Characterizing land-surface processes: A quantitative analysis using air-ground thermal orbits

Jason E. Smerdon

Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY, USA

Hugo Beltrami, Chance Creelman, and M. Bruce Stevens

Environmental Sciences Research Centre, St. Francis Xavier University, Antigonish, Canada

Jason E. Smerdon, Lamont-Doherty Earth Observatory of Columbia University, 61 Route 9W, Palisades, NY 10964, USA. (jsmerdon@ldeo.columbia.edu)

Hugo Beltrami, Environmental Sciences Research Centre, Department of Earth Sciences, St. Francis Xavier University, 1 West Street, Antigonish, Nova Scotia, Canada, B2G 2W5. (hugo@stfx.ca)

Chance Creelman, Environmental Sciences Research Centre, Department of Earth Sciences, St. Francis Xavier University, 1 West Street, Antigonish, Nova Scotia, Canada, B2G 2W5.

M. Bruce Stevens, Environmental Sciences Research Centre, Department of Earth Sciences, St. Francis Xavier University, 1 West Street, Antigonish, Nova Scotia, Canada, B2G 2W5.
Abstract. A quantitative analysis of thermal orbits is developed and applied to modeled air and ground temperatures. Thermal orbits are phase-space representations of air and ground temperature relationships that are generated by plotting daily or monthly ground temperatures against air temperatures. Thermal orbits are useful descriptive tools that provide straightforward illustrations of air and ground temperature relationships in the presence of land-surface processes related to snow cover, soil freezing, and vegetation effects. The utility of thermal orbits has been limited, however, by the lack of quantitative analyses that describe changes in orbits across different environments or in time. This shortcoming is overcome in the present study by developing a linear regression analysis of thermal orbits that allows changes to be tracked in time, space and as a function of depth within the subsurface. The theory that underlies the thermal-orbit regression analysis is developed herein and the utility of the application is demonstrated using controlled model experiments.
1. Introduction

The thermal regime of the continental subsurface is well established as an important heat reservoir within the climate system [Beltrami et al. 2002, 2006; Pielke 2003; Levitus et al. 2005; Huang 2006], as well as a control over crucial conditions and processes such as the extent and stability of permafrost [e.g. Romanovsky et al. 2002; Lawrence and Slater 2005], soil respiration and carbon fluxes [e.g. Risk et al. 2002a, 2002b; Kellman et al. 2007; Riveros-Iregui et al. 2007], and the length of growing seasons [e.g. Foster et al. 1992; Stone et al. 2002; Walsh et al. 2005]. Modeling studies have further shown the importance of realistic simulations of thermal conditions in the terrestrial subsurface and associated air-ground interactions for characterizing the land-surface energy balance and related climate feedbacks [e.g. Zhu and Liang 2005; Seneviratne et al. 2006; Davin et al. 2007; Fischer et al. 2007; Miguez-Macho et al. 2005, 2007], the evolution of soil temperature and permafrost [e.g. Sun and Zhang 2004; Smerdon and Stieglitz 2006; Sushama et al. 2006, 2007; Nicolsky et al. 2007; Aleexeev et al. 2007] and the overall partitioning of heat within the climate system [Stevens et al. 2007; MacDougall et al. 2008; González-Rouco et al. 2009]. Temperatures in the terrestrial subsurface are also used as paleoclimatic indicators of surface temperature changes by means of direct temperature-profile inversions [e.g. Huang et al. 2000; Harris and Chapman 2001; Beltrami 2002; Pollack and Smerdon 2004] or through their control over the noble gas content of groundwater during recharge [Stute and Schlosser, 1993]. Given these various areas of import, there has been a concerted effort to characterize and understand the processes that govern the evolution of the subsurface thermal regime and the coupling between the land and atmosphere [e.g. Lin et al., 2003;
A fundamental approach to understanding how the thermal regime of the terrestrial subsurface evolves is through analyses of observed air and ground temperatures. This approach has been a foundation of soil and permafrost research for many years [e.g. Geiger, 1965; Lachenbruch, 1959; de Vries, 1975; Smith, 1975; Goodrich, 1982]. More recently, studies (such as those referenced in the preceding paragraph) have sought to place analyses of air and ground temperatures in the broader context of the coupling between the land and atmosphere, the carbon cycle, and the evolution of regional and global climate phenomena. Given these various purposes, many different kinds of analyses have been developed to assess land-surface conditions and ground temperature evolution over varying time scales. These analyses range from descriptive to quantitative and can include various statistical and spectral approaches [e.g. Carson, 1963; Outcalt and Hinkel, 1992; Hinkel and Outcalt, 1993; Smerdon et al. 2003, 2004, 2006; Frauenfeld et al. 2004; Amenu et al. 2005; Anctil et al. 2008]. Among these, Beltrami [1996] introduced a phase-space representation of air and ground temperatures, termed thermal orbits, that provides a semi-quantitative description of the coupling between air and ground temperatures. This approach plots ground temperature versus air temperature time series (see Figure 1), yielding elliptical phase-space orbits with characteristics that depend on the
amplitude attenuation and phase shift of the propagating subsurface signal, relative to the air temperature signal (analogous to the well-known Lissajous figures in electronic circuit analysis, e.g. Malmstadt et al. [1981]). Importantly, these thermal orbits are bent or perturbed by land-surface processes connected, for example, to snow cover or freeze-thaw cycles, which can be readily identified in thermal orbit figures.

While thermal orbits have proved useful in later studies [e.g. Beltrami 2001; Sushama et al. 2007; Anctil et al. 2008], their utility has been hampered by the fact that they do not alone provide a quantitative measure that can be tracked spatially, temporally or as a function of depth within the subsurface. For instance, Sushama et al. [2007] use mean thermal orbits as one characterization of simulated permafrost conditions in North America using a regional climate model projection for the 21st-century. The plotted thermal orbits in Sushama et al. [2007] are for current and projected permafrost zones and illustrate convincing, but only qualitative, changes in the simulated permafrost zones. It therefore would be useful to develop an analysis that measures changes in thermal orbits for comparisons between other model simulations or observational data. Furthermore, a thermal-orbit analysis has the advantage of summarizing a diverse number of changes in processes and conditions at the land-atmosphere boundary into a single and quantitative measure.

Here we develop a method of characterizing changes in thermal orbits using a least-squares linear regression analysis. We revisit some of the theory that underlies phase-space representations of air and ground temperatures and further discuss how they change given perturbations to the amplitude and phase of propagating subsurface thermal waves. We derive the theoretical regression slope equations for thermal orbits as a function of depth.
in a one-dimensional conductive medium and perform controlled model experiments to
demonstrate the evolution of thermal-orbit slopes under changing air and ground temper-
ature conditions. We further track changes in slopes of thermal orbits at two contrasting
grid locations in the millennial simulation of the ECHO-G coupled Atmosphere-Ocean
General Circulation Model (AOGCM) to explore the evolution of thermal-orbit slopes
over time. Collectively, our study establishes the theoretical foundation of thermal or-
bit regression analyses and demonstrates their utility for future model and observational
studies.

2. Theory

2.1. Background

With respect to studies of air and ground temperature relationships, it has been useful to
approximate annual air temperature signals as harmonic oscillations with annual periods
[e.g. Carson 1963; Beltrami 1996; Smerdon et al. 2003, 2004, 2006; Demetrescu et al.
2007]. The conductive propagation of such an harmonic signal into the subsurface is
well characterized [e.g. Carslaw and Jaeger 1959; Carson 1963; Geiger et al. 1965; van
Wijk 1963]. An idealized model of thermal propagation in the subsurface assumes that
surface air temperatures (SATs) and ground surface temperatures (GSTs) are perfectly
coupled, yielding a surface temperature boundary condition that is equal to the SAT
signal. Furthermore, the subsurface is represented as a one-dimensional semi-infinite
half space that has either layered or homogeneous thermophysical properties with depth.
For the purposes herein, we refer to this collective scenario as the conductive model.
Deviations from this model therefore will be termed non-conductive perturbations that
are manifest as processes that either disrupt coupling between SAT and GST (e.g. snow
cover, vegetation, or evapotranspiration) or alter the conductive transport of heat in the subsurface (e.g. freeze-thaw cycles or advective moisture transport).

For the conductive model, Beltrami [1996] developed the theoretical phase-space equations assuming an harmonic approximation of the SAT function:

\[ T(0, t) = A \cos(\omega t + \epsilon) + T_0, \quad (1) \]

where \( t \) is time, \( \omega \) is angular frequency, \( T_0 \) is the mean temperature and \( A \) and \( \epsilon \) are the initial amplitude and phase, respectively. The steady-state analytic solution of the one-dimensional heat conduction equation, using Equation 1 as the surface temperature boundary condition, yields the downward propagating temperature signal in the subsurface as a function of time and depth (\( z \)):

\[ T(z, t) = Ae^{-kz} \cos(\omega t + \epsilon - kz) + T_0 \quad (2) \]

where the thermal wave vector \( k = (\frac{\pi}{P\kappa})^{1/2} \) is defined in terms of the harmonic period, \( P \), and the thermal diffusivity, \( \kappa \) (for an extended discussion of this solution see Carslaw and Jaeger, 1959). The solution yields a downward propagating thermal wave that is attenuated and phase shifted by amounts dependent on the period of the propagating wave and the thermal diffusivity of the subsurface (this behavior is also well documented in observational records, e.g. Geiger, 1965; Smerdon et al. 2003; Beltrami and Kellman, 2003; Demetrescu et al. 2007). It is these amplitude and phase shifts with depth that yield characteristic elliptical patterns in phase-space representations of air and ground temperatures.
2.2. Thermal Orbit Regression Analysis

The principal contribution of this study is the application and description of a linear regression analysis for characterizations of evolving land-surface processes. The general approach applies a least-squares regression to independent air temperature and dependent ground temperature variables, i.e. a thermal-orbit representation of the two temperature time series. Our goal, therefore, is to understand the evolution of thermal orbit regression lines within a conductive regime and how changes in land-surface processes will change the nature of the determined regression.

Consider the parameterized versions of Equations 1 and 2:

\[ x(t) = \cos(\omega t + \epsilon), \]  \hspace{1cm} (3)

and

\[ y(t) = e^{-kz} \cos(\omega t + \epsilon - kz), \]  \hspace{1cm} (4)

where the parameter \( t \) is time, \( x(t) \) is the SAT function with unit amplitude, and \( y(t) \) is the signal at depth that has been attenuated and phase shifted proportional to the wave vector \( k \) (we assume a surface function of unit amplitude and a mean of zero). When \( y(t) \) is plotted versus \( x(t) \) the parameterized equations trace an ellipse. A linear regression on \( x \) defines a line that passes through the two points on the ellipse with vertical tangent lines (This is discussed within statistics texts under various subjects. See, for example, Wonnacott and Wonnacott (1972) on normalizing bivariate populations.). Using the calculus of parameterized curves, the first derivative of the function \( y(t) \) with respect to \( x(t) \) is
\[
\frac{dy}{dt} = \frac{dx}{dt} = e^{-kz}[\cos(kz) - \cot(\omega t + \epsilon) \sin(kz)]. \tag{5}
\]

Equation 5 therefore defines the slope of all lines tangent to the ellipse; vertical tangents occur where the slope becomes infinite. At a specific depth \( z \), Equation 5 diverges at two times (both denoted below as \( t_r \)) within the principal argument of the cotangent function:

\[
t_r = -\frac{\epsilon}{\omega} \quad \text{or} \quad \frac{\pi - \epsilon}{\omega}. \tag{6}
\]

If each harmonic function has a mean of zero, the regression line will also pass through the origin. The slope of any thermal orbit regression line therefore will be defined by the attenuation and phase of the signal at depth in the following form:

\[
\frac{y(t_r)}{x(t_r)} = \frac{e^{-kz} \cos(\omega t_r + \epsilon - kz)}{\cos(\omega t_r + \epsilon)}. \tag{7}
\]

To demonstrate the evolution of thermal orbits as a function of depth, Figure 1 plots Equation 2 versus Equation 1 at select depths down to 10 m assuming an harmonic signal with annual period and unit amplitude. An estimated thermal diffusivity of \( 3.1 \times 10^{-7} m^2 s^{-1} \) is used as representative of the subsurface at Fargo, North Dakota [Smerdon et al. 2003] where observational records of air and ground temperatures have been collected for more than two decades and extensively analyzed (e.g. Schmidt et al. 2001; Smerdon et al. 2003; 2004; 2006). In Figure 2 we represent the idealized conductive scenario by plotting the theoretical slopes of the thermal orbits as a function of depth, using Equation 7 and the estimated thermal diffusivity from Fargo. As Figures 1 and 2 demonstrate, the slopes of the thermal orbit regression line begin close to 1, but as the amplitude of the
ground temperature signal is attenuated and phase shifted, the orbit is rotated clockwise and the slope is reduced. Phase shifts make the slope negative for a range of depths in the subsurface, before the slope converges to zero as the annual ground temperature signal is almost completely attenuated by depths of about 10 m. Figure 2 also includes thermal orbit regression slopes determined from the observational data at Fargo using the near-surface ground temperature (0.01 m) as the surface temperature boundary condition to approximate the conductive model. Slopes were determined from daily data and from the annual signals extracted from Fourier decompositions of the Fargo time series [Smerdon et al. 2003]; the two observational results were determined over the period from 1 September 1980 to 31 August 1999. Observed and modeled behavior is clearly consistent, with only small differences resulting from near-surface diffusivity variations at Fargo that are not captured in the homogeneous model adopted for the theoretical calculation [Smerdon et al. 2003].

Given a fixed period of the propagating harmonic signal, the thermal diffusivity of the subsurface dictates the behavior with depth of the orbit slopes in the conductive model. To demonstrate this dependence, Figure 2 also plots thermal-orbit slopes as a function of depth given thermal diffusivities that are ± 50% of the estimated value at Fargo. A reduced diffusivity yields slopes that decline more rapidly with depth and thus converge to zero at a shallower depth within the subsurface. The converse is true for increased diffusivities.

2.3. Perturbation of Thermal Orbits

Following the regression analysis presented in the preceding subsection, we explore here the effect of changes in ground temperature means, amplitudes and phases representative...
of perturbations to the conductive model. We begin with air and ground temperature functions as presented in Equations 1 and 2, assuming the air temperature has a mean of zero, unit amplitude and an initial phase of zero. In the subsequent discussion, we consider the ground temperature at 1 m and a subsurface with a thermal diffusivity of $1 \times 10^{-6} m^2 s^{-1}$.

Figure 3 explores thermal orbit transformations given changes in the annual mean of either the air or ground temperature signal or both. For changes in the air temperature mean, the center of the thermal orbit is shifted horizontally in phase space, while changes in mean ground temperature move the orbit vertically. When changes occur in both air and ground temperature means, the center of the orbit moves diagonally in phase space; when mean changes in air and ground temperatures are identical the orbit center will move exactly along the 1-to-1 line. Most importantly, Figure 3 demonstrates that mean changes in the air and ground temperatures move the location of the center of the orbit, but they do not change the slope of the regression line.

Figure 4 illustrates changes in thermal orbits due to amplitude reductions at 1 m. The air temperature signal is held constant for all orbits shown in Figure 4, whereas the amplitude of the 1-m ground temperature signal is reduced by 75, 50 and 25% relative to the original amplitude of the conductively propagated signal at 1 m. The principal change linked to reduced amplitudes is a rotation away from the original orbit, yielding a reduction in the slope of the orbit regression line that depends on the amplitude change. The demonstration in Figure 4 can be generalized to the following rule for the effect of non-conductive amplitude changes on thermal orbits: increases in amplitude at a given
subsurface depth will cause increases in the measured thermal orbit regression slope, whereas reductions in the amplitude will cause the regression slope to decrease.

Figure 5 explores the impact of phase changes on thermal orbits at 1 m. The air temperature signal is again held constant for all orbits shown in Figure 5, while the phase of the 1-m ground temperature signal is shifted by $\pi/8$, $\pi/6$ and $\pi/4$ radians. Similar to the amplitude reductions, the thermal orbit regression slopes are reduced as phase shifts are increased. These shifts are principally due to an increase in the length of the minor axis of the thermal orbits and a progressive shift in the location of the points on the ellipse where the vertical tangents exist (the points of intersection for the regression lines; see section 2.2). Based on the results of Figure 5, the effect of phase shifts on thermal orbit regression lines can be generalized as: positive phase shifts at depth are associated with a reduction in the regression slope, whereas negative phase shifts are associated with increases in the regression slope.

Variations in multiple conditions and processes at the land-surface can give rise to the amplitude and phase shift changes described above. Snow cover insulates the ground from cold winter temperatures and effectively reduces the amplitude of the annual GST signal. Secular changes in the number of snow-cover days, snow cover thickness or the timing of snow cover arrival and melt can all impact the degree to which annual GST amplitudes are reduced relative to the SAT [e.g. Bartlett et al. 2004]. Freeze-thaw cycles can also impact both the amplitude of the propagating subsurface signal and its phase. Temperature amplitudes are muted by the latent heat of fusion as soil layers are held constant at 0 °C while soil moisture freezes or thaws [e.g. Kane et al. 2001]. The freezing process also delays the timing of minimum temperatures in soil layers and therefore impacts the
phase of the downward propagating signal. Soil moisture, precipitation and vegetation also impact peak subsurface temperatures in the summer. Vegetation or land-use changes alter the radiation balance at the land-surface and can either increase or decrease maximum ground temperatures relative to air temperatures. Moisture availability in the subsurface also impacts the rates of evapotranspiration at the land surface and cools the subsurface relative to the SAT [Smerdon et al. 2006]. Interannual changes in all of these processes and conditions can therefore alter the character of the SAT-GST relationship and are thus manifest in thermal orbit regression slopes.

3. Modeling Experiments

Given the above theoretical discussion, the remainder of the manuscript uses modeled data to demonstrate the application of the thermal orbit regression analysis. In the following subsections, we describe the two models employed and the results of the experiments that we perform.

3.1. One-Dimensional Land-Surface Model Experiments

We use a one-dimensional finite-difference land surface model (LSM) in several experiments hereinafter. The LSM was originally designed to study snow-ground thermal interactions and the thermal regime of the terrestrial subsurface [Goodrich, 1982; Stevens et al., 2007]. The simulations in this study include the effects of snow cover and latent heat exchanges associated with freezing and thawing in the subsurface; no hydrological effects are considered within the model including thermophysical changes associated with moisture variability or surface energy fluxes associated with evapotranspiration. Vegetation effects also are not included in the LSM simulations.
Both the thermophysical properties of the simulated subsurface and the upper and lower boundaries of the model are prescribed. The lower boundary condition can be placed at a specific depth and constrained by either constant temperature or heat flux. For this study we use a zero-flux lower boundary condition at 100-m depth. A thermal diffusivity of $1 \times 10^{-6} \text{m}^2\text{s}^{-1}$ was used for all simulations. All runs were done with quarter-daily timesteps and 0.1 m node spacing.

We use the 1996-97 year (Sept.-Aug.) of observed SAT data from the Pomquet observatory in Nova Scotia, Canada [Beltrami, 2001] to drive a baseline simulation with the LSM. This single year of data was repeated for 100 years as spin-up to remove the effects of initial conditions and the last year of the simulation is used as the baseline year. Our first experiment was to increase the mean of the Pomquet SAT data by 17 °C, relative to baseline. This has the effect of emulating a more southerly site in which all daily SAT values are above zero, resulting in no snow cover or subsurface freezing.

Figure 6 plots daily and monthly baseline and mean-shifted ground temperatures derived from the LSM. The illustration shows the impact of snow cover and freezing on the propagating ground temperature signal given an SAT signal that is identical except for its annual mean. Given the absence of these cryogenic processes in the mean-shifted simulation, the SAT is transmitted dominantly by conduction. By contrast, the baseline simulation is attenuated and subject to cryogenic effects at 0 °C where subsurface temperatures are held constant during freezing and thawing; the minimum ground temperatures are also shifted later in the year. These effects are more clearly evident in the 2-m ground temperatures, where many of the high-frequency fluctuations (periods shorter than annual) have been damped out.
Based on the theoretical discussions in Section 2.3 and the demonstrated differences in the ground temperatures shown in Figure 6, the expected differences between the thermal orbits derived for the baseline and mean-shifted simulations would be a diagonal translation of the center of the shifted orbit and an increase in its regression slope relative to baseline. In Figure 7 we plot the thermal orbits for the two simulations. The mean-shifted orbit has clearly been translated in phase space reflecting the mean increases in both the air and ground temperatures. More importantly, the absence of cryogenic effects in the mean-shifted simulation causes an increase in the regression orbit slope from 0.31 ± 0.02 to 0.36 ± 0.02 at the observed 2-m depth. These slopes have been determined for thermal orbits generated from mean daily temperatures and the uncertainties reflect one standard error in the regression coefficients.

To complement the results in Figure 7, which represent locational differences, we have also used the LSM to calculate orbit changes representative of a specific location. Again beginning with the baseline simulation we impose ±1 °C per century trends in the annual mean of the SAT data. Figure 8 plots the orbits and calculated regression lines for the baseline and the hundredth year of the two trend experiments (although the regression lines have been determined using daily data, only the monthly data are plotted for visual clarity). Once again we note the diagonal translation in the center of the orbits and an increase or decrease in the slope of the regression lines; the orbit slopes of the positive and negative temperature change experiments are 0.34 ± 0.02 and 0.29 ± 0.02, respectively, compared to the original orbit slope of 0.31 ± 0.02. These results are consistent with the expectation that more (less) cryogenic clipping occurs for a temperature decrease (increase), relative to the baseline run.
All of the above changes in regression slopes are consistent with physical expectations. Differences between the baseline and mean-shifted experiments are significant within one standard error of the regression coefficients. The subsequent trend experiments, however, yield estimated regression slopes with overlapping uncertainty ranges. While these latter experiments are indeed consistent with expectation, their measured slopes cannot said to be statistically different from baseline. The implication of these results is that modest changes in SAT, and thus the accompanying cryogenic changes modeled above, may require multiple years of data to derive statistically significant changes in regression slopes. Monthly instead of daily data may also more effectively estimate the thermal orbit regression slopes. Unless used over multiple years, however, monthly aggregations can have the effect of increasing the standard error in a single-year estimate because fewer degrees of freedom are involved. Daily data may also be filtered to extract annual signals [Smerdon et al. 2003], which can then be used to estimate thermal-orbit slopes. Results from such an approach are provided in Figure 2.

It is also important to note that the standard error in the regression coefficient may not be an effective evaluation of the uncertainty in the thermal orbit regression slope because it reflects the adherence of the data to a linear model. By contrast, our approach recognizes that the linear model does not describe the underlying ellipses that we seek to characterize and our purpose is only to identify the points on the ellipse where vertical tangents occur. The problem of identifying these two points is therefore more constrained and standard error estimates likely overestimate the degree of uncertainty in the thermal orbit regression slopes. Consider, for instance, the theoretical thermal orbit presented in Figure 4. The slope of the original orbit (shown in black) was calculated to be 0.69. This
value can be considered the true slope of the orbit because the data define a perfect ellipse.
In the context of least-squares regression, however, the points of the ellipse fall away from
the linear regression line and give rise to residuals that define a standard error in the slope
of ± 0.01. While this standard error is modestly smaller than the uncertainty reported in
the slopes determined for the model data above, the value suggests that standard errors
will overestimate the uncertainties in the regression slopes that we seek to define.

3.2. ECHO-G Experiments

In the remaining experiments we employ simulations from the ECHO-G AOGCM, which
consists of the atmospheric component ECHAM4 and the ocean component HOPE-G
[Legutke and Voss, 1999]. The ECHO-G soil model (ECHAM4, see Roeckner et al. 1996)
has five vertical layers that increase in thickness with depth to 9.834 m. The model uses
a uniform thermal diffusivity of $\kappa = 7.5 \times 10^{-7} m^2 s^{-1}$ and heat capacity per unit volume
of $\rho g C_g = 2.4 \times 10^6 J m^{-3} K^{-1}$. Thermophysical subsurface properties are isotropic in
time and space. Energy is exchanged at the air-ground interface in the form of long-
and short-wave infrared sensible heat fluxes, and latent heat due to moisture fluxes and
cryogenic processes. Water is input to the soil model via precipitation, and is output to
the atmosphere via transpiration, evaporation, and advection. Snow cover, accumulation
and melting are bundled into the soil model, as well as the resulting changes in albedo and
surface roughness parameters. Vegetation cover is not temporally dynamic, but influences
evapotranspiration exchange between plants in the soil model and the atmosphere.

There exist two 1000-year forced simulations for the ECHO-G AOGCM and a control
simulation with constant present-day forcing. The two millennial forced runs were pro-
duced with different initial conditions but the same forcing based on estimates of solar
variability, greenhouse gas concentrations and stratospheric volcanic aerosols (FOR1 and FOR2, as labeled by González-Rouco et al. [2009]). The FOR1 simulation is used for subsequent experiments herein. See González-Rouco et al. [2009, and references therein] for additional details and model verification of the ECHO-G model.

We select simulated temperatures from two grid nodes of the FOR1 simulation (grid spacings in the ECHO-g integrations are defined by a T30 Gaussian grid (approximately a 3.75° latitude-longitude grid)): one in the Northern Hemisphere (50.1°N, 123.75°W) where air temperatures go below freezing and snow cover is present, and one in the Southern Hemisphere (24.1°S, 30.0°E) where air temperatures do not go below freezing and all precipitation falls as rain. We plot in Figure 9 the thermal orbits for the first ten years of the FOR1 simulation for the two grid-cell locations; the plots use the ECHO-G soil layer with a mean depth of 0.78 m. Similar to the results shown in Figure 7, we note a difference in the location of the center of each thermal orbit and differences in the thermal orbit regression slopes consistent with expectations: 0.73 ± 0.02 and 0.68 ± 0.02 for the southern and northern locations, respectively. These sites have the same subsurface diffusivity, tying the differences in the thermal orbits exclusively to the differences in the meteorological conditions and land-surface processes active at the two locations. Consistent with the uncertainty discussion in Section 3.1, the regression lines for the thermal orbits have been calculated using the monthly means from ten years of data as a means of improving the standard error in the regression coefficients.

Perhaps the most useful application of the thermal-orbit regression method is its ability to track changes in land-surface processes over time. We demonstrate this by tracking the thermal-orbit regression slopes over the thousand years of simulation at the two ECHO-G
locations. Regression slopes are determined using the monthly values of air and ground temperatures in ten-year intervals. The standard error of the regression coefficients are also calculated for each decade. The SAT time series and calculated thermal-orbit regression slopes and uncertainties are shown in Figure 10 for the southern and northern locations. Over the course of the simulation, the evolution of the slopes in the southern grid show no appreciable changes, whereas secular changes in the slopes from the northern grid are clearly evident; both locations have secular SAT changes (Figure 10). The calculated uncertainties also indicate that the observed behaviors at the two locations are statistically distinct and that the secular behavior observed for the northern location exceeds the estimated statistical uncertainty. Most interestingly, there is no significant correlation between the decadally averaged regression slopes and SAT at the southern location ($r = -0.096$), whereas these two variables are strongly correlated at the northern location ($r = 0.803$). This result underscores the fact that mean SAT changes do not change the regression slope unless they are tied to additional changes in land-surface processes. In the specific comparison considered herein for the ECHO-G data, this fact is reflected in the northern location where changes in snow cover and freeze-thaw cycles occur over the duration of the simulation caused by temperature decreases into the 16th and 17th centuries before warming toward present day. At the southern location, the model simulates relative temperature changes similar to the northern location, but the mean annual SAT does not go below freezing. The thermal-orbit regression slope therefore is insensitive to the mean SAT changes (e.g. Figure 3), and only responds to the changes in land-surface processes occurring at the northern site.
Note that both the LSM and ECHO-G model involve the placement of a zero-flux lower boundary condition. The depth of the lower boundary can affect the annual signal [e.g. Sun and Zhang 2004; Smerdon and Stieglitz 2006; Stevens et al. 2007; Alexeev et al. 2007; Nicolsky et al. 2007], which is of greatest interest in the thermal orbit application. The lower boundary of the LSM was set to 100 m, well below a depth that will impact the annual signal. The lower boundary of the ECHO-G land-surface scheme is considerably shallower, however, and located at approximately at 10 m. While this boundary-condition depth likely has small effects on the annual signal, it is deep enough to minimize effects compared to most AOGCMs that set the boundary between about 3 and 4 m. Nevertheless, effects due to the lower boundary in the soil temperatures simulated by ECHO-G are likely present [González-Rouco et al. 2009] and could affect the regression slopes determined above. Given that the effect of the lower boundary is to artificially enhance the amplitudes of propagating signals, however, the errors would likely mute the magnitude of slope changes in the regression analysis. The ECHO-G experiments are therefore conservative relative to those measured in simulations in which the lower-boundary effects are not present.

4. Discussion

The differences in calculated slopes presented herein are predominantly larger than the range of reported standard errors, with the exception of the LSM experiments that explored ± 1 °C changes in mean SAT. These latter results were determined for one year of data and indicate that data aggregation may be important for measuring changes in land-surface conditions that are small in magnitude. We have demonstrated one such approach using monthly data over ten-year intervals in the ECHO-G experiments. Each
application of the thermal-orbit method will, however, require an individual assessment of uncertainty and observational data will also infuse additional errors. Future applications of thermal-orbit analyses should therefore be accompanied by proper evaluations of statistical uncertainties, with the caveat that alternative uncertainty estimates may be more appropriate than the standard error calculations presented herein.

Several additional points should also be mentioned. The first regards the subsurface depth that should be used for thermal orbit calculations. As shown in Figure 6, near-surface daily variability is present in the first several tens of centimeters. This variability will obscure the characterization of an annual thermal orbit. Furthermore, in cases where freezing processes extend below the surface, it is suggested that ground temperatures be used below the maximum depth of freezing (or below the active layer in permafrost regions). Not only will this allow a more effective characterization of the annual thermal orbit, the signal below this depth will incorporate all changes in cryogenic processes within the subsurface and therefore more completely characterize changes in land-surface processes. It also should be noted that the character of the thermal-orbit regression slopes will vary with subsurface thermal properties. Contrary to the ECHO-G data, for instance, in which the model diffusivity is spatially invariant, thermal properties need to be accounted for in studies that compare results for locations or models that use different thermal diffusivities. The ultimate consequence is that a measured slope at one depth and diffusivity should not be expected to yield the same result for the same depth at a site with different thermal properties. One means of overcoming this disparity would be to report a ratio between a theoretical slope of a conductively propagated SAT signal and the measured slope that includes the additional land-surface processes at the location. Alternatively,
results could be reported for equivalent thermal depths based on the thermal properties of
the given locations [e.g. Beltrami et al. 1997]. Note that for observational data, the effective
thermal diffusivities of the subsurface can be estimated in a straightforward fashion [e.g. Smerdon et al. 2003].

With the above caveats in mind, the principal benefit of the thermal-orbit regression
analysis is its specific sensitivity to changes in land-surface processes. We have shown this
convincingly for cryogenic processes, namely the effects of snow cover and soil freezing.

The analysis would work similarly for vegetation changes that damp the amplitudes of
propagating SAT signals in the subsurface in a manner similar to cryogenic processes [e.g. 
Smerdon et al. 2004; 2006]. The ability to characterize these changes in a single analysis
has multiple benefits. In modeling applications where land-surface processes are parameterized or calculated in different ways, thermal-orbit analyses can be used as an assessment of the character of the coupling across multiple modeling schemes. As demonstrated with the ECHO-G data, it is also straightforward to track regression slopes in time, making them a useful diagnostic tool for comparing how multiple models simulate the evolution of land-surface processes. The analysis will also allow model validation in areas where observational data are available. Given that the thermal-orbit analysis includes the effect of multiple land-surface processes, it is a comprehensive measure of a model’s ability to simulate relevant land-surface processes. With regard to spectral analysis of air and ground temperatures, the method is also dependent on both the phase and the amplitude of the propagating signal. This joint dependence may yield more robust evaluations of air and ground temperature relationships than methods that track just amplitude or phase behavior independently with depth [e.g. Smerdon et al. 2003, 2006; Demetrescu et al.
2007]. More generally, the thermal-orbit regression slope may be a simple measure of the strength of coupling between the land and atmosphere, making it useful for investigations into the dynamic role that the land surface plays in the climate system (a similar analysis for soil moisture is also possible). A measure of coupling strength would also be useful for land-surface modeling efforts that have attempted to provide parsimonious descriptions of land-surface processes and subsurface thermodynamics [e.g. Stieglitz and Smerdon, 2007].

5. Conclusions

The purpose of this study has been to develop a quantitative means of describing changes in air-ground thermal orbits, and therefore a measure of changes in land-surface processes. Thermal orbits have been useful for describing qualitative relationships between air and ground temperatures since they were originally presented by Beltrami [1996], but the quantitative analysis presented herein greatly enhances the applicability of the thermal orbit representation. We have used theoretical descriptions and controlled model experiments to demonstrate this applicability. A key to understanding the utility of the thermal orbit regression analysis is the fact that it is insensitive to changes in SAT unless they are accompanied by related changes in land-surface processes. This is a critical aspect of the thermal orbit analysis that makes it a comprehensive evaluation of changes in processes occurring at or below the land surface that couple air and ground temperatures. This quality is the foundation of the method’s utility. While some caveats associated with the method still require additional investigation, the success demonstrated in the present study suggests that the analysis has the potential for wide applicability.
Acknowledgments. This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Atlantic Innovation Fund (AIF) and The Canadian Foundation of Climate and Atmospheric Sciences (CFCAS). MBS is supported by an NSERC PGS-D postgraduate scholarship (Canada). Part of this work was carried out while JES was a visiting James Chair Professor at STFX. Special thanks to J. Fidel González-Rouco for providing the ECHO-G data used in this study. We are also grateful to JFGR and Henry N. Pollack for their insightful conversations and comments on the content of this manuscript. LDEO contribution no. XXXX.

References


Foster, J. L., J. W. Winchester, and E. G. Dutton (1992), The date of snow disappearance on the Arctic tundra as determined from satellite, meteorological station and radiomet-


Walsh, J. E. et al. (2005), Cryosphere and Hydrology, in *Arctic Climate Impacts Assessment*, Cambridge University Press, New York, NY.


Figure 1: Theoretical thermal orbits at various depths using a thermal diffusivity of $3.1 \times 10^{-7} m^2 s^{-1}$, as estimated from the observational data at Fargo, North Dakota. The air temperature signal is assumed to have an amplitude of one and a mean of zero. The solid black line in the figure is the one-to-one line.

Figure 2: Theoretical and observed thermal-orbit slopes at Fargo, North Dakota. Theoretical slopes are calculated using a thermal diffusivity of $3.1 \times 10^{-7} m^2 s^{-1}$, as determined from the observational data at Fargo. Measured slopes are determined from the mean daily data and from the mean annual signal extracted from the Fargo time series using a Fourier spectral decomposition [Smerdon et al., 2003]. The two dotted lines in the figure correspond to theoretical slopes using diffusivities that are 50% larger or smaller than the value of the Fargo diffusivity.

Figure 3: Changes in thermal orbits due to mean changes in air and ground temperatures. Theoretical thermal orbits are generated using an harmonic air temperature signal of unit amplitude and the subsequent 1-m ground temperatures in a semi-infinite half space with a diffusivity of $1 \times 10^{-6} m^2 s^{-1}$. Red, green and blue orbits represent respective changes in air, ground and both air and ground temperatures by 0.5 and 1.0 K. The black orbit is the original thermal orbit in which the air and resulting ground temperatures have a mean of zero.
Figure 4: Change in thermal orbits due to a change in ground temperature amplitudes. The black orbit is the original thermal orbit assuming a surface air temperature of unit amplitude and a mean of zero. Blue orbits represent results for 1-m ground temperature signals with amplitudes that are 75, 50 and 25% of the original 1-m amplitude. No phase changes in the ground temperature signal have been applied except for those dictated by the diffusive thermal transport in the subsurface, i.e. the phase of the ground temperature signal is the same in all orbits shown.

Figure 5: Change in thermal orbits due to a change in phase of the ground temperature signal. The black orbit is the original thermal orbit assuming a surface air temperature of unit amplitude and a mean of zero. Blue orbits represent results for 1-m ground temperature signals that have phases increased by $\pi/8$, $\pi/4$ and $\pi/2$ radians. No amplitude changes in the ground temperature signal have been applied except for those dictated by conductive thermal transport in the subsurface, i.e. the amplitude of the ground temperature signal is the same in all orbits shown.

Figure 6: Simulated ground temperatures using the LSM and the single observational year (1996-97) from the Pomquet station. The dark- and light-blue lines correspond to the baseline LSM simulation in which the original Pomquet surface air temperatures were used. The orange and red lines represent the LSM simulation that uses the Pomquet surface air temperatures with a mean shift of 17 °C. The plots show both the mean daily and monthly temperatures. Right and left vertical axes have been shifted to yield summer temperatures that overlay each other in both simulations.
Figure 7: Equilibrium thermal orbits computed from the air and ground temperatures shown in Figure 6. The original orbit was calculated using the actual Pomquet data and the shifted orbit was calculated using a 17 °C shift in the annual mean of the Pomquet data so that there were no daily air temperatures below 0 °C and no snow cover. This mean shift increases both air and ground temperature means and has translated the center of the mean-shifted orbit in phase space as shown. Linear regression lines (shown in red) are computed using daily mean data, results of which are provided in the figure.

Figure 8: Equilibrium thermal orbits computed from a repeated year of the Pomquet observational data compared to thermal orbits computed in the hundredth year of increasing and decreasing trends (±1°C per century) in the mean annual air temperature. These mean shifts increase or decrease both air and ground temperatures and translate the center of the orbits in phase space as shown. Only monthly data are shown for ease of viewing, but linear regression lines (shown in red) are computed using daily mean data, results of which are provided in the figure.

Figure 9: Example thermal orbits calculated for the first ten years of the ECHO-G FOR1 millennial simulation. Northern and southern grids are centered at 50.1°N, 123.75°W and 24.1°S, 30.0°E, respectively. Linear regression lines (shown in red) are computed using ten years of mean monthly data, results of which are provided in the figure.
Figure 10: Decadal mean time series for (a) the air temperatures selected from the northern and southern locations in the ECHO-G simulation; and (b) the thermal orbit slopes computed for each of the selected locations. The southern grid shows no appreciable change in the mean orbit slope over the duration of the simulation and has no significant correlation with the mean annual air temperature. The thermal orbit slopes for the northern grid display considerably more variability, centennial trends and have a significant correlation with the surface air temperature. Shaded regions in the bottom panel correspond to one standard error in the linear regression coefficients determined for the thermal orbits.
17 °C Shift in Mean Air Temperatures

Original Temperatures

Linear Regression Results
Original Orbit: \( y = 0.31x + 4.6 \)
Mean Shift in Air Temp. (17 °C): \( y = 0.36x + 14 \)
Linear Regression Results

Original Orbit: \[ y = 0.31x + 4.6 \]
1 °C Temperature Increase: \[ y = 0.34x + 4.9 \]
1 °C Temperature Decrease: \[ y = 0.29x + 4.2 \]
Southern Grid:
\[ y = 0.73x + 5.1 \]
Northern Grid:
\[ y = 0.68x + 1.6 \]