A surface model for water and energy balance in cold regions accounting for vapor diffusion

Enkhbayar Dandar\textsuperscript{1,2,4}, Maarten W. Saaltink\textsuperscript{2,3}, Jesús Carrera\textsuperscript{1,3}, Buyankhishig Nemer\textsuperscript{4}

\textsuperscript{1}Institute of Environmental Assessment and Water Research (IDAEA), CSIC, c/ Jordi Girona 18-26, Barcelona, 08034, Spain
\textsuperscript{2}GHS, Department of Civil and Environmental Engineering, Universitat Politecnica de Catalunya, UPC, Jordi Girona 1-3, Barcelona, 08034, Spain
\textsuperscript{3}Associated Unit: Hydrogeology Group (UPC-CSIC)
\textsuperscript{4}Department of Geology and Hydrogeology, School of Geology and Mining Engineering, MUST, Baga Toiruu 46-114, Ulaanbaatar, 14191, Mongolia

Correspondence to: Enkhbayar Dandar (denkhbayar@gmail.com)

Abstract. Computation of recharge in subarctic climate regions is complicated by phase change and permafrost, causing conventional conceptual land surface models to be inaccurate. Actual evaporation tends to fall very low in the Budyko curve and surface runoff tends to be much larger than what would be expected in terms of potential evapotranspiration. We develop a two-compartments water and energy balance model that accounts for freezing and melting and includes vapor diffusion as a water and energy transfer mechanism. It also accounts for the effect of slope orientation on radiation, which may be important for mountain areas. We apply this model to weather data from the Terelj station (Mongolia). We find that direct surface runoff is small and concentrated at the beginning of spring due to snowmelt. Recharge is relatively high and delayed with respect to snowmelt because a portion of it is associated to thawing at depth, which may occur much later. Finally, but most importantly, we find that vapor diffusion plays an important quantitative role in the energy balance and a relevant qualitative role in the water balance. Except for a few large precipitation events, most of the continuous recharge is driven by vapor diffusion fluxes. Large vapor fluxes occur during spring and early summer, when surface temperatures are moderate, but the subsoil remains cold, creating large downwards vapor pressure gradients. Temperature gradients reverse in fall and early winter, but the vapor diffusion fluxes do not, because of the small vapor pressure differences at low temperature. The downwards latent heat flux associated to vapor diffusion is largely compensated by heat conduction, which is much larger than in temperate regions and upwards on average.
1. Introduction

This work is motivated by the assessment of water resources in the Upper Tuul River basin, around Ulaanbaatar (Mongolia) and, in general, by subarctic continental climate regions, characterized by very low temperatures, low rainfall and, yet, sizable runoff. This causes such regions to fall very low in the Budyko curve (see, e.g., Figure 3b of Hanasaki et al., 2007). That is, total runoff is much larger than what would be expected in terms of potential evapotranspiration and rainfall. We conjecture that increased runoff may be caused by condensation (deposition) of air moisture. Condensation and freezing may be especially significant during the spring when air temperature and moisture increase, which drives water vapor to the cold soil. This mechanism was proposed by Shvetzov (1978) and it is also mentioned by Gusev and Nanosova (2002), but has not been analyzed in detail. Since moisture condensation data are not available in such regions, and phase changes are driven by energy availability, water and energy balances are required. Water and energy balances are the basic building blocks for any hydrological model and specifically of land surface schemes (also called large area hydrological models) linking meteorological data with runoff, and thus water availability. Hydrological modeling in cold regions has been largely motivated by climate change (Nickolsky et al., 2007) because of the fear that permafrost melting may release large amounts of methane to the atmosphere, thus accelerating global warming (Anisimov, 2007). Efforts have also been devoted to the construction of land surface schemes to be coupled to atmospheric circulation models (Ek et al., 2003), which cover the whole Earth. However, these tend to be simple to facilitate coupling to atmospheric processes, so that they may not account for all processes taking place at the soil surface. Detailed models are possible, but complex (e.g., Sapriza-Azuri et al., 2015; Bao et al., 2016). Yet, proper understanding of the hydrological cycle and water and energy fluxes in boreal regions is important not only for climate change but also for water management (Dandar et al., 2016).

Classical hydrological model codes, such as WatBal (Yates, 1996) and SWAT (e.g., Arnold et al., 1998; Hülsmann et al., 2015) calculate the basic terms of a river basin water balance, such as stored soil water, evapotranspiration and surface runoff. In its original form, SWAT calculates snowmelt by means of a water balance of the snow cover and a so-called temperature-index method that estimates the snow cover temperature from previous ones and air temperature. Some later versions of SWAT use energy balances of the snow cover instead of the temperature-index method (e.g., Fuka et al., 2012). Both methods have been compared thoroughly (Debele et al., 2010; Verdhen et al., 2014). Only recently SWAT has been extended by considering energy balances in the soil as well (Qi et al., 2016). Unfortunately, this only considers conduction as a mechanism for heat transport in the soil. For our purposes we need to take into account vapor diffusion flux in both the water and the energy balance of the soil.

The aim of this work is, first, to understand the hydrological processes and to evaluate water resources in cold and semi-arid regions and, second, to assess the importance of vapor diffusion. We do this by developing a hydrological scheme that accounts for the processes that are relevant for subarctic climates, including the effect of slope on radiation, which should be relevant for high latitudes, water phase changes and soil vapor diffusion.
2. Methodology

2.1. Model description

Water and energy balances in land surface hydrological models are typically expressed on one layer that extends to root depth, where plants can extract water. However, since we are interested in water and energy dynamics including seasonal fluctuations, we formulate the balances over two layers (see Figure 1). The surface layer extends some 16 cm, so as to accommodate the roots of typical grass in the Tuul basin and to dampen daily temperature fluctuations. The subsoil (with length $L_{ss}$) layer accommodates the "active" layer that freezes and thaws seasonally. The input and output terms for the water balance of the surface layer include precipitation (as rain or snow), evapotranspiration (including both ice deposition and sublimation), infiltration into the subsoil and vapor diffusion to or into the subsoil. Those for the subsoil are infiltration, vapor diffusion from the surface and recharge to the aquifer. The energy balance considers solar radiation, latent and sensible heat fluxes, heat conduction between the two layers and energy released due to phase changes. The model also takes into account the slope of the surface. The above fluxes can be written as a function of meteorological data and two state variables: mass of water ($kg/m^2$) and energy ($J/m^2$). Details of each mass balance term are given below.

![Figure 1. Schematic diagram of the water balance (a) and energy balance (b) models](image-url)

2.2. Water balance

The water balance for the surface and the subsoil layers can be formulated as:

$$\frac{\partial m_{sf}}{\partial t} = P - ET_{sf} - I - SR - J_D$$  (1)
\[ \frac{\partial m_{sf}}{\partial t} = I - ET_{sf} - R + J_D \]  

(2)

where subscripts “sf” and “ss” refer to surface and subsoil layers, respectively, \( m \) is the mass of water (liquid or ice) (kg m\(^{-2}\)), and fluxes (all in kg m\(^{-2}\) s\(^{-1}\)) include precipitation (P), evapotranspiration (ET), infiltration (I, positive downwards), surface runoff (SR), vapor diffusion (JD, positive downwards), and recharge (R).

Evapotranspiration is commonly used to describe both evaporation and transpiration (Brutsaert, 1982), including sublimation, the conversion of solid water directly to vapor (Zhang et al., 2004). ET is usually the most important term for returning energy to the atmosphere. Traditional methods for estimating evapotranspiration can be divided be into those that use temperature (Hargreaves and Samani, 1985), radiation (Priestley and Taylor, 1972) and an aerodynamic approach (McMahon et al., 2013 and Katul et al., 1992). We chose the latter, because in dry climates ET is controlled by water availability rather than incoming radiation. The aerodynamic method is based on Dalton law, which establishes that ET is proportional to the difference of vapor pressures between air and soil. It is the basis of the method of Penman (1948) and has been used to estimate evaporation from water surfaces (Rosenberg et al., 1983; Xu and Singh, 2002), or bare soil (Ripple et al., 1970) and evapotranspiration from vegetated surfaces (Blad and Rosenberg, 1976). We assume that one part of the water evaporates from the soil surface controlled by an aerodynamic resistance and another part evaporates through plant transpiration controlled by both aerodynamic and stomata resistance. Moreover, we distinguish between plant transpiration from the soil surface and the subsoil. We express ET from soil surface and subsoil as:

\[ ET_{sf} = \frac{M}{RT_{\text{air}}} \left( \frac{1}{r_s + r_s} - \frac{1}{r_s} \right) \left[ f_a \left( p_{e,sf} - p_{e,air} \right) \right] \]  

(3)

\[ ET_{ss} = \frac{M}{RT_{\text{air}}} \left( \frac{1}{r_s + r_s} - \frac{1}{r_s} \right) \left[ (1 - \alpha) \beta \left( p_{e,ss} - p_{e,air} \right) \right] \]  

(4)

where M is the molar mass of water (0.018 kg mol\(^{-1}\)), R is the gas constant (8.314 J mol\(^{-1}\) K\(^{-1}\)), \( T_{\text{air}} \) is the air temperature (K), \( r_s \) and \( r_s \) are the aerodynamic and stomata resistances, respectively (s m\(^{-1}\)), \( \alpha \) is the fraction of total transpiration from the surface layer and \( \beta \) is the fraction of vegetation cover. When the soil is frozen, no transpirations is assumed and \( \alpha \) and \( \beta \) are given a zero value. Factor, \( f_a \), represents the reduction of evaporation due to the lack of water in the surface and subsoil layers (see below), \( p_{e,sf} \) and \( p_{e,ss} \) are the saturated vapor pressures (Pa) at the surface and subsoil layers and \( p_{e,air} \) is the actual vapor pressure in the atmosphere (Pa). The actual vapor pressure can be calculated from the relative humidity and saturated vapor pressure (\( p_{e,air} \)), which we computed with Murray’s (1967) equation as a function of temperature.

The aerodynamic resistance, \( r_s \), describes the resistance from the vegetation upward and involves friction from air flowing over vegetative surfaces whereas the stomata surface resistance, \( r_s \), describes the resistance of vapor flow through stomata, total leaf area and evaporating soil surface (Shuttleworth, 1979). A general form for the aerodynamic resistance to evapotranspiration (or sublimation) and sensible heat is (Evett et al., 2011):
where \( z \) is height at which wind speed, temperature and relative humidity are measured (m), \( z_0 \) is the roughness length (m), \( k \) is von Karman’s constant (\( k = 0.4 \)) and \( u_z \) is the wind speed (m s\(^{-1}\)). The roughness length can vary over five orders of magnitude (from \( 10^{-5} \) m for very smooth water surfaces to several meters for forests and urban areas) and increases gradually with increasing height of roughness elements (Arya, 2001). The stomata surface resistance can be calculated by (Allen et al., 1998):

\[
\frac{r_s}{r_p} = 0.5 \text{ LAI}
\]

where \( r_p \) is the bulk stomata resistance of a well-illuminated leaf (s m\(^{-1}\)). Monteith and Unsworth (1990) suggest that \( r_p = 100 \) s m\(^{-1}\) for grassland. LAI is the leaf area index (leaf area per unit soil surface area).

Factor, \( f \), in equations (3) and (4) expresses the decreases of \( ET \) with water content in the surface and subsoil layers, according to:

\[
f_{ss/sf} = \begin{cases} 
1 & \text{if } m_{ss/sf} > m_{ss/sf}^{fc} \\
\frac{m_{ss/sf} - m_{ss/sf}^{wp}}{m_{ss/sf}^{fc} - m_{ss/sf}^{wp}} & \text{if } m_{ss/sf} < m_{ss/sf}^{fc} \\
0 & \text{otherwise}
\end{cases}
\]

where superscripts fc and wp refer to field capacity and wilting point, respectively.

For the calculation of infiltration (I) we assume that it is limited by a maximum infiltration capacity, which only liquid water exceeding the field capacity can infiltrate. Then, for a time step \( \Delta t \) the infiltration can be formulated as:

\[
I_{k+1}^{s/sf} = \max \left( 0, \min \left( \frac{m_{ss/sf}^{k+1} - m_{ss/sf}^{wp}}{\Delta t} + P - ET_{sf} - I_D, I_{max} m_{sf/sf} \right) \right)
\]

where \( m_{ss/sf}^{k+1} \) is the mass of water in the surface layer at a present time step and \( m_{sf/sf} \) is the liquid mass of water in the surface layer (kg m\(^{-2}\)). The maximum infiltration (\( I_{max} \)) equals the saturated hydraulic conductivity (\( K_{sat} \)).

In a similar way we calculate surface runoff (SR) by assuming that only water exceeding the maximum water content in the surface layer (\( m_{sf/sf}^\phi \)) can runoff, where \( m_{sf/sf}^\phi \) (kg m\(^{-2}\)) equals the porosity (\( \phi \)) multiplied by water density and length of the soil surface (\( L_{sf} \)). For the calculation of recharge (R) we assume that only liquid water exceeding the field capacity in the subsoil layer can percolate to the aquifer.

Vapor diffusion using Fick’s Law, is written as:

\[
I_D = \frac{M}{RT_{sf} L_{sf}} \frac{m_{sf}^\phi - m_{sf}}{m_{sf}^\phi} [P_{v sf} - P_{v ss}]
\]
where $D$ is the diffusion coefficient ($\text{m}^2\text{s}^{-1}$), $T_{sf}$ is the temperature of the surface layer (K), $L_{sf}$ is the surface layer's length (m). Note that we use $L_{sf}$ as length between the two layers rather than $(L_{sf} + L_{ss})/2$ because temperature gradients, which control vapor pressure, are expected to be largest near the soil surface. The value to be adopted for the diffusion coefficient deserves some discussion. The diffusion coefficient of water vapor in air is $0.3 \cdot 10^{-4} \text{m}^2\text{s}^{-1}$ (Cussler, 1997). This value should be reduced due to reduced open area and tortuosity in a porous medium. However, for reasons that are subject to debate (Ho and Webb, 2006), vapor diffusion is enhanced (Cass et al., 1984). Gran et al., (2011) required using values above $2 \cdot 10^{-4} \text{m}^2\text{s}^{-1}$. Given these uncertainties we have adopted $10^{-4} \text{m}^2\text{s}^{-1}$ as base value, and then analyzed the sensitivity of the model to this parameter.

2.3. Energy balance

Energy balance for the two layers is written as:

$$\frac{\partial U_{sf}}{\partial t} = e_{i,f}P - e_{g}ET_{sf} - (e_{g} - e_{i})ET_{ss} - e_{i}I - e_{i}SR + R_{n} - H - G - e_{g}J_{D}$$

(10)

$$\frac{\partial U_{ss}}{\partial t} = e_{i}I - e_{i}ET_{ss} - e_{i}R + G + e_{g}J_{D}$$

(11)

where $U$ is the total energy of each layer ($\text{J m}^{-2}$), $R_{n}$ is the net radiation ($\text{J m}^{-2}\text{s}^{-1}$), $H$ is the sensible heat flux ($\text{J m}^{-2}\text{s}^{-1}$) and $G$ is the heat conduction ($\text{J m}^{-2}\text{s}^{-1}$) between surface and subsoil. Advection heat fluxes include the fluxes of water, ice (snow) or vapor ($P$, $ET$, $SR$, $I$, $J$, $R$) multiplied by the corresponding internal energies of liquid water, ice or vapor ($e_{i}$, $e_{g}$, $e_{i}$, $J$ kg$^{-1}$), which are linear functions of temperature (minus the fusion latent heat for ice or the vaporization latent heat for vapor). Although a distinction is made between evapotranspiration from the surface and the subsoil ($ET_{sf}$ and $ET_{ss}$), equations (10) and (11) assume that the heat loss occurs only at the surface because the actual phase change takes place in the plant, which is part of the surface layer.

The total energy ($U$) is a sum of energy for liquid water, ice and solid (Figure 2):

$$U = me = m_{i}e_{i} + m_{j}e_{j} + m_{s}e_{s}$$

(12)

where subscript $s$ presents the solid part of the soil layers. During melting or freezing, the temperature is fixed at the melting temperature of 0°C while at higher temperature all water is liquid ($m = m_{i}$) and at lower temperature all water is ice ($m = m_{s}$). Therefore, we can calculate the temperature and mass of ice and water as a function of $U$ and $m$ as:

$$T = \frac{U + m\Lambda_{melt}}{mc_{i} + mc_{s}} \quad \text{and} \quad m = m_{i} \quad \text{if} \quad -m\Lambda_{melt} > U$$

(13)

$$T = 0 \quad \text{and} \quad m_{i} = -\frac{U}{\Lambda_{melt}} \quad m_{j} = m - m_{i} \quad \text{if} \quad -m\Lambda_{melt} < U < 0$$

(14)
\[ T = \frac{U}{mc_t + m_ic_i} \quad \text{and} \quad m = m_t \quad \text{if} \quad U > 0 \quad (15) \]

where \( \Lambda_{\text{melt}} \) is the melting heat (3.34 \times 10^5 J kg\(^{-1}\)), \( c_i \) is specific heat of water (4184 J kg\(^{-1}\) °C\(^{-1}\)), \( c_i \) is specific heat of ice (2092 J kg\(^{-1}\) °C\(^{-1}\)) and \( c_o \) is specific heat of soil (843 J kg\(^{-1}\) °C\(^{-1}\)). Equations (13) to (14) are also illustrated by Figure 2.

**Figure 2.** Temperature versus total energy

**Radiation**

Radiation is the main energy input for land surface models. Hence, it is not surprising that it has received a lot of attention. We follow a somewhat modified version of the approaches of Tian el al., (2001) and Allen et al., (2006) and divide radiation between shortwave, received from the sun, and longwave radiation, emitted by the Earth and the atmosphere. Net radiation (\( R_n \)) is the sum of the net shortwave radiation (\( R_{ns} \)) and the net longwave radiation (\( R_{nl} \)):

\[ R_n = R_{ns} + R_{nl} = (1 - A)R_S + R_{L_{up}} + R_{L_{down}} \quad (16) \]

where \( A \) is the albedo, fraction of solar radiation (\( R_S \)) that is reflected by the surface. The albedo depends on surface types (small for vegetated surface and high values for snow). The net longwave radiation is equal to the received (downward, \( R_{L_{down}} \)) minus the emitted (\( R_{L_{up}} \)) radiation.

**Solar radiation on a horizontal and on an inclined surface**

The solar radiation on a horizontal surface (\( R_{S_{hor}} \)) is measured or can be calculated from the atmospheric transmissivity (\( \tau_a \)), which is the fraction of extraterrestrial solar radiation (easy to compute, see Appendix) that makes it to the land surface. Atmospheric transmissivities highly sensitive to cloudiness and moisture content. It can be estimated from the relative sunshine hours (Allen et al.,1998) or through the method of Hargreaves and Allen (2003):

\[ \tau_a = K_{R} \left( \frac{T_{\text{max}}}{T_{\text{air}}} - \frac{T_{\text{min}}}{T_{\text{air}}} \right) \quad (17) \]
where $T_{\text{air}}^{\text{max}}$ and $T_{\text{air}}^{\text{min}}$ are the daily maximum and minimum air temperature and $K_H$ is an empirical constant. Allen et al., (1998) recommends $K_H = 0.16$ for interior and $K_H = 0.19$ for coastal regions.

Shortwave radiation may reach the land surface directly from the sun, reflected by the surrounding or scattered by the atmosphere (diffuse solar radiation). The distinction is relevant for inclined surfaces in the shade, which only receive reflected and diffuse solar radiation. Numerous relations can be found in the literature (Noorian et al., 2008) to estimate the fraction of diffuse radiation over the total solar radiation ($f_{\text{diff}}$). We adopted the one of Boland et al., (2008), which is simple and statistically sound:

$$f_{\text{diff}} = \frac{1}{1 + \exp (8.6 \tau_{\text{a}} - 5)}$$

Solar radiation on an inclined surface can be calculated from the solar radiation on a horizontal surface by means of the following expression (e.g., Tian et al., 2001):

$$R_S = R_{S,\text{hor}} \left[ (1 - f_{\text{diff}}) \frac{\max (p^T \mathbf{s}, 0)}{s_{\text{up}}} + f_{\text{diff}} f_{\text{sv}} + A (1 - f_{\text{sv}}) \right]$$

where $f_{\text{sv}}$ is the sky view factor (see equation A14 in the Appendix). The first term represents the direct solar radiation, the second term the diffuse radiation and the third term the solar radiation reflected from the surroundings. Definitions of $p^T \mathbf{s}$ and $s_{\text{up}}$ are shown in the Appendix (equations A1 and A2). To perform daily energy balances, we calculate the daily averaged solar radiation on an inclined surface, assuming the atmospheric transmissivity ($\tau_a$) to be constant during the day, as follows:

$$R_S = \left[ (1 - f_{\text{diff}}) \int_{-\omega_{\text{as}}}^{\omega_{\text{as}}} \frac{\max (p^T \mathbf{s}, 0)}{s_{\text{up}}} \, dt + f_{\text{diff}} f_{\text{sv}} + A (1 - f_{\text{sv}}) \right] R_{S,\text{hor}}$$

where $\omega_{\text{as}}$ is the sunset angle, the integrals of $s_{\text{up}}$ and $\max (p^T \mathbf{s}, 0)$ are given by equations A6 and A13 of the Appendix.

Longwave radiation

Longwave radiation appears simple, but actual parameterization is hard (Herrero and Polo, 2014; Zabel et al., 2012). Upward longwave radiation is calculated from Stefan-Boltzmann law:

$$R_{L,\text{up}} = -\varepsilon_s \sigma T_s^4$$

where $\varepsilon_s$ is the surface emissivity, $\sigma$ is the Stefan-Boltzmann constant ($5.7 \cdot 10^{-8}$ J s^{-1} m^{-2} K^{-4}$) and $T_s$ is the surface temperature (K). The soil surface emissivity is usually close to 1 (Saito and Šimůnek, 2009). However, small changes in $\varepsilon_s$ may cause an imbalance between upwards and downwards longwave radiation balance, thus having a large effect on net radiation. We adopted a constant value of 0.94, but also tested 0.99 for sensitivity analysis purposes.

The Earth's surface also receives longwave radiation emitted by the atmosphere and surrounding surfaces. It can be calculated from the same law:
where $\varepsilon_a$ is the emissivity of the atmosphere and $T_{air}$ the absolute temperature of the atmosphere. Note that a fraction equal to the sky view factor ($f_{sv}$) originates from the atmosphere and another part ($1-f_{sv}$) from the surroundings. Clear sky emissivity is obtained from the empirical expression of Brutsaert (1975):

$$\varepsilon_{air,cs} = 1.24 \left( \frac{P_{v,air}}{T_{air}} \right)^{1/7}$$   \hspace{1cm} (23)

The cloudy sky emissivity ($\varepsilon_a$) is obtained from $\varepsilon_{air,cs}$ using the expression that Sicart et al., (2006) derived empirically for a subarctic continental climate in Yukon (Canada):

$$\varepsilon_a = \varepsilon_{air,cs}(1 + 0.44 h_r - 0.18 \tau_a)$$   \hspace{1cm} (24)

where $h_r$ is relative humidity.

### Sensible heat

The sensible heat flux is calculated using the aerodynamic resistance ($r_a$) and soil surface resistance ($r_{sf}$):

$$H = \frac{\rho_a c_a}{r_a + r_{sf}} (T_{sf} - T_{air})$$   \hspace{1cm} (25)

where $\rho_a$ is the air density (1.22 kg m$^{-3}$) and $c_a$ is the specific heat of air (1013 J kg$^{-1}$ K$^{-1}$). The soil surface resistance, $r_{sf}$, is calculated from the thermal conductivity of the soil:

$$r_{sf} = \frac{0.5 L_{sf} \rho_a c_a}{\lambda_s}$$   \hspace{1cm} (26)

where $\lambda_s$ is the thermal conductivity of soil.

### Heat conduction

Heat conduction of soil can be calculated from Fourier’s Law as:

$$G = \frac{\lambda}{(0.5 L_{sf} + 0.5 L_{ss})} (T_{sf} - T_{ss})$$   \hspace{1cm} (27)

where $\lambda$ is a representative soil thermal conductivity ($\lambda = \lambda_i^{\phi} \lambda_s^{1-\phi}$), where $\phi$ is the porosity and $\lambda_i$, is the thermal conductivity of ice, $\lambda_s$, when water is frozen or that of liquid water, $\lambda_s$, otherwise (Côté and Konrad, 2005).

### 2.4. Numerical solution and implementation

We solve the water and energy balance equations (Eqs. (1), (2), (10) and (11), respectively) using a semi-implicit finite differences scheme with a time step of one day. The term semi-implicit means that all variables are treated explicitly (i.e., using the values from the previous day), with two exceptions to ensure stability. First, vapor pressures in equations (3), (4)...
and (9) are linearized and treated implicitly (that is, $p_{s+1} - p_{a+T} = (T_{s+1} - T_{a}) \Delta + (1 - h_2)p_{a+T}$, where $\Delta$ is the slope of the saturation vapor pressure curve (Pa °C$^{-1}$)). This type of linearization is frequent (e.g., Penman, 1948) because vapor pressure is highly sensitive to temperature and treating it explicitly may cause instability. Second, we perform a preliminary water balance to approximate the water available for evaporation, this is followed by the energy balance which yields not only energy and temperature, but also actual evaporation, that is used for the final water balance.

The algorithm was implemented in a spreadsheet that is available for the cases discussed below at http://h2ogeo.upc.edu/es/investigacion-hidrologia-subterrania/software.

### 2.5. Data of meteorological stations and parameters

The model was tested using meteorological data from 2000 through 2004 of the Terelj station (elevation 1540 m, 47.98°N, 107.45°E), located in northern Mongolia some 40 km east of Ulaanbaatar. This station records daily meteorological data (maximum and minimum $T$, precipitation, snow depth, wind and relative humidity) provided by the Institute of Meteorology and Hydrology of Mongolia. The area is mountainous with grassland and forest of Larix and Pinus. Forests dominate the north face of mountains while grassland dominates the south face of mountains and flat areas (Dulamsuren et al., 2008; Ishikawa et al., 2005). The region contains discontinuous and sparsely insular permafrost (Gravis et al., 1972; Sharkhuu, 2003; Jambaljav et al., 2008).

The average daily maximum and minimum air temperature is 5.06°C and -11.5°C, respectively. Mean air temperature averaged -3.2°C for the studied period. Annual precipitation averaged 334 mm/year, with 80% falling between June and September. Snow usually falls between mid-October and mid-April, with a maximum thickness of 31 cm. The average wind speed is 1.5 m/s and average relative humidity 70.12%.

We used parameters from the literature (Table 1) to define a base model that assumes that the surface is horizontal and covered by grass. Jambaljav et al., (2008) noted that the north and south facing slopes of mountains in the Terelj area are about 10-40°. Therefore, for the sensitivity analysis, we take into account both the north and south face with a slope of 20 degrees. Tuvshinjargal et al., (2004) used an albedo of 0.21 for grass meadow to calculate the surface energy balance. There are no other data for albedo from this area, especially snow albedo. The albedo ($A$) was taken from Oke (1987) as 0.6 during periods with snow cover and 0.23 for grass and soil surface. Most surfaces have emissivities larger than 0.9 (Arya, 2001). So, we used $\varepsilon = 0.94$ for the base model. We assumed that all the roots of grass are in the surface layer ($L_{sf}$) which means the evapotranspiration only occurs from the surface layer. Thus, we used $a = 1$.

The surface roughness length ($z_0$) is defined by surface types such as soil, vegetation and snow. We used $z_0 = 0.04$ for grass and $z_0 = 0.002$ for snow surface (see figure 10.5 of Arya, 2001). The leaf area index (LAI) is defined by vegetation types. Asner et al., (2003) give LAI of 2.1 m$^2$ m$^{-2}$ for grass. According to the National Soil Atlas of Mongolia (1981), the soil of the study area belongs to the Gleysols-umbrisol and Cryosols-leptic type. The soil texture is mainly a silt-clay-loam. The wilting point ($\theta^w$), field capacity ($\theta^f$), porosity ($\phi$) and saturated hydraulic conductivity ($K_{sat}$) are 0.11, 0.342, 0.365 and 4.2·10$^{-7}$ m s$^{-1}$, respectively, and were obtained from Schroeder et al., (1994). Length of the surface layer ($L_{ss}$) and subsoil layer are of 0.16 m and 1.5 m, respectively. Thermal conductivities were obtained from Bristow (2002).
For the sensitivity runs we changed eight parameters: dip of the surface ($\theta$), roughness length ($z_0$), soil emissivity ($\varepsilon_s$), saturated conductivity ($K_{sat}$), wilting point ($\theta_{wp}$), vapor diffusion coefficient ($D$), surface length ($L_{sf}$) and subsoil length ($L_{ss}$).

Table 1. Parameter and values for Base model

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
<th>Units</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope ($\theta$)</td>
<td>0</td>
<td>degree</td>
<td></td>
</tr>
<tr>
<td>Latitude ($\phi$)</td>
<td>48</td>
<td>degree</td>
<td></td>
</tr>
<tr>
<td>Albedo (A)</td>
<td>0.23</td>
<td>for grass and soil</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>for snow</td>
<td>Oke, 1987</td>
</tr>
<tr>
<td>Soil emissivity ($\varepsilon_s$)</td>
<td>0.94</td>
<td>-</td>
<td>Arya, 2001</td>
</tr>
<tr>
<td>Fraction of transpiration ($\alpha$)</td>
<td>1</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Vegetation cover ($\beta$)</td>
<td>0.6</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Leaf area index (LAI)</td>
<td>2.1</td>
<td>for grass</td>
<td>Asner et al., 2003</td>
</tr>
<tr>
<td></td>
<td>0.04</td>
<td>for grass</td>
<td>Arya, 2001</td>
</tr>
<tr>
<td></td>
<td>0.002</td>
<td>for snow</td>
<td></td>
</tr>
<tr>
<td>Surface roughness length ($z_0$)</td>
<td>0.04</td>
<td>m</td>
<td>Arya, 2001</td>
</tr>
<tr>
<td></td>
<td>0.002</td>
<td>m</td>
<td></td>
</tr>
<tr>
<td>Diffusion coefficient ($D$)</td>
<td>$10^{-4}$</td>
<td>m$^2$s$^{-1}$</td>
<td></td>
</tr>
<tr>
<td>Field capacity ($\theta^{(0)}$)</td>
<td>0.342</td>
<td>m$^3$m$^{-3}$</td>
<td>Schroeder et al, 1994</td>
</tr>
<tr>
<td>Wilting point ($\theta_{wp}$)</td>
<td>0.11</td>
<td>m$^3$m$^{-3}$</td>
<td>Schroeder et al, 1994</td>
</tr>
<tr>
<td>Porosity ($\phi$)</td>
<td>0.365</td>
<td>-</td>
<td>Schroeder et al, 1994</td>
</tr>
<tr>
<td>Saturated hydraulic conductivity ($K_{sat}$)</td>
<td>$4.2 \times 10^{-7}$</td>
<td>m s$^{-1}$</td>
<td>Schroeder et al, 1994</td>
</tr>
<tr>
<td>Surface depth ($L_{sd}$)</td>
<td>0.16</td>
<td>m</td>
<td></td>
</tr>
<tr>
<td>Subsoil depth ($L_{ss}$)</td>
<td>1.5</td>
<td>m</td>
<td></td>
</tr>
<tr>
<td>Thermal conductivity ($\lambda$)</td>
<td>2.9</td>
<td>for soil</td>
<td>Bristow (2002)</td>
</tr>
<tr>
<td></td>
<td>0.57</td>
<td>for water</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.2</td>
<td>for ice</td>
<td></td>
</tr>
</tbody>
</table>

3. Results

Results are summarized in Figure 3, which displays the evolution of the water fluxes (precipitation, evapotranspiration, recharge, surface runoff and vapor diffusion), water contents, heat fluxes (net radiation, latent heat, sensible heat, vapor convection and heat conduction) and temperature (air, surface and subsoil) of the base model during the last two years. Table 2 and 3 summarize the balances averaged over the 5 years for the base model and for the sensitivity.

Direct surface runoff is very small. Figure 3.b only shows some surface runoff at the beginning of April 2003 during snowmelt. No surface runoff occurred during the snowmelt of 2004 probably because the accumulated snow on the surface was small that year. The sensitivity analysis suggests that the limiting factor is the infiltration capacity. When saturated hydraulic conductivity is reduced by a factor of 10, direct surface runoff increases dramatically and infiltration reduces. The reduced infiltration also implies increased ET, so that the overall runoff (SR plus recharge) is also reduced. The small surface runoff is consistent with the lack of very intense rainfall events and with the absence of indications of surface erosion.
Figure 3. Daily evolution of (a) precipitation and recharge, (b) other water fluxes, (c) water content at surface (solid line) and subsoil (dashed line), (d, e) heat fluxes, and (f) temperatures of the base model.
Infiltration and recharge are also relatively low. Infiltration occurs mainly after heavy rainfall events and it is not shown in figure 3 because similar pattern as recharge. Most infiltration transforms into recharge because, in the absence of deep rooted plants, the subsoil is always close to field capacity. Recharge from rainfall infiltration can vary a lot from year to year, due to the irregular occurrence of heavy rainfall events. However, a significant amount of recharge occurs throughout the spring and summer driven by vapor diffusion into the subsoil. While the rate is small (it can hardly be seen in Figure 3a, so we zoom it in Figure 4), it occurs throughout the late spring and summer, after the subsoil has started to thaw. Overall, it is about half of recharge from direct rainfall infiltration, but much more regular (it occurs every year) and quite robust, in that it displays little sensitivity to model results (See Table 2).

Vapor diffusion between the surface and subsoil layers is positive (downwards) during spring and early summer, because then the temperature and, therefore, vapor pressure is higher in the surface than in the subsoil. The flux fluctuates during late winter, when the subsoil has started to warm, so that vapor diffuses upwards during cold days and downwards during warm days. Diffusion is consistently upwards during autumn and winter, but the rate is very low because the saturated vapor pressure is low and changes little with temperature (that is, $dp_{v,sat}/dT$ is small) at low temperatures. Therefore, there is a net downward vapor flux. Its amount is not very large, but as mentioned above, it is what drives recharge during spring and early summer.

![Figure 4. Zoom of recharge (red dashed line) and vapor diffusion (blue line) during 2003 (Temperatures are also shown). Note that, except for the two heavy rainfall events of May and August, recharge during the warming period (after the subsoil has started to thaw) is identical to the vapor diffusion flux.]

Vapor diffusion is basically controlled by the diffusion coefficient ($D$), which is the only parameter that affects the vapor diffusion flux significantly (Table 2). Reducing $D$ leads obviously to a reduction of vapor diffusion. A similar effect results from increasing the soil surface thickness, which results in an apparent reduction of the gradient. As less water is transported downwards, evapotranspiration increases and recharge decreases.
As expected, evapotranspiration is the main sink of water. In fact, it is limited by water availability during the warm season, being high only after rainfall events and during melting (Figure 3.b). The evapotranspiration is about 85% of rainfall (Table 2), which is similar to the results obtained by Ma et al., (2003) for the Selenge River basin, northern Mongolia. Note that the evapotranspiration is very small during winter because low temperatures hinder vaporization. In fact, it is negative (i.e., ice deposition) during January and February (Figure 5), when the soil is colder than the air. Sublimation only becomes relevant in March. Cumulative sublimation (some 29 kg m\(^{-2}\) year\(^{-1}\) in 2003) is low compared to typical values of cold regions (see, e.g., Zhou et al., 2014), but large compared to winter rainfall.

The whole cycle is driven by radiation, which follows the usual seasonal patterns, high during late spring and early summer and low in winter, when net radiation may become negative, partly due to the high albedo of snow (Figure 3.d). The radiation balance is highly sensitive to orientation of the slope (Table 3). Obviously, the south face receives more radiation than the north face. But this is largely compensated by an increase in sensible heat flux. The sensible heat increases when latent heat decreases. That is, heat is returned to the atmosphere either as latent heat when water is available for evaporation, or as sensible heat when the soil is dry. According to the energy balance (Table 3), the sensible heat is higher than the latent heat flux, which reflects the dry climate of region. As a result the effect of slope and orientation is smaller than we had anticipated (Tables 2 and 3). The large increase in radiation of south facing slopes only results in a small increase in evaporation and latent heat because there is little water and a parallel reduction of infiltration and recharge decreases (Table 2).

**Figure 5.** Zoom of ET (actually, water phase change processes) (black line) along with liquid (blue) and ice (red) water contents during 2003.

The dependence of the two balances on slope is non-monotonic, which points to the complexity of the system, even in the relatively simple model we are presenting here. Radiation is dramatically reduced in north facing slopes, which causes a reduction in ET, but the reduction is very small as discussed above, and not sufficient to cause an increase in infiltration. The reduction in ET is compensated by an increase in surface runoff and latent heat diffusion (vapor convection) downwards.
The non-monotonic dependence of vapor convection on slope also illustrates the robustness of vapor diffusion. It is slightly larger in south-facing slopes than in horizontal land because surface temperatures are also larger. But it is also slightly larger in north-facing slopes than in horizontal land because subsoil temperatures are lower.

**Table 2.** Average in 5 years of water fluxes: evapotranspiration (ET), infiltration (I), surface runoff (SR), vapor diffusion ($J_D$) and Recharge (R). Precipitation is 334 kg m$^{-2}$ year$^{-1}$

<table>
<thead>
<tr>
<th>Water Bal. [kg m$^{-2}$ year$^{-1}$]</th>
<th>ET</th>
<th>I</th>
<th>SR</th>
<th>$J_D$</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base model</td>
<td>284.9</td>
<td>30.9</td>
<td>0.1</td>
<td>18.1</td>
<td>48.8</td>
</tr>
<tr>
<td>South face</td>
<td>290.8</td>
<td>23.8</td>
<td>0.0</td>
<td>19.5</td>
<td>43.3</td>
</tr>
<tr>
<td>North face</td>
<td>283.8</td>
<td>30.2</td>
<td>1.0</td>
<td>19.0</td>
<td>49.1</td>
</tr>
<tr>
<td>$z_0$ (x 2)</td>
<td>289.8</td>
<td>26.8</td>
<td>0.0</td>
<td>17.5</td>
<td>44.1</td>
</tr>
<tr>
<td>$e_s$ (0.94 → 0.99)</td>
<td>280.6</td>
<td>35.5</td>
<td>1.0</td>
<td>16.9</td>
<td>52.2</td>
</tr>
<tr>
<td>$K_u$ (x 0.1)</td>
<td>288.2</td>
<td>12.8</td>
<td>14.9</td>
<td>18.1</td>
<td>30.6</td>
</tr>
<tr>
<td>$\Theta_{wp}$ (x 2)</td>
<td>252.9</td>
<td>66.0</td>
<td>2.1</td>
<td>13.1</td>
<td>79.1</td>
</tr>
<tr>
<td>$D$ (x 0.25)</td>
<td>296.1</td>
<td>32.6</td>
<td>0.0</td>
<td>5.3</td>
<td>37.8</td>
</tr>
<tr>
<td>$L_{sd}$ (x 2)</td>
<td>318.5</td>
<td>4.2</td>
<td>0.0</td>
<td>11.4</td>
<td>15.6</td>
</tr>
<tr>
<td>$L_{sn}$ (1.5 → 2.15)</td>
<td>267.9</td>
<td>30.2</td>
<td>0.0</td>
<td>36.0</td>
<td>66.2</td>
</tr>
</tbody>
</table>

The limited availability of water causes most rainfall to evaporate. It also implies a low sensitivity of ET parameters to water and energy fluxes. For example, increasing roughness length ($z_0$) decreases the aerodynamic resistance ($r_a$), which leads to small increases in both latent and sensible heat fluxes and a parallel small decrease in infiltration and recharge. Similarly, increasing land surface emissivity from 0.94 to 0.99 reduces considerably net radiation as more longwave radiation is emitted, but is compensated by a decrease of sensible and, to a lesser extent, latent heat fluxes.

**Table 3.** Average in 5 years of energy fluxes: net radiation ($R_n$), Latent heat ($e_s$ET), sensible heat (H), heat conduction (G) and vapor convection ($e_sJ_D$).

<table>
<thead>
<tr>
<th>Energy Bal. [MJ m$^{-2}$ year$^{-1}$]</th>
<th>$R_n$</th>
<th>$e_s$ET</th>
<th>H</th>
<th>G</th>
<th>$e_sJ_D$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base model</td>
<td>2064.4</td>
<td>717.1</td>
<td>1345.6</td>
<td>-45.8</td>
<td>45.9</td>
</tr>
<tr>
<td>North face</td>
<td>2610.7</td>
<td>732.3</td>
<td>1876.8</td>
<td>-49.1</td>
<td>49.4</td>
</tr>
<tr>
<td>$z_0$ (x 2)</td>
<td>1708.4</td>
<td>714.6</td>
<td>992.3</td>
<td>-48.0</td>
<td>48.1</td>
</tr>
<tr>
<td>$e_s$ (0.94 → 0.99)</td>
<td>2181.4</td>
<td>729.5</td>
<td>1450.5</td>
<td>-44.2</td>
<td>44.3</td>
</tr>
<tr>
<td>$K_u$ (x 0.1)</td>
<td>1739.1</td>
<td>706.4</td>
<td>1030.9</td>
<td>-42.9</td>
<td>42.9</td>
</tr>
<tr>
<td>$\Theta_{wp}$ (x 2)</td>
<td>2067.2</td>
<td>725.4</td>
<td>1340.1</td>
<td>-45.4</td>
<td>45.9</td>
</tr>
<tr>
<td>$D$ (x 0.25)</td>
<td>2041.5</td>
<td>636.6</td>
<td>1402.1</td>
<td>-34.4</td>
<td>33.1</td>
</tr>
<tr>
<td>$L_{sd}$ (x 2)</td>
<td>2072.4</td>
<td>745.5</td>
<td>1325.8</td>
<td>-14.3</td>
<td>13.6</td>
</tr>
<tr>
<td>$L_{sn}$ (1.5 → 2.15)</td>
<td>1980.5</td>
<td>801.8</td>
<td>1177.9</td>
<td>-28.6</td>
<td>29.0</td>
</tr>
<tr>
<td>$L_{sd}$ (1.5 → 2.15)</td>
<td>2032.0</td>
<td>675.2</td>
<td>1355.5</td>
<td>-91.4</td>
<td>91.1</td>
</tr>
</tbody>
</table>
The most surprising energy balance terms are heat conduction (soil heat flux) and vapor convection (Figure 3.e). Conductive heat flux is usually considered seasonal, with yearly averages close to zero. Downward heat fluxes in summer are usually balanced by upward fluxes in winter (e.g. Alkhaier et al., 2012). In our case, even though the subsoil remains frozen for long (more than 7 months, compared to less than six the surface layer), there is a net flux upwards, to compensate the latent heat convection associated to vapor diffusion, which is downwards, as was discussed before. Therefore, it is not surprising that all factors that reduce the soil heat flux cause an increase in vapor convection, and vice versa.

The temperature oscillates more at the surface than at the subsoil layer (Figure 3.f). The differences of temperature of the layers are higher in summer. One can observe clearly the periods of melting and freezing of the subsoil layer with temperatures of 0°C. The annual average surface layer temperature was 0.8°C for 2003 and 2.0°C for 2004 while the air temperature was -3.6°C for 2003 and -2.4°C for 2004. For subsoil layer, -0.5°C and 0.7°C, in 2003 and in 2004 respectively.

An increase of the subsoil length \( L_{ss} \) leads to temperatures in the subsoil that oscillate less due to the increased heat storage capacity. This leads to larger temperature differences between surface and subsoil, which according to our model (equation 9, Figure 6) leads to larger vapor diffusion. As more water is transported downwards, evapotranspiration decreases and recharge increases.

![Figure 6. Daily evolution of surface and subsoil temperature in 2004.](image)

4. Discussion and Conclusions

We have developed a water and energy balance model that contains two layers and attempts to represent all terms relevant for simulating land surface hydrological processes, including all possible phase changes and, singularly, vapor diffusion. The model has been applied using meteorological data from the Terelj station, northern Mongolia and typical soil properties of the region. Results are consistent with local observations by others:

- Direct surface runoff is negligible and restricted to snowmelt periods.
- Liquid infiltration and subsequent recharge are restricted to a few heavy rainfall events. However, a sizable recharge (about half of that from rainfall events) occurs continuously during late spring and early summer.
Evapotranspiration is limited by water availability, as it accounts for 85% of rainfall. Sublimation is restricted to late fall and spring, but it is also large, compared to winter snowfall. Ice deposition occurs most days during January and February.

Sensible heat is higher than latent heat flux, which reflects the dry climate of the region and low precipitation.

The active layer remains frozen during the winter with periods of freezing or thawing of some three months, a length of time that increases when the thickness of the active layer increases.

In summary, results are qualitatively consistent with observations. Notably, total runoff would be too small, compared to observations, if vapor diffusion is reduced. The most singular result of the simulations is the relative importance of vapor diffusion, which is downwards during spring and early summer, when temperature and, therefore, vapor pressure are higher in the surface than in the subsoil. The upwards vapor diffusion flux is much smaller than the downward one, because vapor pressure is a non-linear function of temperature. This downward flux transforms into recharge, which is continuous, although fluctuating during that period.

In summary, the net downward vapor flux is relevant both in terms of water balance, accounting for a sizable portion of recharge, and energy balance, causing a net upwards flux of heat. We conclude that land surface schemes should account for vapor diffusion. We notice that, being a diffusive process, it may be included in such schemes at a moderate effort. Still, further research is needed to ascertain the right values of diffusion coefficient to be used and the way of discretizing Fick’s Law, that is, the choice of length over which diffusion takes place in equation (9).

Appendix

Position of the sun

For the calculation of the position of the sun and the zenith, it is convenient to define two unit vectors: \( \mathbf{p} \), which is orthogonal to the land surface and \( \mathbf{s} \), which points to the sun (Figure A1). Their first component points eastwards, the second northwards and the third upwards. Vector \( \mathbf{p} \) can be calculated from the strike (\( \sigma \)) and dip (\( \theta \)) (see also figure A2):

\[
\mathbf{p} = \begin{pmatrix} p_{\text{east}} \\ p_{\text{north}} \\ p_{\text{up}} \end{pmatrix} = \begin{pmatrix} \cos \sigma \sin \theta \\ -\sin \sigma \sin \theta \\ \cos \theta \end{pmatrix}
\] (A1)

Vector \( \mathbf{s} \) depends on the time of the day, the sun declination (\( \delta \)), solar angle (\( \omega \)) and the latitude (\( \phi \)). It can be calculated according to Sproul (2007) as:

\[
\mathbf{s} = \begin{pmatrix} s_{\text{east}} \\ s_{\text{north}} \\ s_{\text{up}} \end{pmatrix} = \begin{pmatrix} -\cos \delta \sin \omega \\ \sin \delta \sin \phi \cos \omega - \cos \delta \sin \phi \cos \omega \\ \cos \delta \cos \phi \cos \omega + \sin \delta \sin \phi \cos \omega \end{pmatrix}
\] (A2)
The product of both vectors \( \mathbf{p} \cdot \mathbf{s} \) equals the cosine of the angle between them. Note, that on a horizontal surface \((\theta = 0)\), \( \mathbf{p} \cdot \mathbf{s} \) is \((0,0,1)\) and \( \mathbf{p} \cdot \mathbf{s} = s_{up} \). At night \( s_{up} < 0 \), at daylight \( s_{up} > 0 \) and at sunrise and sunset \( s_{up} = 0 \). Furthermore, an inclined surface is in the shade when \( \mathbf{p} \cdot \mathbf{s} < 0 \).

The sun declination \((\delta)\) is the angle between the direction of the sun and the equator. It can be calculated by a yearly sinusoidal function:

\[
\delta = -\delta_{\text{max}} \sin \left(2\pi \frac{t - t_e}{d_a}\right)
\]  

(A3)

where \(\delta_{\text{max}}\) is the maximum sun declination (\(0.4091 \text{ rad} = 23.26^\circ\)), \(t\) is time \((s)\), \(t_e\) is time at September equinox, approximately September 21\(^{st}\), \((s)\) and \(d_a\) is the duration of year (\(\approx 365.241\) days \(\approx 3.15568 \times 10^7\) s).

374

**Figure A1.** Illustration of vectors \( \mathbf{p} \) and \( \mathbf{s} \)

376

The extraterrestrial solar radiation on a horizontal surface can be simplified by:

\[
R_{\text{s,et,hor}} = S_0 f_e s_{up} = S_0 f_e (\cos \delta \cos \varphi \cos \omega + \sin \delta \sin \varphi)
\]

If \( s_{up} > 0 \)

\[
R_{\text{s,et,hor}} = 0
\]

otherwise

(A4)

where \(S_0\) is the sun constant \((1367 \text{ J m}^{-2} \text{ s}^{-1})\) and \(f_e\) is the factor that corrects for the eccentricity of the earth’s orbit. It can be calculated from (Allen et al., 1998):

\[
f_e = 1 + 0.033 \cos (2\pi \frac{t - t_{ph}}{d_a})
\]

(A5)

where \(t_{ph}\) is the time at perihelion, approximately January 3\(^{rd}\), \((s)\).

380

For the daily averaged extraterrestrial solar radiation on a horizontal surface, integrating equation (A4) between sunset and sunrise and dividing by the duration of a day \((d_a)\) gives:

\[
R_{\text{s,et,hor}} = \frac{S_0 f_e}{d_a} \int_{-\omega_{as}}^{\omega_{as}} s_{up} \, dt = \frac{S_0 f_e}{d_a} (\cos \delta \cos \varphi \sin \omega_{as} + \omega_{as} \sin \delta \sin \varphi)
\]

(A6)

where the sunset angle, \(\omega_{as}\), is the solar angle when \(s_{up}\) equals 0:

\[
\omega_{as} = \cos^{-1} \left(\max \left(\min (\tan \varphi \tan \delta, 1), -1\right)\right)
\]

(A7)
The min and max functions guarantee that $\omega_{\text{az}} \in [0, \pi]$. Thus, equation (A6) also works for days when the sun doesn't set or rise (i.e., in polar regions).

**Correction for an inclined surface**

The strike ($\sigma$) of an inclined plane is the orientation of a horizontal line on this plane, expressed as an angle relative to the north in clockwise direction. The dip ($\theta$) is the maximum angle between a horizontal plane and the incline plane (figure A2).

The extraterrestrial solar radiation on an inclined surface is the solar radiation without taking into account the reduction of it by the atmosphere. It can be expressed as:

$$R_{\text{xet,inc}} = S_0 f_s \text{max}(p^T s, 0)$$

If $s_{\text{up}} > 0$

$$R_{\text{xet,inc}} = 0$$

otherwise

(A8)

Note that for a horizontal surface ($\theta = 0 \Rightarrow p^T s = s_{\text{up}}$) equation (A8) reduces to (A4). For the daily averaged extraterrestrial solar radiation on an inclined surface, we have to integrate equation (A8):

$$R_{\text{xet,inc}} = \frac{S_0 f_s}{d} \int_{-\omega_{\text{az}}}^{\omega_{\text{az}}} \text{max}(p^T s, 0) \, dt$$

(A9)

*Figure A2.* Illustration of strike ($\sigma$) and dip ($\theta$). Note that a dip between $-0.5\pi$ and $0.5\pi$ ($-90^\circ$ and $90^\circ$) refers to a plane with its outside facing upwards. A dip between $1.5\pi$ and $0.5\pi$ ($90^\circ$ and $270^\circ$) or between $-1.5\pi$ and $-0.5\pi$ refers to a plane with its outside facing downwards.

However, we have to take into account that the surface can be in the shade during part of the day, which complicates the calculations. The integral of $\text{max}(p^T s, 0)$ can be calculated by dividing it into different periods, when it is in the shade or not.
\[
\int_{-\omega_{\text{min}}}^{\omega_{\text{max}}} \max(p^r s, 0) \, dt = \max\left(\int_{-\omega_{\text{min}}}^{\omega_1} p^r s \, dt, 0\right) + \max\left(\int_{\omega_1}^{\omega_2} p^r s \, dt, 0\right) + \max\left(\int_{\omega_2}^{\omega_{\text{max}}} p^r s \, dt, 0\right)
\]

\[\omega_1 = \min\left(\max(-\omega_{\text{min}}, \omega_{\text{lo1}}), \omega_{\text{lo2}}\right)\]

\[\omega_2 = \min\left(\max(-\omega_1, \omega_{\text{lo2}}), \omega_{\text{lo3}}\right)\]

where \(\omega_{\text{lo1}}\) and \(\omega_{\text{lo2}}\) are the two solar angles in a day when the inclined surface comes out of or into the shade, that is, when \(p^r s = 0\). They are calculated from:

\[
\begin{align*}
\omega_{\text{lo1}} &= \min(\mod(\pi - \omega_b + \omega_c, 2\pi) - \pi, \mod(-\omega_b - \omega_c, 2\pi) - \pi) \\
\omega_{\text{lo2}} &= \max(\mod(\pi - \omega_b + \omega_c, 2\pi) - \pi, \mod(-\omega_b - \omega_c, 2\pi) - \pi)
\end{align*}
\]

\[\begin{align*}
\omega_b &= \frac{b}{a\sqrt{1 + b^2/a^2}} \\
\omega_c &= \frac{c}{a\sqrt{1 + b^2/a^2}}
\end{align*}\]  

(A11)

else (there is no solution for \(p^r s = 0\))

\[\begin{align*}
\omega_{\text{lo1}} &= -\pi \\
\omega_{\text{lo2}} &= \pi
\end{align*}\]

where \(\mod\) is a function that returns the remainder of the first argument after it is divided by the second argument. This guarantees that \(\omega_1\) and \(\omega_2\) are in \([-\pi, \pi]\). Moreover, the min and max functions guarantee that \(\omega_{\text{lo1}} \leq \omega_{\text{lo2}}\). When there is no solution for \(p^r s = 0\), it means that during the whole day the inclined surface is facing the sun or not. The integral of \(p^r s\) can be calculated by using equations (A1) and (A2):

\[
\int_{\omega_{\text{ini}}}^{\omega_{\text{fin}}} p^r s \, dt = \int_{\omega_{\text{ini}}}^{\omega_{\text{fin}}} (a \sin \omega + b \cos \omega - c) \, dt =
\]

\[\frac{d}{2\pi} \left[-a(\cos \omega_{\text{fin}} - \cos \omega_{\text{ini}}) + b(\sin \omega_{\text{fin}} - \sin \omega_{\text{ini}}) - c(\omega_{\text{fin}} - \omega_{\text{ini}})\right]\]

(A12)

with
\[ a = - \cos \sigma \sin \theta \cos \delta \]

\[ b = \sin \sigma \sin \theta \cos \delta \sin \varphi + \cos \theta \cos \delta \cos \varphi \]  \hspace{1cm} (A13)

\[ c = \sin \sigma \sin \theta \sin \delta \cos \varphi - \cos \theta \sin \delta \sin \varphi \]

**Sky view factor**

The sky view factor \( f_{sv} \), is the proportion of the sky above the inclined surface that is not blocked from view by the surrounding horizontal plane (figure A3). It ranges from 0.5 for a vertical to 1 for a horizontal surface. In general one can use the formula of Badescu (2002):

\[ f_{sv} = \frac{\cos(2\varphi)}{4} + 2 \]  \hspace{1cm} (A14)

**Figure A3. Illustration of the sky view factor; \( f_{sv} \).**

**References**


