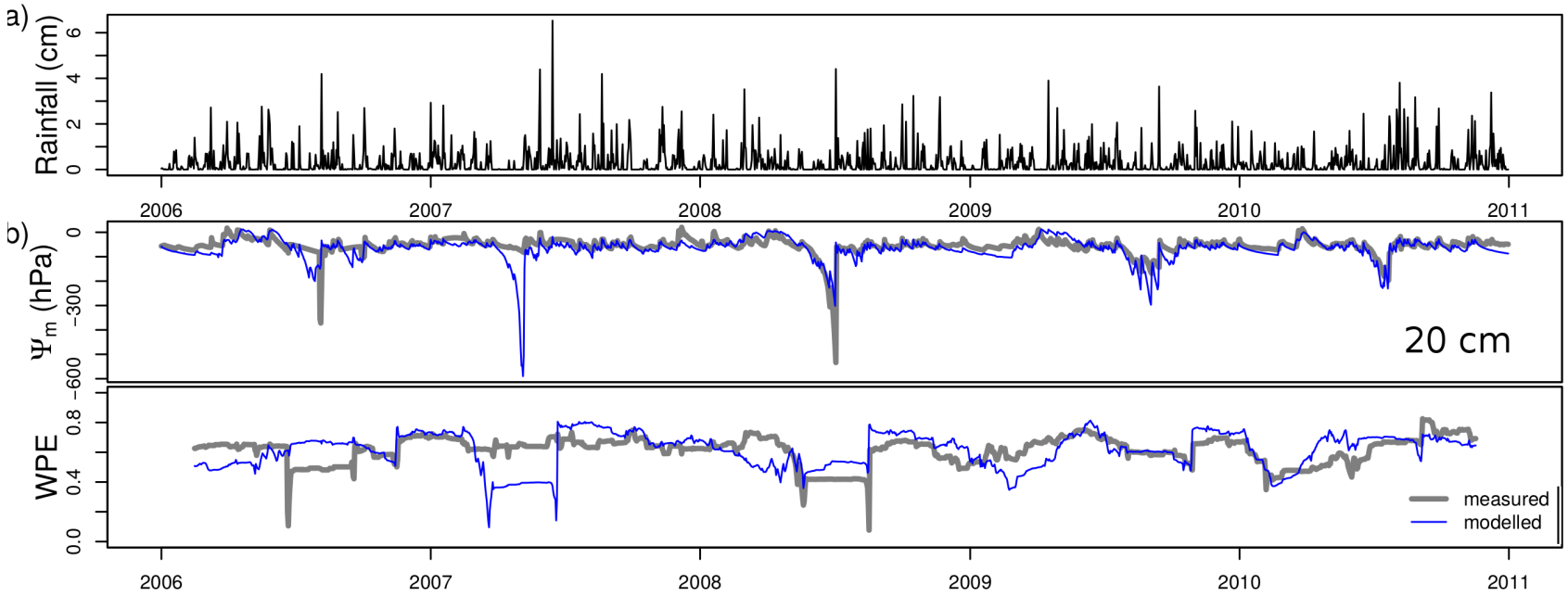
Measured matric potential



Modelled matric potential

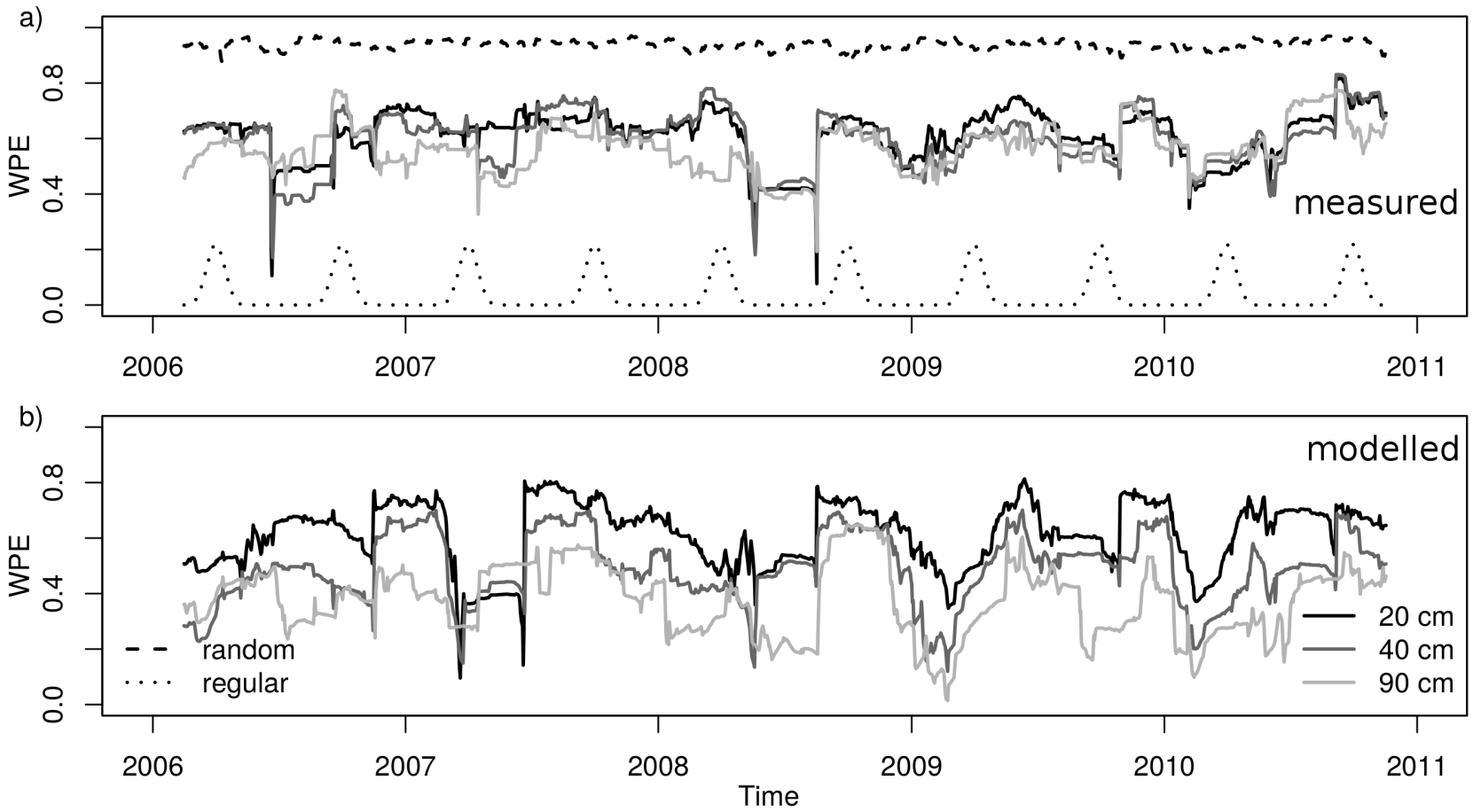


Weighted permutation entropy



Bogner et al., 2017

Entropy in different depths



Bogner et al., 2017

Lessons learned

- WHNSIM captures the complexity of measured matric potential in the upper soil.
- Synchronicity of matric potential in different depths is reproduced.
- The modelled signal in 90 cm is less complex (damped) than the measurements.
- Some process might be missing from the model (in the deeper soil).

Bibliography I

Bandt, C. and B. Pompe (2002). "Permutation entropy: a natural complexity measure for time series". In: *Physical review letters* 88.17, p. 174102.

Bogner, C., B. Aufgebauer, O. Archner, and B. Huwe (2017). "Dynamics of water flow in a forest soil: Visualization and modelling". In: *Energy and Matter Fluxes of a Spruce Forest Ecosystem*. Ed. by T. Foken. Vol. 229. Ecological Studies. Springer, Cham. Chap. 7, pp. 137–156. DOI: 10.1007/978-3-319-49389-3_7.

Fadlallah, B., B. Chen, A. Keil, and J. Príncipe (2013). "Weighted-permutation entropy: A complexity measure for time series incorporating amplitude information". In: *Physical Review E* 87.2, p. 022911.

Bibliography II

Foken, T., P. Gerstberger, K. Köck, L. Siebicke, A. Serafimovich, and J. Lüers (2017). "Description of the Waldstein Measuring Site". In: *Energy and Matter Fluxes of a Spruce Forest Ecosystem*. Ed. by T. Foken. Vol. 229. Ecological Studies. Springer, Cham. Chap. 2, pp. 19–38. DOI: 10.1007/978-3-319-49389-3_2.

Gerstberger, P., T. Foken, and K. Kalbitz (2004). "The Lehstenbach and Steinkreuz catchments in NE Bavaria, Germany". In: *Biogeochemistry of Forested Catchments in a Changing Environment*. Ed. by E. Matzner. Vol. 172. Ecological Studies. Berlin Heidelberg: Springer Verlag, pp. 15–44.

Bibliography III

Golyandina, N and E Osipov (2007). "The "Caterpillar"-SSA method for analysis of time series with missing values". In: *Journal of Statistical Planning and Inference* 137.8, pp. 2642–2653.

Huwe, B. and K. U. Totsche (1995). "Deterministic and stochastic modelling of water, heat and nitrogen dynamics on different scales with WHNSIM". In: *Journal of Contaminant Hydrology* 20.3-4. DOI: 10.1016/0169-7722(95)00073-9, pp. 265–284.