2-2021

**Multiscale Legacy Responses of Soil Gas Concentrations to Soil Moisture and Temperature Fluctuations**

Anthony J. Parolari
Joseph Sizemore
Gabriel Katul

Follow this and additional works at: [https://epublications.marquette.edu/civengin_fac](https://epublications.marquette.edu/civengin_fac)

Part of the Civil Engineering Commons
Multiscale Legacy Responses of Soil Gas Concentrations to Soil Moisture and Temperature Fluctuations

Anthony J. Parolari1, Joseph Sizemore2, and Gabriel G. Katul1,4

1Department of Civil, Construction, and Environmental Engineering, Marquette University, Milwaukee, WI, USA, 2Department of Electrical and Computer Engineering, Marquette University, Milwaukee, WI, USA, 3Nicholas School of the Environment, Duke University, Durham, NC, USA, 4Department of Civil and Environmental Engineering, Duke University, Durham, NC, USA

Abstract The sensitivity of soil carbon dynamics to climate change is a major uncertainty in carbon cycle models. Of particular interest is the response of soil biogeochemical cycles to variability in hydroclimatic states and the related quantification of soil memory. Toward this goal, the power spectra of soil hydrologic and biogeochemical states were analyzed using measurements of soil temperature, moisture, oxygen, and carbon dioxide at two sites. Power spectra indicated multiscale power law scaling across subhourly to annual timescales. Precipitation fluctuations were most strongly expressed in the soil biogeochemical signals at monthly to annual timescales. Soil moisture and temperature fluctuations were comparable in strength at one site, while temperature was dominant at the other. The effect of soil hydrologic, thermal, and biogeochemical processes on gas concentration variability was evidenced by low spectral entropy relative to the white noise character of precipitation. A full mass balance model was unable to capture high-frequency soil temperature influence, indicating a gap in commonly used model assumptions. A linearized model was shown to capture the main features of the observed and modeled gas concentration spectra and demonstrated how the means and variances of soil moisture and temperature interact to produce the gas concentration spectra. Breakpoints in the spectra corresponded to the mean rate of gas efflux, providing a first-order estimate of the soil biogeochemical integral timescale (∼1 min). These methods can be used to identify biogeochemical system dynamics to develop robust, process-based soil biogeochemistry models that capture variability in addition to long-term mean values.

Plain Language Summary The ability to describe how climate change impacts soil carbon and nutrient cycles with models is a necessary tool for ecosystem management and sustainability. One difficulty in developing these predictive models is the so-called “legacy effect”—for example, one wet summer may alter the ecosystem for many years afterward. Soil data and models are used here to quantify the relative strength of short- and long-term variability of soil biogeochemical systems and how it responds to rainfall, soil moisture, and soil temperature. We found that variability in soil biogeochemistry is concentrated at longer timescales of several weeks to months and this is because the soil stores water and heat, retaining a “memory” of past rainfall and temperature. Further, this analysis offered a new perspective on the equations used in current models—models driven by soil moisture and temperature are able to capture the legacy in soil biogeochemical data.

1. Introduction

Soil biogeochemical processes such as carbon and nutrient cycling and mineral weathering remain a recalcitrant problem in climate science, agriculture, and water resource management. Changes in soil thermal and hydrologic regimes have increased greenhouse gas release through thawing permafrost, enhanced peat decomposition, and increased fertilizer application, representing a positive feedback in the climate-carbon cycle system (Griffis et al., 2017; Heimann & Reichstein, 2008; Ise et al., 2008; McCalley et al., 2014). Yet, the sensitivity of soil carbon and nutrient cycles to climate change continues to pose a formidable challenge to the science of global change and sustainability research (Luo et al., 2016). For example, modeled and data-informed estimates of heterotrophic respiration and soil organic carbon sequestration carry uncertainties as high as 50–100% (Konings et al., 2019; Tian et al., 2015). Improved model representation of soil biogeochemical processes and how they respond to soil and atmospheric conditions remains, to say the least, a “first-order” problem to date.
Soil moisture and temperature variations are the primary physical drivers of soil biogeochemical process rates in the critical zone. Soil moisture and temperature influence rates of root and microbial respiration, decomposition, nutrient and metal cycling, and mineral weathering through physical, chemical, and biological pathways (Cook & Knight, 2003; Duncan et al., 2015; Jarecke et al., 2016; Kumagai & Kume, 2012; Manzoni & Porporato, 2009; Olshansky et al., 2019; Riveros-Iregui et al., 2007; Sierra et al., 2017; Wang et al., 2015; Warrinnier et al., 2020; Zhang et al., 2015). In general, reaction rates increase with soil temperature due to thermodynamic constraints, while the effect of soil moisture is more complex (Moyano et al., 2013). In dry soils, reaction rates increase with soil moisture as microbial osmotic stresses are overcome and soil pore connectivity increases substrate mobility (Manzoni & Katul, 2014). In contrast, as soil moisture increases above field capacity, metabolic process rates begin to decrease as metabolism shifts from aerobic to anaerobic pathways. This shift from aerobic to anaerobic processes results from a decrease in soil diffusivity in saturated soils, which limits resupply of atmospheric oxygen to the soil through diffusion. When oxygen is in limited supply, anaerobic metabolic processes dominate because they are more energetically variable in the absence of oxygen (Burgin et al., 2011). When anaerobic processes dominate, soil CO₂ production is reduced, while production of CH₄, N₂O, and N₂ is elevated, along with reduction of redox active metal hydroxides (e.g., Fe, Mn) (Burgin & Groffman, 2012; Jarecke et al., 2016; Liptzin et al., 2011; Warrinnier et al., 2020). Shifts in active soil biogeochemical processes and associated changes in pore water chemistry can occur rapidly during infiltration or recharge events (Austin et al., 2004; Duncan et al., 2015; Jarecke et al., 2016; Olshansky et al., 2019).

These processes are all mediated by the dynamics of complex networks of mineral aggregates, organic matter, microbes, and mycorrhiza. Pools of reaction substrates and mediators have residence times that range from days to decades or more (Bolker et al., 1998). In addition, vertical transport and gradients in substrate availability can induce storage effects that decouple deep processes from surface fluxes (Ryan et al., 2018). This wide range of timescales indicates that below-ground processes can retain a memory, or a legacy, of past soil and atmospheric hydrologic states (Vargas et al., 2011, 2010). Both soil moisture and temperature variability within the rooting zone are characterized by long-range auto-correlations that operate on differing timescales (Katul, Porporato, Daly, et al., 2007; Katul, Porporato, & Oren, 2007; Franzke et al., 2020). Disentangling the signature of such long-range dependence on physical and biochemical drivers of soil biogeochemical processes is a necessary first step for developing parsimonious modeling strategies to address uncertainties in the climate-carbon cycle feedback and soil sustainability.

Long-range dependence in auto-correlation is routinely analyzed using spectral methods. These methods comprise both linear and nonlinear time-series approaches to data analysis and model discovery (Halley & Inchausti, 2004; Kantz & Schreiber, 2004; Priestley, 1981). A first step toward identifying long-range dependence in an environmental process is analysis of the power spectrum. Power spectra derived from data and models can be combined to test model structures and to identify governing processes and error sources (Katul, Porporato, Daly, et al., 2007; MacMynowski & Tziperman, 2010; Hensley et al., 2018). When the process is described by a linear differential equation, this analysis can be carried out analytically (Katul, Porporato, Daly, et al., 2007; MacMynowski & Tziperman, 2010). Other methods, such as wavelet analysis (Torrence & Compo, 1998; Vargas et al., 2010) and various machine learning approaches (e.g., Rasp et al., 2018; Shen, 2018), complement the power spectrum, allowing for time-dependent estimation, and improved prediction; however, at the expense of additional degrees of freedom in the analysis.

To explore long-range dependence in soil biogeochemical systems and its dependence on climate, the spectral properties of soil gas concentration time series are analyzed in relation to precipitation, soil moisture, and temperature fluctuations. Specifically, we ask what modes of variability in precipitation, soil moisture, and soil temperature impact the strength and timescale of variability of physical and biogeochemical processes that control soil gas concentration fluctuations. The question is addressed using a combination of linear dynamical systems models and high frequency concentration data at two sites, one with soil oxygen and the other with soil carbon dioxide measurements.
2. Methods

2.1. Soil Dynamics in the Time Domain

As a logical starting point, a homogeneous soil defined by an active root-zone depth, $z_r$, porosity, $n$, and volumetric soil water content, $\theta = sn$ is considered, where $s$ is the degree of saturation ($s = 1$ when all soil pores are filled with water). A mass balance of a slightly or sparingly soluble gas such as carbon dioxide or oxygen, $M_g$, can be expressed in terms of its sources and sinks as

$$\frac{dM_g}{dt} = F_d[c,s,T] - F_r[c,s,T], \quad (1)$$

where $F_d$ is the diffusion flux from the atmosphere into the soil (for oxygen) or conversely (for carbon dioxide), $F_r$ is the respiration flux along with other source-sink terms (negative for oxygen, positive for carbon dioxide), $c$ is the gas phase concentration, and $T$ is soil temperature. In general, $F_d$ and $F_r$ can be functions of $c$, $s$, and $T$. The total mass is the sum of the gas and liquid phase contributions, which is assumed to be in chemical equilibrium with the gas phase, so that

$$M_g(s,T) = n c R_f(s,T)c, \quad (2)$$

where $R_f(s,T) = (1 - s) + s K_H RT$ is a variable retardation factor that depends on $s$ and $T$, $K_H$ is the Ostwald solubility coefficient, and $R$ is the universal gas constant (Davidson et al., 2006; Simunek, Jiri and Suarez, Donald L, 1993).

The gas phase diffusivity depends on $s$ and $T$, which when combined with a bulk model for $F_d$ results in

$$F_d[c,s,T] = D_b f_d(s)f_{dT}(T)[c_{atm} - c(T)], \quad (3)$$

where $D_b$ is the molecular diffusivity of the gas in free air normalized by a diffusion length scale, and $c_{atm}$ is the atmospheric gas concentration above the soil surface. The soil moisture and temperature adjustment factors in Equation 3 are assumed to be $f_d(s) = n^2[1 - s(t)]^{1.5}$ and $f_{dT}(T) = (T(t) / 293)^{1.25}$ (Daly et al., 2009; Moyano et al., 2013; Suwa et al., 2004; Turcu et al., 2005). This representation of $F_d$ is lumped or vertically integrated and does not explicitly resolve the contributions of each soil layer as conducted in detailed multilayer models described elsewhere (Daly et al., 2009; Fang & Moncrieff, 1999; Simunek, Jiri and Suarez, 1993; Suwa et al., 2004). It is assumed here that for carbon dioxide and oxygen, liquid-phase diffusion is negligible relative to gas-phase diffusion, which is a plausible assumption (Daly et al., 2009; Moldrup et al., 2000).

Lastly, gas-phase production or consumption through respiration and other biogeochemical transformations are represented as a product of a concentration-dependent respiration model, $r(c)$, adjusted by limitations arising from $s$ and $T$. That is,

$$F_r[c,s,T] = f_{rs}(s)f_{r,T}(T)r(c). \quad (4)$$

In general, the respiration function, $r(c)$, can take on a number of forms. Typical forms range from a constant source term to nonlinear Michaelis-Menten dependence on multiple substrates (Davidson et al., 2012; Zhang et al., 2015). As noted earlier, the respiration flux will be positive for carbon dioxide (produced) and negative for oxygen (consumed). The soil moisture and temperature adjustment functions in Equation 4 are assumed to be $f_{rs}(s) = e^{-ks}$ and $f_{r,T}(T) = \frac{Q_{10}}{Q_{10}^{T-25}}/10$ (Barron-Gafford et al., 2011; Suwa et al., 2004; Zhang et al., 2015).

Each of the variables (i.e., $c$, $s$, $T$) and their functions can be written in the form, $x(t) = \bar{x} + x'$, where the overbar denotes the temporal mean and the apostrophe denotes fluctuations around the mean. A linearized form of the model (1) can be obtained by making this substitution and retaining only the first-order terms, as

$$\frac{nc_R}{dt} \frac{dc'}{dt} = D_b\left(f_{ds}f_{dT}\Delta c' + f_{ds}\Delta c f_{dT} + f_{dT}\Delta c f_{ds} - f_{ds}f_{dT}c'\right) - k_r\left(f_{rs}f_{r,T}c' + f_{rs}\Delta c f_{r,T} + f_{r,T}\Delta c f_{rs} + f_{rs}\Delta c f_{r,T}c'\right) \quad (5)$$

PAROLARI ET AL.
where a linear respiration function \( k \) was assumed, \( \Delta C = C_{\text{atm}} - \bar{C} \) is the mean concentration deviation from the atmospheric value, and fluctuations in the retardation factor were assumed negligible.

### 2.2. Soil Dynamics in the Frequency Domain

The ordinary differential equation describing the soil gas mass balance defined in Equations 1–4 can be transformed from the time (or \( t \)) domain to the frequency (or \( f \)) domain by the Fourier transform to yield,

\[
\int_{-\infty}^{\infty} \frac{dM}{dt} e^{i\omega t} dt = \int_{-\infty}^{\infty} F_d[c,s,T] e^{i\omega t} dt - \int_{-\infty}^{\infty} F_c[c,s,T] e^{i\omega t} dt
\]

where the Fourier transform of a function \( g(t) \) is defined as

\[
\hat{g}(\omega) = \int_{-\infty}^{\infty} g(t) e^{-i\omega t} dt
\]

The Fourier transform of the linear model, Equation 5, can be solved analytically to obtain,

\[
\hat{c}(f) = \frac{\bar{C}}{\beta_c - if} \left[ \tau_d \left( \frac{C_{\text{atm}}}{\bar{C}} - 1 \right) \left( \hat{f}_{s,f} \hat{f}_{s,T}(f) + \hat{f}_{s,T} \hat{f}_{d,s}(f) \right) - \tau_i \left( \hat{f}_{s,f} \hat{f}_{s,T}(f) + \hat{f}_{s,T} \hat{f}_{d,s}(f) \right) \right]
\]

where

\[
\beta_c = \tau_d \hat{f}_{s,f} \hat{f}_{s,T} + \tau_i \hat{f}_{d,s} \hat{f}_{d,T}
\]

and \( \tau_d = D_0 / (nz \, R_f) \) and \( \tau_i = k_i / (nz \, R_f) \) are the unmodified diffusion and respiration frequencies of the linear system, respectively. The frequencies are in units of inverse time [T⁻¹] and the adjustment functions are unitless [−]. The \( \beta_c^{-1} \) is the integral timescale of the biogeochemical system.

The soil moisture and temperature adjustment functions take on different ranges of values and it is convenient to continue with normalized variables. The Fourier transforms of the variables and the adjustment functions normalized by their standard deviation are equivalent, i.e., \( \hat{f}_T(f) = \hat{T}(f) / \sigma_T = \hat{f}_{d,T}(f) / \sigma_{d,T} = \hat{f}_{s,T}(f) / \sigma_{s,T} \) (see Figure S1). With that substitution,

\[
\hat{c}(f) = \frac{\bar{C}}{\beta_c - if} \left[ \omega_s \hat{f}_s(f) + \omega_T \hat{f}_T(f) \right]
\]

where

\[
\omega_s = \tau_d \left( \frac{C_{\text{atm}}}{\bar{C}} - 1 \right) \hat{f}_{s,f} \hat{f}_{s,T} - \tau_i \hat{f}_{s,s} \hat{f}_{s,s}
\]

and

\[
\omega_T = \tau_d \left( \frac{C_{\text{atm}}}{\bar{C}} - 1 \right) \hat{f}_{d,T} \hat{f}_{d,T} - \tau_i \hat{f}_{d,s} \hat{f}_{d,s}
\]

The frequencies \( \omega_s \) and \( \omega_T \) are referred to as the soil moisture and temperature gain coefficients, respectively. Note that the coefficients in Equation 9 are crossing, meaning that the soil moisture gain is multiplied by the temperature transform and vice versa.

Equation 9 then leads to the squared Fourier amplitudes defining the power spectrum of \( c(t) \) as

\[
E_c(f) = \left( \frac{\bar{C}^2 \omega_s^2}{\beta_c^2 + f^2} \right) + \left( \frac{\omega_T^2}{\sigma_T^2} \right)
\]

where the power spectrum is \( E_c(f) = |\hat{c}(f)|^2 \).
Equation 12 relates the distribution of soil gas concentration variance across frequencies, $f$, to that of the external fluctuations in environmental parameters, $s$ and $T$. To a first order, the gas concentration spectrum, $E_c(f)$, is a superposition of $E_T(f)$ and $E_s(f)$ that is modulated relative to these spectra by the timescales inherent to soil physical (diffusion) and biogeochemical (dissolution, respiration) processes. These effects are represented by the coefficients $\omega_c$, $\omega_s$, and $\omega_T$, which each contain contributions from the timescales $\tau_r$ and $\tau_d$ and the means and standard deviations of the $s$ and $T$ adjustment functions. Naturally, this simplistic representation can be criticized ad-infinitum given the bulk representation and the multiple linearizations adopted. Nonetheless, it does serve as a reference (or a null) model to interpret multiscaled data and provide causal links between environmental variability (e.g., precipitation and temperature) and soil gas concentration fluctuations.

Another advantage of this representation is a clear definition of the biogeochemical system spectral transfer function, which relates the ratio of the concentration spectra (output) to environmental variability (input) (Stoy et al., 2009). The variance-normalized spectral transfer function for input variable, $i$, is defined as $|H_i(f)|^2 = E_i(f)/\sigma_i^2$. The transfer functions quantify the frequency response of the biogeochemical system for a given input (MacMynowski & Tziperman, 2010; Stoy et al., 2009). Transfer functions for soil moisture, temperature, and precipitation are presented below. Although precipitation is not included in the model here, the spectral characteristics of precipitation are expressed in the soil moisture signal at the observed frequencies (Katul, Porporato, Daly, et al., 2007).

### 2.3. Spectral Entropy

The spectral entropy (SE) quantifies the relative amount of deterministic, process-driven information contained within each time series. The SE here is defined as the Shannon information entropy of the power spectral density function normalized by the variance, $E(f)/\sigma^2$. As discussed elsewhere, applying the Shannon information entropy to spectra (instead of probability of the time series of the variable) is acceptable because the normalized spectrum abides by the same probability axioms of information entropy (Wesson et al., 2003). Hence, the discrete probability of variance associated with frequency bin $i$ can be defined as,

$$p_i = \frac{E_i(f)\Delta f}{\sigma^2}$$

and the SE can then computed as,

$$SE = -\sum_{i} p_i \log(p_i).$$

where $m$ is the number of frequency bins over which the sum is conducted.

The range of SE is confined to $0 < SE < \log(m)$ and the maximum possible value, $\log(m)$, corresponds to white noise. Larger values of SE indicate a more uniform distribution of variability across frequencies and, therefore, a more stochastic (and less deterministic) signal. In the limit of $p_i = 1/m$, all frequencies contribute equally to the variance. This state can be labeled as the most ‘disordered’ because there is no dominant timescale explaining the variability in gas concentration fluctuations. On the other hand, smaller values of SE indicate a less uniform distribution of variance across frequencies (or timescales) and thus a more deterministic, process-driven state. In this case, the activity or variability of the process is controlled by few modes with predefined timescales. If those timescales are identified (and linked to $\tau_r$ and $\tau_d$), then the associated biogeochemical processes operating on such timescales can be delineated and used in any causal explanation.

Because the spectral density functions are of different lengths (between the study sites), the SE of each time series is reported relative to its maximum possible value, i.e., $0 < SE/\log(m) < 1$ (or relative entropy). In terms of analog to entropy in statistical mechanics (i.e., $S_b = k_B \log(W_b)$, where $k_B$ is the Boltzmann constant and $W_b$ is the number of arrangements the microstate can attain in a system) it is worth noting that variability (or energy) at each $f$ is analogous to a “micro” state, whereas SE reflects an aggregate or “macro” state of the fluctuations.
3. Data

Measurements from two field experiments are used to assess the spectral links between precipitation, soil moisture, and temperature, and gas concentration. The first experiment was conducted in an urban agricultural soil in Milwaukee, Wisconsin, from April to September, 2019. Volumetric soil water content, soil temperature, and oxygen concentration (Apogee SO110) were measured at 5-min intervals at a depth of 15 cm. Hourly precipitation was obtained from the Milwaukee airport rain gauge located ∼10 miles southeast of the study location. The second dataset was obtained from the Duke Forest free air carbon dioxide enrichment experiment conducted in a Loblolly pine forest near Durham, North Carolina, 1998–2005. These data included half-hourly measurements of precipitation above the canopy, interception losses, volumetric water content, soil temperature, and soil carbon dioxide (Vaisala GMT 220) concentration at multiple depths. The data used here were measured by soil sensors nearest to the surface, installed between 2 and 11 cm depth (Daly et al., 2008, 2009). The two datasets provide different soil, vegetation, and climate contexts for the analysis. Soil properties and model parameters for each site are provided in Table 1. Time-series data are provided in the supplemental material, Figure S2.

The power spectra of the environmental variables and gas concentrations were approximated with the general expression,

$$E(f) \propto \frac{1}{\beta^\alpha + f^\alpha}$$

where $\alpha$ is the slope of $E(f)$ versus $f$ in log-log space and $\beta$ is an integral frequency where the spectrum displays a breakpoint in the slope. The estimated values of $\alpha$ and $\beta$ are summarized in Table 2. When $\beta$ is small relative to the frequencies measured, the well-known $1/f$ noise is recovered (Vasseur & Yodzis, 2004). The physical origin of this expression is apparent in the role of $\beta$ in Equation 12 and is derived for reference in Appendix A.

4. Results and Discussion

The generic spectral properties of the forcing variables (precipitation, soil moisture, and soil temperature) and the soil biogeochemical state variables (oxygen and carbon dioxide concentrations) are first introduced. Next, the behavior of the soil biogeochemical system is demonstrated using the Shannon Entropy as a measure of organization and the linear
transfer functions as an indicator of how the soil system processes physical hydro-climatic inputs into biogeochemical outputs. Finally, the spectral linkage between the forcing and response variables is discussed using the simplified model introduced above.

4.1. Empirical Power Spectra

For the Milwaukee site, the power spectrum of precipitation was approximately “white” for subdaily frequencies \((f^{-1} < 24 \text{ h})\) and was approximately “pink”, decaying according to a power-law \(f^{-1}\), at higher frequencies (Figure 1a). The power spectrum of soil moisture decayed at a rate \(f^{-1.7}\) for low frequencies and \(f^{-2.7}\) for high frequencies, indicating that spectral decay of soil moisture was larger relative to that of precipitation by a factor of 1.7. The power spectrum of temperature decayed at a rate \(f^{-1.2}\) for low frequencies and \(f^{-3.5}\) for high frequencies, demonstrating large damping of high frequency variability expected for the high thermal inertia of soils. The temperature spectrum also showed a strong diurnal signal at 12 and 24 h, as expected.

The power spectrum of soil oxygen retained some modes expressed by the driving variables, but also contrasted their behavior particularly at low frequencies. Similar to temperature, the oxygen spectrum showed peaks at 12- and 24-h frequencies, indicating the diurnal control of temperature on respiration (Riveros-Iregui et al., 2007; Zhang et al., 2015). Soil oxygen exhibited three spectral scaling regimes, with slopes of \(-0.9\) for low frequencies, \(-1.8\) for intermediate frequencies, and \(-3.8\) for high frequencies. For intermediate and high frequencies, the soil oxygen spectrum displayed similar behavior to the soil moisture and temperature spectra. However, it decayed more slowly for low frequencies less than \(\sim 10^{-2} \text{ hr}^{-1}\), which corresponds to approximately weekly timescale.

At the Duke Forest site, precipitation and soil moisture showed similar behavior as the Milwaukee site (Figure 1b). The precipitation spectrum decayed at frequencies higher than \(\sim 1 \text{ day}\) and the soil moisture spectrum was dampened relative to precipitation, decaying according to \(f^{-1.75}\) relative to the precipitation spectrum. Similar to the Milwaukee site, the Duke Forest temperature spectrum was more dampened at high frequencies, decaying at a rate \(f^{-1.25}\) for low frequencies and \(f^{-2.75}\) for high frequencies. The temperature spectrum also showed a strong diurnal signal at 12 and 24 h.

In contrast to oxygen concentration at the Milwaukee site, the Duke Forest carbon dioxide power spectrum was more similar to soil moisture and temperature. The carbon dioxide spectrum exhibited two similar scaling regimes, decaying at a rate \(f^{-1.5}\) for low frequencies and \(f^{-2.25}\) for high frequencies.

The behaviors of the precipitation and soil moisture spectra shown here have been observed and described previously. Decay of the precipitation spectrum at frequencies larger than 1 day reflects serial correlation among intrastorm rainfall pulses, whereas the approximately white noise at lower frequencies reflects the independence of individual storms. Independence of daily rainfall amounts is often assumed in Poisson-distributed precipitation models driving many stochastic ecohydrological models (e.g., Parolari et al., 2014). The soil moisture spectrum decays much more rapidly than precipitation due to storage in the root zone (i.e., the so-called reddening effect), the implications of which to the climate system are reviewed elsewhere (Delworth & Manabe, 1988). The steepening of the soil moisture spectrum relative to the precipitation spectrum by a power of 2, i.e., \(E_{s}(f) \sim f^{-2}E_{p}(f)\) was well predicted by a bulk soil water balance with evapotranspiration losses specified as a linear function of soil moisture (Katul, Porporato, Daly, et al., 2007; Nakai et al., 2014; Ghannam et al., 2016). The powers estimated here, 1.7 and 1.75, are similar to this theoretical prediction.

4.2. Spectral Entropy

The spectral entropy values calculated from the observed time series are summarized in Table 3. The values are reported relative to a white noise spectrum such that a value of 1 corresponds to a completely random, white noise (or flat) spectrum and a value of 0 corresponds to a completely deterministic spectrum where all of the power is concentrated at a single frequency.

The precipitation series are the most random with a relative spectral entropy of 0.92 for Milwaukee and 0.97 for Duke Forest. Therefore, both precipitation series are nearly white.
Figure 1. Empirical power spectral density normalized by variance, $E(f)/\sigma^2$, for precipitation, soil moisture, temperature, and soil gas concentration for (a) oxygen data at Milwaukee, Wisconsin, April–September 2019 and (b) carbon dioxide data at Duke Forest, North Carolina, 1998–2005. The black lines represent the power spectra fit to Equation 15 and the estimated slopes are annotated. The estimated slopes and breakpoints are also tabulated in Table 2.
The soil moisture and temperature series are more organized in the spectral domain compared to precipitation. The spectral entropy at Milwaukee was 0.33 for temperature and 0.27 for soil moisture, whereas the values were 0.18 and 0.35 at Duke Forest, respectively. Because these values are closer to 0 than 1, they indicate that the soil physical drivers of respiration are process-dominated, but retain \( \sim 20\% - 35\% \) of the stochastic variability in the environmental forcing. This result is consistent with the substantial spectral decay relative to the precipitation spectrum (Figure 1). The smaller spectral entropy for temperature relative to soil moisture at Duke Forest is likely related to either stronger diurnal temperature variations or larger spectrum coefficients at the lowest frequencies.

Finally, the spectral entropy of \( \text{O}_2 \) at Milwaukee was 0.45, whereas that of \( \text{CO}_2 \) at Duke Forest was 0.31. For Milwaukee, this indicates that the \( \text{O}_2 \) series retained more of the stochastic character of the precipitation series relative to temperature and soil moisture. At Duke Forest, the opposite was observed. That is, the \( \text{CO}_2 \) series had a spectral entropy similar to soil moisture and greater than temperature. Again, this result is consistent with the power spectra. The \( \text{O}_2 \) spectrum decayed slowly at low frequencies relative to the other variables, making it the most similar to the white noise spectrum. In contrast, the \( \text{CO}_2 \) spectrum was most similar to the soil moisture spectrum.

The spectra and their spectral entropies are evidence for a physical mechanism that modulates the relation between precipitation, soil moisture, temperature, and respiration fluctuations across multiple scales. In the next section, empirical transfer functions are introduced to quantify the transfer of energy from physical to biogeochemical variables across frequencies.

### 4.3. Empirical Transfer Functions

The soil biogeochemical system filtered fluctuations in the physical drivers at different timescales, as shown by the empirical transfer functions (i.e., \( H(f) \)) illustrated in Figure 2. For frequencies where the transfer function exceeds 1, the soil gas concentration exhibits more variability at that frequency than the environmental driver (and vice versa).

![Figure 2](image-url)
At the Milwaukee site, oxygen concentration was more variable than soil moisture and temperature across most scales (Figure 2a). O₂ exhibited less variability than precipitation at high frequencies and approximately the same at low frequencies. The soil moisture and temperature transfer functions were of similar order of magnitude. However, soil moisture was strongest at intermediate frequencies between two days and 10 h, while temperature was largest at frequencies greater than 10 h. Therefore, the soil acted as a low-pass filter on precipitation, high-pass filter on temperature, and a band-pass filter on soil moisture fluctuations. The multiscale character of the soil O₂ spectrum can be related to how the driving variables control oxygen variability across timescales: soil moisture and temperature contributed similarly at intermediate timescales and temperature dominated at subdaily timescales.

At Duke Forest, carbon dioxide concentration was more variable than temperature across most frequencies and slightly less variable than soil moisture (Figure 2b). Similar to Milwaukee, precipitation variability was damped in the CO₂ signal at high frequencies and amplified at low frequencies (<10⁻² hr⁻¹). Therefore, the soil acted as a band-pass filter on temperature and a low-pass filter on precipitation. Temperature and precipitation variability were primarily expressed in CO₂ variability at this site. Temperature and CO₂ concentration exhibit strong seasonality at this site (see Figure S2). Further, strong relations between throughfall depth and CO₂ concentration jumps during rain events were observed and analyzed elsewhere (Daly et al., 2008), supporting the findings here.

### 4.4. Modeled Power Spectra

The soil gas concentration power spectra were determined by numerically solving the full model (Equations 1–4) and with the simplified linear model (Equation 12).

The full model performed well compared to the time-series data for both sites (Figure 3). The O₂ concentration at Milwaukee is much more strongly influenced by diurnal temperature fluctuations (Figure 3a) than the CO₂ concentration at Duke Forest, which is more strongly impacted by rainfall pulses that arrive every few days (Figure 3c). The most conspicuous discrepancy between the model and data is the large predicted decrease in O₂, which occurred during periods of high soil moisture.

The full model also reproduced the gas concentration spectra at both sites. For Milwaukee (Figure 4a), the O₂ spectrum estimated from the full model matched that from the data best for low frequencies, \( f < 10^{-1} \text{ hr}^{-1} \) (i.e., \( \sim 10 \text{ h} \)). In contrast, the full model underpredicted the empirical power spectrum at higher frequencies. Note that the model captured the spectral slope at high frequencies (i.e., \( \sim 4 \)), but overestimated the slope at intermediate frequencies. For Duke Forest (Figure 4b), the full model captured the empirical power spectrum across all but the highest frequencies. At high frequencies, \( f > 0.5 \text{ h}^{-1} \) (i.e., \( \sim 2 \text{ h} \)), the spectral decay was overestimated by the full model. Overestimates of spectral decay at both sites indicate that the full model does not accurately preserve spectral content at high frequencies.

The empirical transfer functions and numerical simulations informed further analysis of the linearized model to provide analytical foresight into how the soil biogeochemical system filters fluctuations in the physical driving variables. Contributions of each driving variable and process to the gas concentration dynamics depends on the soil system coefficients \( \beta_s, \omega_s, \) and \( \omega_T \). Estimates of these coefficients from the full model parameters in Table 1 are shown in Table 4. The magnitudes of the coefficients are compared below in a discussion of the primary drivers of gas concentration dynamics at each site.

The coefficients \( \beta_s^{-1} \) can be interpreted as the integral timescales of the soil hydrologic (\( \beta_s \)), thermal (\( \beta_T \)), and biogeochemical (\( \beta_c \)) systems, an index of the soil “memory.” Soil storage of moisture, energy, and chemical constituents effectively acts as a low-pass filter on stochastic environmental drivers. Frequency modulation due to storage is determined by the ratio of the storage capacity to the flux. (Delworth & Manabe, 1988; Katul, Porporato, Daly, et al., 2007). For example, the soil moisture memory timescale was previously derived as \( \beta_s^{-1} = \frac{n_s}{ET_{\text{max}}} \), the ratio of soil water holding capacity to the maximum evapotranspiration rate (Katul, Porporato, Daly, et al., 2007). Similarly, the estimated parameter \( \beta_T^{-1} \) can be interpreted as a memory of the surface energy balance associated with temperature fluctuations damped by thermal dissipation from the surface. The biogeochemical memory estimated here takes a similar form, i.e., \( \beta_c^{-1} = \frac{n_s}{R_f} \left[ D_h \bar{f}_{d,T} \bar{f}_{d,T} + k_f \bar{f}_{f,f_T} \right] \). This expression is the ratio of the soil gas storage capacity to the
mean flux due to diffusion and respiration. Note that the soil gas storage capacity includes the retardation factor which represents equilibrium reactions that control the gas phase concentration. A more soluble gas has a larger storage that takes longer to ‘empty’ via diffusion and, therefore, is expected to amplify more the input signal at low frequencies. It is anticipated that the biogeochemical integral timescale will also depend on soil pH, alkalinity, and other properties (Calabrese et al., 2017).

The integral timescales appear as breakpoints in the slope of the $E(f)$ versus $f$ graph in log-log space. The model estimated values are $102 \, \text{h}^{-1}$ for Milwaukee and $39.5 \, \text{h}^{-1}$ for Duke Forest (i.e., $\sim 1 \, \text{min}$), respectively (Table 4). These values are larger than the largest frequencies measured at either site and, therefore, are not apparent in the observed or modeled spectra. The hydrologic memory, estimated here for Duke Forest as 112 days, is comparable to (Duke Forest) or smaller than (Milwaukee) the lowest frequencies measured and, therefore, is also not apparent in the data. The most prominent breakpoint identified in the concentration spectrum is that associated with the temperature spectrum, or the soil thermal memory. The thermal memory was estimated as 0.56 days for Milwaukee and 3.5 days for Duke Forest.

Figure 3. Results of the numerically integrated full model: (a)–(b) Milwaukee O$_2$ and (c)–(d) Duke Forest CO$_2$. 
The coefficients $\omega_s$ and $\omega_T$ weight the contributions of the soil moisture and temperature spectra in the gas concentration spectrum. According to Equations 10 and 11, these coefficients characterize the bulk rate of gas movement into or out of the soil, or the difference between the mean diffusion and respiration rates, multiplied by the standard deviations of the adjustment functions.

For the O$_2$ dynamics at Milwaukee, the dominant controls were the soil moisture effect on respiration and the temperature effect on diffusion. These effects were of a similar order of magnitude (Table 4), indicating that a combination of soil moisture and temperature variability is important for predicting the soil gas concentration variability, consistent with the empirical transfer functions. A comparison of the full linear model and an approximation driven by temperature only demonstrates that the temperature spectrum dominates at intermediate frequencies and the soil moisture spectrum dominates at high and low frequencies (Figure 4a). Note, because the coefficients are crossing, it is the magnitude of the soil moisture effect on respiration that emphasizes the temperature spectrum at intermediate frequencies and the temperature effect on diffusion that emphasizes the soil moisture spectrum at high and low frequencies. The linear model structure elucidates how the two variables, $s$ and $T$, interact to produce the concentration variance and its spectrum.

The coefficients $\omega_s$ and $\omega_T$ weight the contributions of the soil moisture and temperature spectra in the gas concentration spectrum. According to Equations 10 and 11, these coefficients characterize the bulk rate of gas movement into or out of the soil, or the difference between the mean diffusion and respiration rates, multiplied by the standard deviations of the adjustment functions.

For the O$_2$ dynamics at Milwaukee, the dominant controls were the soil moisture effect on respiration and the temperature effect on diffusion. These effects were of a similar order of magnitude (Table 4), indicating that a combination of soil moisture and temperature variability is important for predicting the soil gas concentration variability, consistent with the empirical transfer functions. A comparison of the full linear model and an approximation driven by temperature only demonstrates that the temperature spectrum dominates at intermediate frequencies and the soil moisture spectrum dominates at high and low frequencies (Figure 4a). Note, because the coefficients are crossing, it is the magnitude of the soil moisture effect on respiration that emphasizes the temperature spectrum at intermediate frequencies and the temperature effect on diffusion that emphasizes the soil moisture spectrum at high and low frequencies. The linear model structure elucidates how the two variables, $s$ and $T$, interact to produce the concentration variance and its spectrum.

### Table 4

<table>
<thead>
<tr>
<th></th>
<th>Milwaukee</th>
<th>Duke Forest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Diffusion</td>
<td>Respiration</td>
</tr>
<tr>
<td>Soil moisture gain, $\omega_s$</td>
<td>0.034</td>
<td>-0.45</td>
</tr>
<tr>
<td>Temperature gain, $\omega_T$</td>
<td>0.44</td>
<td>0.031</td>
</tr>
<tr>
<td>Integral timescale, $\beta_c$</td>
<td>92.1</td>
<td>9.9</td>
</tr>
</tbody>
</table>

Note. The diffusion and respiration contributions refer to the first and second terms, respectively, in Equations 10 and 11. For example, the soil moisture gain, $\omega_s$, is composed of diffusion, $\tau_d \left( \frac{c_{am}}{c} - 1 \right) \sigma_{d,T}$, and respiration, $-\tau_r \sigma_{r,T}$, components.
For Milwaukee, the numerical model and its linear approximations all underpredict the \( \text{O}_2 \) spectrum at frequencies greater than \( f = 10^{-1} \ \text{hr}^{-1} \), which corresponds to approximately 10 h timescale (Figure 4a). The empirical transfer functions (Figure 2a) demonstrate that temperature variability is expressed in the \( \text{O}_2 \) fluctuations at these high frequencies. This inconsistency between the models and data highlights that the temperature adjustment factors of respiration and diffusion, \( f_{r,T} \) and \( f_{d,T} \), respectively, used in the model may not be sensitive enough to retain temperature variability at these frequencies. Evidence exists for diurnal regulators of heterotrophic respiration other than temperature, such as photosynthesis and root exudation of carbohydrates (Dettto et al., 2012; Stoy et al., 2007; Tang et al., 2005; Wan & Luo, 2003), which may offer one hypothesis to close this gap.

Similarly, for the \( \text{CO}_2 \) dynamics at Duke Forest, the soil moisture effect on respiration and temperature effect on diffusion were the dominant controls (Table 4). The soil moisture effect on diffusion and the temperature effect on respiration were less important. The linear model including only the temperature spectrum gives a good approximation of the \( \text{CO}_2 \) spectrum (Figure 4b). The model is consistent with the empirical transfer functions (Figure 2b), which show temperature fluctuations are amplified broadly across all frequencies.

The linkages between the environmental variables and concentration spectrum can be further illustrated by constructing a minimal model guided by the relative magnitudes of the coefficients in the linear model. For Duke Forest, \( \omega^2_T \gg \omega^2_f \) and \( \omega^2_f \gg \omega^2_r \). In addition, note that production of \( \text{CO}_2 \) is independent of the soil \( \text{CO}_2 \) concentration—that is, the reaction is not limited by the reactant concentration. In that case, a constant respiration function, \( r(c) = r \), applies. It follows that the concentration terms in the respiration expression of Equation 5 are dropped and \( \omega_r = \tau_df_r, f_{d,T} \). Under those conditions, Equation 12 can be written as,

\[
E_r(f) = \left[ \frac{\tau_d}{\tau_r} \left( \frac{c_{\text{atm}}}{c} - 1 \right)f_d, \sigma_{d,T} = \tau_r, f_{d,T}, \sigma_{T,F} \right]^2 \frac{1}{\beta_T^{1.5} + f^{1.5} f^{1.25}}.
\]

Equation 16 predicts that the \( \text{CO}_2 \) gas concentration spectrum exhibits two regimes following the temperature spectrum. At low frequencies, the \( \text{CO}_2 \) spectrum displays a nearly pink noise regime with \( E_r \sim f^{-1.5} \) for \( f < \beta_T \). At high frequencies, the spectrum follows a "red" regime with \( f^{-2.75} \), for \( f > \beta_T \). The linear model captures all the essential regimes in the observed \( \text{CO}_2 \) spectrum (Figure 4b). This multiscale feature of the \( \text{CO}_2 \) spectrum is emphasized when the spectrum is multiplied by \( f \), yet is obscured by sensor noise at high frequencies.

5. Conclusions

The spectrum of soil gas concentrations in relation to hydro-climate variables was studied using empirical and modeling approaches. Power spectra, transfer functions, and spectral entropy were quantified from concurrent time-series of precipitation, soil moisture, temperature, and soil gas concentrations (\( \text{O}_2 \) and \( \text{CO}_2 \)) at two locations with contrasting soil, climate, and vegetation types. Mass balance models of varying complexity (full nonlinear and linear) were then applied to the data to identify the mechanisms linking variability in external driving variables to the soil biogeochemical state variables.

The data demonstrated certain similarities and differences between the two sites. At both sites, the mean effects of soil moisture and temperature were important, while soil temperature variability dominated the soil gas concentration variability and was key to modeling the system in the time and frequency domains. At the Milwaukee site, the contributions of soil moisture and temperature variability were of similar magnitude with soil moisture more dominant at intermediate frequencies and temperature more important for high, subdaily frequencies. In contrast, at the Duke Forest site, soil temperature variability was the variable most strongly expressed by the soil biogeochemical system, whereas soil moisture variability was less apparent across all frequencies and precipitation was expressed at the lowest frequencies of multiple weeks. The soil \( \text{O}_2 \) data were more stochastic in character and the soil \( \text{CO}_2 \) data appeared more organized, or process-driven. Therefore, the two systems processed precipitation, soil moisture, and temperature fluctuations differently.

Theoretical expressions were derived for the power spectrum of soil gas concentration that generally reflected the conclusions from the data. In general, the soil gas concentration spectrum is a superposition...
of the soil moisture and temperature spectra that is damped by the diffusion and respiration processes. The weighting of the two spectra is determined by the relative values of the soil moisture and temperature fluctuations. Milwaukee O₂ variability at low frequencies was well predicted by a model driven by temperature and soil moisture, while the full model was unable to capture the high frequency variability due to temperature. This highlighted a potential gap in currently modeling capabilities at these scales. Duke Forest CO₂ data were well predicted across all frequencies by temperature fluctuations only. In both cases, the “memory” of the soil biogeochemical system was reasonably approximated by the timescale associated with diffusion, on the order of a few minutes. This indicates that strong damping by the soil biogeochemical system only occurs at very high frequencies. However, damping of soil moisture and temperature due to storage is propagated to the soil gas concentrations, which is associated with timescales greater than several days or weeks. The soil thermal memory of 0.5–3.5 days was the damping timescale most apparent in the system only occurs at very high frequencies. Therefore, models that operate at the daily scale and higher may be able to capture most soil gas concentration variability.

### Appendix A: Power Spectrum of Environmental Variables

For reference, a general expression for the power spectrum of an environmental variable is derived here. Consider a variable \( x(t) \) that varies in time only, is driven by random noise \( \xi(t) \), and decays according to a function \( L \) \([x(t)]\). A balance equation for \( x(t) \) can be written as,

\[
\frac{dx(t)}{dt} = \xi(t) - L[x(t)]
\]

(A1)

with the corresponding Fourier transform,

\[
-if\hat{\xi}(f) = \hat{x}(f) - \hat{L}(f).
\]

(A2)

For a linear loss function, i.e., \( L(x) = -kx \), Equation A2 can be solved for the Fourier transform of \( x(t) \),

\[
\hat{x}(f) = \frac{1}{k-if}\hat{\xi}(f).
\]

(A3)

The power spectrum is then obtained as the squared amplitudes of the Fourier spectrum,

\[
E_x(f) = \left|\frac{1}{k-if}\hat{\xi}(f)\right|^2 = \frac{1}{k^2 + f^2}E_{\xi}(f).
\]

(A4)

Equation A4 shows that the power spectrum of the variable \( x(t) \) is related to the spectrum of the noise \( \xi(t) \) as well as the rate of the decay process, \( k \). The rate \( k \) is interpreted as the integral timescale or memory of the decay process. When \( f \ll k, E_x(f) \propto E_{\xi}(f) \) by the factor \( k^{-2} \). Alternatively, when \( f \gg k \), the spectrum of \( x(t) \) is dampened relative to \( \xi(t) \), i.e., \( E_x(f) \propto f^{-2}E_{\xi}(f) \). Therefore, the signatures of the external noise and internal process on the variability of \( x(t) \) can be distinguished from the shape of the power spectrum.

### Data Availability Statement

Data are available from the Dryad Digital Repository https://doi.org/10.5061/dryad.v41ns1rtv (Parolari, 2020).

### References


---

**Acknowledgments**

A. J. Parolari and J. Sizemore acknowledge support from the Fund for Lake Michigan and the Marquette University Opus College of Engineering Summer Undergraduate Research Fellowship. G. G. Katul acknowledges support from the U.S. National Science Foundation (NSF-AGS-1644382, NSF-IOS-1754893, and NSF-AGS-2028633).


