Effects of spatial aggregation on the multi-scale estimation of evapotranspiration

A. Ershadi a,⁎, M.F. McCabe a,b, J.P. Evans c, J.P. Walker d

a School of Civil & Environmental Engineering, the University of NSW, Sydney, Australia
b Water Desalination and Reuse Centre, King Abdullah University of Science and Technology (KAUST), Jeddah, Saudi Arabia
c Climate Change Research Centre, the University of NSW, Sydney, Australia
d Department of Civil Engineering, Monash University, Clayton, Australia

A R T I C L E I N F O

Article history:
Received 3 September 2012
Received in revised form 1 December 2012
Accepted 8 December 2012
Available online XXXX

Keywords:
Landsat
MODIS
Surface Energy Balance System (SEBS)
Flux aggregation
Input aggregation
Upscaling
Land surface temperature
Uncertainty
Roughness
Aerodynamic resistance

A B S T R A C T

The influence of spatial resolution on the estimation of land surface heat fluxes from remote sensing is poorly understood. In this study, the effects of aggregation from fine (< 100 m) to medium (approx. 1 km) scales are investigated using high resolution Landsat 5 overpasses. A temporal sequence of satellite imagery and needed meteorological data were collected over an agricultural region, capturing distinct variations in crop stage and phenology. Here, we investigate both the impact of aggregating the input forcing and of aggregating the derived latent heat flux. In the input aggregation scenario, the resolution of the Landsat based radiance data was increased incrementally from 120 m to 960 m, with the land surface temperature calculated at each specific resolution. Reflectance based land surface parameters such as vegetation height and leaf area index were first calculated at the native 30 m Landsat resolution and then aggregated to multiple spatial scales. Using these data and associated meteorological forcing, surface heat fluxes were calculated at each distinct resolution using the Surface Energy Balance System (SEBS) model. Results indicate that aggregation of input forcing using a simple averaging method has limited effect on the land surface temperature and available energy, but can reduce evapotranspiration estimates at the image scale by up to 15%, and at the pixel scale by up to 50%. It was determined that the predominant reason for the latent heat flux reduction in SEBS was a decrease in the aerodynamic resistance at coarser resolutions, which originates from a change in the roughness length parameters of the land surface due to the aggregation. In addition, the magnitude of errors in surface heat flux estimation due to input aggregation was observed to be a function of the heterogeneity of the land surface and evaporative elements. In examining the response of flux aggregation, fine resolution (120 m) heat fluxes were aggregated to coarser resolutions using a range of common spatial interpolation algorithms. Results illustrate that a simple averaging scheme provides the best choice for flux aggregation compared to other approaches such as nearest neighbour, bilinear interpolation or bicubic interpolation, as it not only preserves the spatial distribution of evapotranspiration, but most importantly also conserves the mass balance of evaporated water across pixel and image scales.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Evapotranspiration (ET) is a complex process that incorporates interactions across a range of terrestrial and atmospheric variables, including the land surface temperature, air temperature, wind speed and humidity, as well as vegetation height and density (Brutsaert, 1982). As a consequence, evapotranspiration can be highly variable in space and time, particularly over heterogeneous surfaces. Given the importance of evapotranspiration in characterising aspects of the hydrological cycle, understanding the nature and degree of this variability has been an ongoing effort in the hydrological and related sciences (Anderson et al., 2003; Brunsell & Anderson, 2011; Entekhabi & Eagleson, 1989; McCabe & Wood, 2006; Settle & Drake, 1993). While there are established methods to estimate surface heat fluxes at the point scale (e.g. scintillometry and eddy covariance techniques), such local scale estimates cannot easily be extrapolated beyond the field to basin scales: although there are some approaches that attempt to do this (Jung et al., 2009). Given the spatial and temporal variability of the evapotranspiration process (McCabe et al., 2005), a practical method for the routine estimation of spatially distributed heat fluxes at both field and basin scales is through the use of remote sensing techniques (Allen et al., 2007a; Anderson et al., 2003; Bastiaanssen et al., 1998a; Norman et al., 1995; Su, 2002).

A number of remote sensing evapotranspiration models like SEBS (Su, 2002), SEBAL (Bastiaanssen et al., 1998a) and METRIC (Allen et al., 2007b) have been developed and validated at the patch scale using fine resolution satellite imagery (e.g. Landsat and ASTER) (Allen et al., 2007a; Bastiaanssen et al., 1998b; Choi et al., 2009; Tasumi & Allen, 2007; Timmermans et al., 2007; van der Kwast et al., 2009). However, there are cases when these methods are used with much coarser resolution data from sensors like AATSR (Advanced
Along Track Scanning Radiometer), MODIS (Moderate Resolution Imaging Spectroradiometer) and AVHRR (Advanced Very High Resolution Radiometer) (Elhag et al., 2011; Gibson et al., 2011; Golmen et al., 2012; Jia et al., 2003; Zwart & Bastiaanssen, 2007) in order to broaden the scope of their application. Unfortunately, the effects of subsequent changes in the spatial resolution on modelling performance and the implicit scaling influences that occur as a result of these are addressed in relatively few studies (Gebremichael et al., 2010; Hong et al., 2009; Long et al., 2011; McCabe & Wood, 2006; Tian et al., 2012).

Validation of turbulent heat fluxes using remote sensing algorithms can be significantly influenced by the spatial resolution of the data (Su et al., 1999). In particular, aggregation of the input forcing can have mixed influences on the evaluation of the resultant heat flux (Brunsell & Gillies, 2003; McCabe & Wood, 2006; Su et al., 1999). Likewise, validating coarse resolution measurements is generally more difficult due to the additional uncertainty introduced by the scale discrepancy between ground measurements and the coarse spatial resolution imagery (Gebremichael et al., 2010; Hong et al., 2009; Long et al., 2011). Furthermore, the mismatch between the variable being represented and the resolution at which it can be retrieved provides another level of uncertainty in the estimation process. For instance, the effects of spatial resolution of land surface temperature and roughness parameters on heat and vapour transfer are not well understood, particularly as the satellite resolution increases (Becker & Li, 1995; Brunsell & Anderson, 2011).

To address such issues, a number of studies have evaluated the aggregation (or up-scaling) effects on heat flux estimation (Brunsell & Gillies, 2003; Entekhabi & Eagleson, 1989; Familetti & Wood, 1994; Hong et al., 2009; Kustas & Norman, 2000; Kustas et al., 2004; McCabe & Wood, 2006; Nakagawa et al., 2001; Su et al., 1999; Wang & Currit, 2011). In general, aggregation can be applied either on the input forcing of the evapotranspiration models, or it can be applied to the fluxes derived from fine resolution input fields (i.e. aggregate then calculate versus calculate then aggregate). To examine these different approaches further, the concepts of ‘input aggregation’ and ‘flux aggregation’ are explored.

Aggregation of the input forcing has an immediate influence on the representative heterogeneity of the surface and affects the land surface control on heat flux generation (Brunsell & Gillies, 2003). One of the underlying assumptions for most physically based evapotranspiration models is the requirement for homogeneous conditions across a pixel, including homogeneity in both land surface (vegetation type, roughness, temperature) and meteorological conditions. To date, the effect of spatio-temporal variability of surface and atmospheric fields on heat flux generation remains poorly explained and quantified (Brunsell et al., 2008).

In addition to the aggregation of input forcing, resultant surface heat fluxes may require subsequent aggregation for a range of purposes e.g. to allow spatially consistent comparison and evaluation of General Circulation Model (GCM) and Regional Climate Model (RCM) outputs (Jiménez et al., 2011; Mueller et al., 2011). Likewise, GCM and RCM models require input forcing with a coarse spatial resolution that is generally much larger than the spatial resolution of remote sensing sensors. Therefore, an aggregation procedure is used to bridge the scale gap between remote sensing derived fluxes and the input requirements for large scale models (Hong et al., 2009). Further, flux aggregation is useful (and sometimes necessary) in comparisons of heat fluxes derived from geostationary images and those from polar-orbiting satellites (Brunsell & Anderson, 2011).

Moran et al. (1997) evaluated the effect of radiance aggregation on temperature and consequently on the sensible heat flux over a semi-arid rangeland in Arizona, finding negligible change in the land surface temperature, but large errors (more than 50%) in the sensible heat flux across heterogeneous areas having small vegetation elements within the pixels. The authors indicated that the uncertainty in flux estimation by input aggregation was due mainly to the non-linearity of the relations between the sensor signals, estimated variables and fluxes, and the inherent heterogeneity of the landscape. Hong et al. (2009) examined the aggregation of radiance from Landsat ETM+ resolution (30 m) to MODIS resolution (250 m) using the SEBAL model (Bastiaanssen et al., 1998a) and found that the peak of the histogram of latent heat flux increased 10–25% due to input aggregation. In a related study, Gebremichael et al. (2010) found that both input and flux aggregation procedures produced similar spatial patterns in SEBAL. Recently, Long et al. (2011) found that input aggregation of Landsat data to MODIS resolutions resulted in similar spatial mean values of sensible heat flux but with smaller spatial standard deviations.

For studies examining flux aggregation, Moran et al. (1997) found that errors in the aggregation of turbulent fluxes were highly influenced by the heterogeneity of the site and due mainly to variations in atmospheric stability, aerodynamic roughness, and patchy vegetation structures. Separate to the underlying surface heterogeneity, Srirhar et al. (2003) evaluated the performance of the nearest neighbour, bilinear, and bicubic interpolation methods for aggregation of evapotranspiration, finding that nearest neighbour and bilinear methods provided better performance than bicubic interpolation. Hong et al. (2009) found that flux aggregation using simple averaging and nearest neighbour methods can preserve the mean value of the original image and that the nearest neighbour method performed better than simple averaging by preserving the spatial variability of the fluxes.

While the majority of previous aggregation studies have employed semi-empirical evapotranspiration methods (e.g. SEBAL), a physically-based approach is used here for simulation of the land surface interactions and heat flux estimation. Doing this provides an opportunity to directly quantify the effect of input aggregation on each of the contributing components of heat flux estimation, including the land surface temperature, roughness parameters, aerodynamic resistance, and available energy. Also, while flux aggregation techniques have been examined previously (see above), an evaluation of these approaches based on conservation of the evaporative mass balance has not been examined. Preservation of the evaporated water volume across scales provides a superior measure of performance of the flux aggregation than considering spatial statistical aspects of the aggregation alone.

In this research effort, we examine the following hypotheses: a) that the effect of aggregation on input variables and parameters is not equal and that the roughness parameters are more significantly influenced by aggregation than other input variables; b) that errors due to the input aggregation are a function of the land surface heterogeneity and the size of the evaporative elements; and c) that a simple averaging approach is the best candidate for flux aggregation based on preservation of the hydrological mass balance.

2. Description of study area and data sources

The focus of these investigations is a heterogeneous agricultural region located in a semi-arid environment in the south-east of Australia. The 10.8 × 10.8 km region is situated in the agriculturally rich and economically important Murrambuggee catchment (a sub-catchment of the Murray Darling Basin), comprising natural drylands in the northern and eastern directions and active irrigation areas elsewhere. An irrigation canal passes through the study area from the south to the northeast (see Fig. 1).

2.1. Meteorological data

A meteorological station located in the centre of the study site provided the necessary meteorological forcing data. Observed variables included half-hourly air temperature, humidity, wind speed and atmospheric pressure, along with a Kipp and Zonen CNR1 four way net radiometer that provided detailed radiation budget components.
Fig. 1. Location of the study area in Australia’s Murray-Darling Basin (top-left), with the Murrumbidgee sub-catchment (top-middle) and the Coleambally irrigation area (top-right) also identified. Bottom panels show the Landsat band combination (7-4-2) colour composites for selected days at 120 m resolution. Green indicates vegetation and pink to magenta indicates bare soil. MIA represents the Murrumbidgee Irrigation Area, CIA the Coleambally Irrigation Area, and MIL the Murrumbidgee Irrigation Limited regions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 2. Variations of net radiation ($R_n$ in W m$^{-2}$), air temperature ($T_a$ in °C), and wind speed ($u_a$ in m/s) for selected days. The thick grey line shows the time of the Landsat satellite overpass.
Fig. 2 presents the daily variations of net radiation, air temperature and wind speed for the selected days.

2.2. Satellite data

Landsat 5 TM overpasses (Path 93, Row 84) were evaluated during the cropping calendar, with three satellite images representing different crop-growth stages selected on 6 September 2010, 18 November 2010 and 13 February 2011 (see Fig. 1). As shown in Fig. 2, 6 September 2010 is not a clear sky day, although no clouds are apparent in the acquired Landsat image. However, other times within the 30 minute period of tower observations (corresponding to the Landsat overpass) might have been affected by clouds. Therefore, meteorological variables are possibly uncoupled from the land surface conditions seen in the Landsat image of 6 September 2010, which results in uncertainties for evapotranspiration estimations based on this image. Images and tower records for 18 November and 13 February present clear-sky conditions.

The raw Landsat radiance data were aggregated across spatial increments of 120 m, from 120 m (the native resolution) up to the MODIS equivalent 960 m scale. Land surface temperature was calculated at each resolution from the aggregated radiance data. This last step is particularly important in estimating the Ts, as the relationship between radiance and Ts is non-linear. Aggregating the temperature directly (as opposed to averaging the radiances and then calculating the temperature), would subsequently increase the uncertainty in retrievals (McCabe et al., 2008). Vegetation structural parameters including vegetation height (h_v) (Su, 2001), leaf area index (LAI) (Ross, 1976), and fractional vegetation cover (f_c) (Campbell & Norman, 1998) are calculated from fine resolution NDVI (Normalized Difference Vegetation Index) and then aggregated to coarser resolution by simple averaging.

2.2.1. Land surface temperature, radiative fluxes and vegetation indices

Digital numbers in all bands of the Landsat images were converted to top of atmosphere radiance and reflectance values, and subsequently to land surface temperature following the methodology of Chander and Markham (2003) and Chander et al. (2007), after atmospheric correction using MODTRAN 5 software (Berk et al., 2008). For the atmospheric correction, temperature, and water vapour profiles were determined from the MODTRAN products of the MODIS sensor. For each of the aggregation scenarios, the upward longwave (LW_u) radiation was calculated using the spatially equivalent land surface temperature. In all cases, the emissivity and albedo were obtained from aggregated Landsat products. The downward longwave and shortwave radiation components were assumed uniform over the period of tower observations (corresponding to the Landsat overpass) which might have been affected by clouds. Therefore, meteorological variables are possibly uncoupled from the land surface conditions seen in the Landsat image of 6 September 2010, which results in uncertainties for evapotranspiration estimations based on this image. Images and tower records for 18 November and 13 February present clear-sky conditions.

2.3. Surface Energy Balance System (SEBS) model

SEBS (Su, 2002) is a physically based approach for the estimation of actual evapotranspiration using combined inputs from remote sensing and in-situ observations. The main forcing data to the SEBS model include the land surface temperature, vegetation height and density, air temperature, humidity, and wind speed. The principal element of the SEBS model is its robust formulation for estimation of the sensible heat flux using either Monin–Obukhov Similarity Theory (MOST) (Monin & Obukhov, 1945) for the atmospheric surface layer (ASL) domain or the Bulk Atmospheric Similarity Theory (BAST) (Brutsaert, 1999) for the mixed layer domain of the atmospheric boundary layer. In the majority of cases (and as employed here), MOST equations are used unless the roughness of the surface is high and/or the ASL is low. The MOST equations used in SEBS include stability-dependent flux-gradient functions for momentum and heat transfer according to

\[ u_g = \frac{u}{k} \left[ \ln \left( \frac{z-d_0}{z_{0m}} \right) - \Psi_m \left( \frac{z-d_0}{L} \right) + \Psi_h \left( \frac{z_{0m}}{L} \right) \right] \]  \hspace{1cm} (2)

\[ \theta_s - \theta_a = \frac{H}{\kappa u_p c_p} \left[ \ln \left( \frac{z-d_0}{z_{0h}} \right) - \Psi_h \left( \frac{z-d_0}{L} \right) + \Psi_h \left( \frac{z_{0h}}{L} \right) \right] \]  \hspace{1cm} (3)

\[ L = -\frac{\rho c_p \gamma \theta \theta_s}{\kappa g H} \]  \hspace{1cm} (4)

where \( z \) is the reference height above the land surface for measurement of the meteorological variables in metre, \( u \) is the friction velocity (m·s\(^{-1}\)), \( \rho \) is the density of the air (kg·m\(^{-3}\)), \( \kappa = 0.41 \) is the von Karman’s constant, \( c_p \) is specific heat capacity of air in J·kg\(^{-1}\)·K\(^{-1}\), \( L \) is the Obukhov length (m) with \( g \) the acceleration due to gravity (m·s\(^{-2}\)), \( d_0 \) is the zero-plane displacement height, \( z_{0m} \) is the roughness height for momentum transfer, and \( z_{0h} \) is the roughness height for heat transfer (in metre), \( \theta_s \) is the potential land surface temperature (K), \( \theta_a \) is the potential air temperature at height \( z \), and \( \theta \) is the atmospheric virtual potential temperature (K). Potential temperature (\( \theta \)) is calculated as:

\[ \theta = T \left( \frac{P_0}{P} \right) \frac{R_e \theta_s}{\rho q} \]  \hspace{1cm} (5)

where \( T \) is ambient temperature (for air or land surface in Kelvin), \( P \) is the corresponding atmospheric pressure (Pascal), \( P_0 \) is standard atmospheric pressure (101,325 Pa), and \( q \) is gas constant for dry air in J·kg\(^{-1}\)·K\(^{-1}\). The virtual potential temperature (\( \theta_v \)) is calculated as:

\[ \theta_v = (1 + 0.61q) \theta \]  \hspace{1cm} (6)

where \( q \) is the specific humidity in kg·m\(^{-3}\).

\( \Psi_m \) and \( \Psi_h \) are the stability correction functions for momentum and heat transfer. For stable conditions, the expressions proposed by Beljaars and Holtslag (1991) and evaluated by van den Hurk and Holtslag (1997) are used for atmospheric stability corrections in the atmospheric surface layer, and the functions proposed by Brutsaert (2005) are used for stability corrections in the mixed layer. Eqs. (2) to (4) can be solved in an iterative manner using the methodology given by Brutsaert (1982).

The roughness length for momentum transfer and heat transfer (\( z_{0m} \) and \( z_{0h} \)) are important parameters used in MOST and BAST equations, and are functions of the bio-meteorological conditions of the land surface. These two key parameters are estimated in SEBS using the methodology developed by Su et al. (2001), which employs vegetation phenology, air temperature, and wind speed. The roughness height for momentum transfer can be calculated as:

\[ z_{0m} = h_c \left( 1 - \frac{d_0}{h_c} \right) \exp \left( -\frac{K}{P} \right) \]  \hspace{1cm} (6)
where \( h_c \) is the vegetation height and \( \beta \) is the ratio of friction velocity to the wind speed at the canopy top calculated as \( \beta = c_1 - c_2 \exp\left(-c_3 LAI\right) \) with \( c_1 = 0.32 \), \( c_2 = 0.264 \), \( c_3 = 15.1 \), and the drag coefficient \( C_d = 0.2 \). The roughness length for heat transfer, \( z_0h \), can be derived by assuming an exponential relationship between \( z_0h \) and \( z_0m \) as \( z_0h = \frac{z_0m}{\exp(B - 1)} \) where \( B - 1 \) is the inverse Stanton number. To estimate the \( \kappa B^{-1} \) parameter, the extended model of Su et al. (2001) suggests:

\[
kB^{-1} = \frac{\kappa C_d}{4C_h^2(1-\exp(-\frac{z_0h}{C_h}))} j_c^2 + 2f j_f, \frac{\kappa C_d^2}{C_t} + \kappa B^{-1}j_f^2 \tag{7}
\]

where \( j_c \) is the fractional canopy coverage and \( j_f \) is its complement (for soil coverage), \( C_d \) is the heat transfer coefficient of the leaf, \( C_t \) is the heat transfer coefficient of the soil. As stated by Su (2002), the first term of Eq. (7) follows the full canopy only model of Choudhury and Monteith (1988), the third term is that of Brutsaert (1982) for a bare soil surface, and the second term describes the interaction between vegetation and a bare soil surface. Following Brutsaert (1999), for a bare soil surface the \( \kappa B_{-1} \) is calculated as \( \kappa B_{-1} = 2.46 Re^{-1/4} - \ln(7.4) \) with \( Re \) being the Reynolds number.

After estimation of the sensible heat flux, SEBS uses a scaling method to adjust the derived sensible heat flux between hypothetical dry and wet limits based on the relative evaporation concept. This scaled \( H \) is then used to derive the latent heat flux \( \lambda E \) as a residual term in the general energy balance equation. Further details on the formulation and implementation of the SEBS method are available from Su (2002), Su et al. (2005) and McCabe and Wood (2006). The flowchart of the key calculation steps in the SEBS model is presented in Fig. 3.

3. Methodology

In order to study the effects of aggregation on the distribution and magnitude of surface heat fluxes, SEBS model simulations were performed for two scenarios, comprising 1) input aggregation and 2) flux aggregation. Here, input aggregation first scales the remote sensing forcing data required by SEBS to the relevant resolution at which the heat flux calculations were performed (i.e. aggregate then calculate). To examine flux aggregation, SEBS latent heat flux retrievals were determined using the fine resolution Landsat derived data (in 120 m) and then aggregated to the subsequent resolutions (calculate then aggregate). The aggregation resolutions of this study were 120, 240, 480, 600, 720, 840, and 960 m, and the simple averaging method used for input aggregation. The 120 m resolution is the nominal resolution of the thermal channel (band 6) of the Landsat 5 TM, while the 960 m resolution is the closest integer multiplier to the nominal 1 km resolution of the MODIS daily land surface temperature products (including MOD11A1). The 240 m and 480 m resolutions provide an evaluation of the aggregation transfer effects between 120 m and 960 m and also approach the daily 250 m and 500 m MODIS visible band products which have been used in other approaches for disaggregation of \( T_s \) and flux data (Anderson et al., 2011). A flowchart of the simulation scenarios including the resolutions, interpolation methods, and the source of input data is presented in Fig. 3.

![Fig. 3. Flowchart of input aggregation (left) and flux aggregation (middle) scenarios. SEBS model flowchart is also shown in the right side. L. S. Param. refers to land surface parameters, including vegetation height, LAI, fractional vegetation cover, emissivity, and albedo derived from reflectance bands. Numbers in each box are spatial resolution in metres. For the SEBS flowchart, \( Q_n \) is available energy as \( Q_n = R_n - G_o \).](image-url)
For the input aggregation scenario, the aim is to evaluate the suitability of MODIS resolution land surface temperature for field scale evapotranspiration estimation. With the suspension of Landsat 5 (from November 2011), the only sources of fine resolution land surface temperature data are from the ETM + sensor on-board Landsat 7 and the ASTER sensor on-board Terra, both of which have limited capability in providing the required temporal resolution for many water resource applications. However, daily land surface temperature products can be obtained from MODIS and AVHRR sensors at coarse spatial resolutions, which can be integrated with vegetation parameters derived from optical sensors on-board a number of operational satellites (e.g. Landsat, SPOT, IRS) that provide fine spatial (25 m) but coarse temporal resolutions (16 to 22 days). In contrast to the land surface temperature which changes rapidly, the vegetation condition can be assumed relatively constant at weekly time scales. Hence, for the input aggregation of Landsat data in this study, radiance data (band 6) are directly aggregated to provide the required resolutions of \( T_s \), while reflectance data are used to first calculate the high resolution land surface parameters (e.g. vegetation height, \( LAI \)) and are then aggregated separately. Roughness parameters (\( d_a, z_{0m}, z_{0v} \)) are calculated at each distinct resolution using the aggregated vegetation parameters (\( h_v, LAI, f_v \)).

As noted for the case of input aggregation described above, a simple averaging method was used in up-scaling the higher resolution flux values to the MODIS scale. However, to examine the influence of the choice of interpolation routine on flux aggregation, nearest neighbour, bilinear and bicubic interpolation approaches were also examined. For nearest neighbour, the value of the aggregated pixel is the value from the fine resolution pixel that lies at the centroid of the coarse pixel. For bilinear interpolation, the aggregated pixel value is a weighted average of pixels in the nearest 2-by-2 neighbourhood. In the bicubic interpolation, the aggregated pixel value is a weighted average of pixels in the nearest 4-by-4 neighbourhood. The 1.8 × 1.8 km study area includes 90×90 pixels at 120 m resolution. To prevent edge effects in aggregation from 120 m to 480 m and 1.8×1.8 km study area includes 90×90 pixels at 120 m resolution. However, loss in spatial information is not spatially uniform at the larger pixel with the footprint of the larger pixel.

For near neighbour, the value of the aggregated pixel is the value from the fine resolution pixel that lies at the centroid of the coarse pixel. For bilinear interpolation, the aggregated pixel value is a weighted average of pixels in the nearest 2-by-2 neighbourhood. In the bicubic interpolation, the aggregated pixel value is a weighted average of pixels in the nearest 4-by-4 neighbourhood. The 1.8 × 1.8 km study area includes 90×90 pixels at 120 m resolution. To prevent edge effects in aggregation from 120 m to 480 m and 960 m resolutions, the last two rows and columns were ignored for these scales. As such, aggregation is performed for an 88×88 pixel region. Similarly, an 84×84 region is used for 840 m resolution aggregation. All 90 × 90 pixels are used for the other aggregated resolutions.

To allow for an evaluation of the aggregation effects in the input aggregation scenario, each aggregated product or variable (e.g. radiative and turbulent fluxes) is compared with its corresponding value at the native 120 m resolution. In some previous aggregation studies (e.g. Li et al., 2008), the relationship between increased pixel size and improved agreement with measured heat fluxes (i.e. perhaps as a response to “matching” the larger pixel with the footprint of the eddy covariance instrument) has been evaluated. The aim of this current study is to evaluate the errors and uncertainties in evapotranspiration estimation when the spatial resolution of the input variables and parameters increases, not to compare against in-situ measurements.

4. Results and discussion

For each of the selected days in September, November and February, the SEBS model was used to calculate latent and sensible heat fluxes for the aggregation scenarios identified in Fig. 3. Results for each scenario are presented and discussed in the following sections.

4.1. Input aggregation: effect of surface temperature and vegetation

Spatial maps, density plots, and statistical measures were used to assess the influence of input aggregation on flux retrievals. In the spatial maps and density plots, \( \Delta E \) is presented at 120, 240, 480 and 960 m resolutions. However, for statistical evaluation of the aggregation effects, additional resolutions of 360, 600, 720 and 840 m were also calculated. The statistical measures used for analysis of input aggregation include the spatial mean, relative error, root mean square difference (RMSD) and the coefficient of determination, \( R^2 \) (Kalma et al., 2008; Moore et al., 2009; Timmermans et al., 2007; Willmott, 1982).

Maps of land surface temperature and latent heat flux for each aggregated resolution are shown in Fig. 4. As the magnitude of both \( T_s \) and \( \Delta E \) in September is almost half that of the November and February images, a single consistent colour scheme is not used. Fig. 4 illustrates that there is significant spatial variability across all three days at the 120 m resolution retrievals. The spatial standard deviation of \( \Delta E \) is increased from 28 W m\(^{-2}\) in September to 41 and 54 W m\(^{-2}\) respectively in November and February. This variability is due in part to the agricultural practices and different phenological stages of crops in the study region. In particular, in the central west of the study area there are well defined agricultural fields that have low evapotranspiration rates in November, but which exhibit high values in the February image as a response to irrigation. Although the land surface temperatures in the majority of the November and February images are relatively similar, air temperature in November (24 °C) is higher than air temperature in the February image (20 °C). Therefore, for the same land surface temperature, the temperature gradient between the land surface and the atmosphere in November is 4 °C lower than February. This results in lower sensible heat fluxes in the November image. Moreover, net radiation in November (625 W/m\(^2\)) is higher than February (490 W/m\(^2\)) at the location of the meteorological tower. Consequently, the latent heat flux in the November image is higher than that of the February image.

In Fig. 4b, degradation in the spatial pattern, range, and magnitude of \( \Delta E \) is evident in the aggregated fluxes from 240 m to 960 m. While loss of some spatial detail is evident in the 240 m retrievals of \( T_s \) and \( \Delta E \), the range, magnitude, and spatial patterns of \( \Delta E \) are maintained. However, loss in spatial information is not spatially uniform at the 480 m resolution, and is a function of the size of the more strongly evaporating elements of the scene. At the 960 m pixel resolution, the range, magnitude, and spatial variability in both \( T_s \) and \( \Delta E \) images are noticeably reduced. Similar trends were observed by Li et al. (2008) in an aggregation study ranging across 30 m to 960 m over a semi-arid region.

4.1.1. Image scale errors due to the input aggregation

To provide a quantitative evaluation of the effect of input aggregation on evapotranspiration estimation, density and cumulative density plots of \( \Delta E \) maps for 120, 240, 480, and 960 m resolutions are presented in Fig. 5. For the density plots, the frequency of each interval is normalized by the area under the frequency curve.

As can be seen from Fig. 5, the range of \( \Delta E \) values varies with each aggregation resolution. Likewise, increases in the peak of the density plots with aggregation do not persist across all days. This is in contrast to observations in Hong et al. (2009), who found that the peak of the histogram of latent heat flux increased 10–25% as a response to aggregation. However, this difference might be attributed to the model structure difference between SEBS and SEBAL: in particular, SEBAL’s sensitivity to the choice of hot and cold pixel locations. Aggregation of input forcing shifts the peak of the density plot of the February image towards lower values of \( \Delta E \), which is apparent in the cumulative density plots, indicating that aggregation increased the frequency of lower \( \Delta E \) values. However, it is not generic and depends on the underlying land surface condition, the spatial interpolation method, and the heat flux model. For example, Gebremichael et al. (2010) found contrary results to the present study (i.e. a lower frequency of low \( \Delta E \) values), using a simple averaging aggregation of ASTER thermal images (90 m) to MODIS resolution, due possibly to: a) their methodology (SEBAL); b) differences in study area and eco-hydrological conditions of the surface; or c) different parameterization of the roughness parameters and aerodynamic resistance.
Fig. 4. Spatial maps of a) the land surface temperature and b) the resulting evapotranspiration from the SEBS model using a simple averaging approach.

For an evaluation of the effect of input aggregation on the magnitude and spatial variability of the latent heat flux and related parameters at the satellite image scale, the mean (μ) and standard deviation (σ) of simulated flux values were examined. In Fig. 6a, the relative spatial mean (μr) and relative spatial standard deviation (σr) values for key variables across the aggregated maps are shown. The relative spatial mean for each variable in each aggregated resolution is derived by dividing the spatial mean of the aggregated coarse resolution image by the spatial mean of the original fine resolution image. For example, for NE at 960 m resolution, the expression is μrNE = μNE(960)/μNE(120), where μNE is the spatial average of NE. Similarly, σrNE = σNE(960)/σNE(120), with σ being the spatial standard deviation.

The relative spatial mean and relative spatial standard deviation are derived for key input and flux variables of SEBS across all aggregated resolutions and plotted in Fig. 6. Variables analysed include the land surface temperature (T0), friction velocity (u*), aerodynamic resistance (ra), sensible heat flux (H), latent heat flux (NE), and available energy (Qn = Rn - G0). Here, the aerodynamic resistance (ra) is calculated as

\[ r_a = \frac{1}{k u_*} \left( \ln \left( \frac{z - d_0}{z_{th}} \right) - \psi_{z_0} \left( \frac{z - d_0}{L} \right) + \psi_{z_0} \left( \frac{z_{th}}{L} \right) \right). \]  

The aerodynamic resistance (ra) aids in characterising the roughness of the surface for momentum and heat transfer, and is thus useful in representing the combined effects of vegetation structure and the aerodynamic stability of the atmosphere (Brutsaert, 1982, 2005).

The plots in Fig. 6a illustrate that the relative spatial mean (μr) in the land surface temperature is constant, due to the simple averaging scheme used for aggregation. Relative spatial mean values for Qn are also constant and indicate that the aggregation of emissivity and albedo products does not affect the image scale spatial average of the available energy. However, relative spatial mean values for u* are increased while they are decreased for ra across aggregated resolutions. These responses are due to the combined effects of aggregation on the roughness properties of the surface. Consequently, decreases in the aerodynamic resistance cause an increase in the sensible heat flux in the September and February images. However, in the November image, H is decreased by a decrease in the aerodynamic resistance due to the SEBS algorithm, as it scales H between hypothetical dry and wet limits (Su, 2002). As a response to the change in the sensible heat flux variation, a decrease is evident for NE in the September and February images with an increase in the November image NE. However, the magnitude of change in H and NE is not the same for all selected days, and is related to the change in the roughness properties (z0m, z0h) of the land surface. For example, the February image shows greater sensitivity to the aggregation, resulting in a larger decrease in the relative mean of NE through aggregation.

In contrast to the relative spatial mean (μr), the relative spatial standard deviation (σr) for all selected variables and parameters is decreased through aggregation. While T0 and Qn have relatively similar σr for all days, the rate of decrease in σr for u* and ra is different in the September image than those of the November and February images. This response is related to the change in the spatial variability of roughness length parameters, which can be best evaluated at the pixel scale.

4.1.2. Pixel scale errors due to the input aggregation

To understand the effect of input aggregation at the pixel scale, each coarse resolution pixel is compared against the unaltered 120 m pixels located within it. For example, the land surface temperature value from
Fig. 5. Density (left) and cumulative density (right) plots of $\lambda E$ for three selected days from original 120 m data and aggregation of input forcing to coarser resolutions.

Fig. 6. (a) Variations of the relative spatial mean ($\mu_r$) and (b) relative spatial standard deviation ($\sigma_r$) for $T_s$ (°C), $u^*$ (m s$^{-1}$), $r_a$ (s m$^{-1}$), $H$ (W m$^{-2}$), $\lambda E$ (W m$^{-2}$) and $Q_n$ (W m$^{-2}$) across aggregated resolutions from 240 to 960 m, increasing by 120 m increments.
A 960 m pixel is compared to the 8 × 8 set of 120 m pixels from which it is comprised. To be able to make such comparison, a coarse 960 m resolution pixel can be considered as a set of 8 × 8 pixels all having the same value. The statistical measure for this comparison is the ‘relative error’ (or ‘estimation error’) defined as $E_r = \text{RMSD}/\mu$ (Kalma et al., 2008), where RMSD is the root-mean-square difference between the coarse resolution pixel and its constituent 120 m resolution pixels, and $\mu$ is the spatial mean value of those 120 m pixels. The relative error ($E_r$) is calculated for each coarse resolution pixel of the aggregated images for the key input and flux variables of the SEBS model across all aggregated resolutions, with the mean and percentiles (25th and 75th) of relative error ($E_r$) maps plotted in Fig. 7.

The plots in Fig. 7 illustrate that the pixel scale relative errors in the land surface temperature ($T_s$) and available energy ($Q_n$) are low (less than 5%). However, similar to the satellite image scale results, relative errors in $u_*$ and $r_o$ are higher (with wider percentile ranges) due to a response to the combined effects of aggregation errors on the roughness properties of the surface. Such a finding is in agreement with Moran et al. (1997) who identified negligible change in the land surface temperature aggregation, but large errors (greater than 50%) in the sensible heat flux over a heterogeneous study area.

As meteorological variables such as wind speed and air temperature are assumed constant for the study area (and hence for all pixels at all resolutions), only the land surface parameters derived from the Landsat image impact on the pixel scale spatial variability of $u_*$ and $r_o$. The friction velocity ($u_*$) is related to the roughness height for momentum transfer ($z_{200}$) and the instability of the atmosphere caused by such roughness. However, $r_o$ is related to both $z_{200}$ (via $u_*$) and $z_{50}$ (roughness height for heat transfer) and hence has more variability than $u_*$ across all resolutions in all three days. This influence is clearly apparent in the wider bounds of the 25th and 75th percentiles of $r_o$ compared to those of $u_*$.

To evaluate the effect of input aggregation on the spatial variability of relative errors, pixel scale $E_r$ maps of aerodynamic resistance ($r_o$) and latent heat flux ($AE$) are presented in Fig. 8. It is clear that the spatial distributions of errors in both $r_o$ and $AE$ are related to the variation of the land surface and may be associated with changes in the roughness height parameters. For example, in the 240 m error maps of Fig. 8a, pixels with high relative errors are linked with the location of the irrigation canals and the borders of agricultural fields where the land surface type (especially the roughness of the surface) changes at the pixel scale. However, by increasing pixel size, such scale effects reduce and the land surface type at the landscape scale influences the magnitude and distribution of the relative errors instead (e.g. for the border of drylands and irrigation areas).

Fig. 8b illustrates the presence of large pixel scale errors (greater than 40%) in $AE$ estimation at the 960 m resolution for February and September. The November images have relative errors of approximately 20% in agricultural areas at this same resolution. Differences in the relative errors of available energy ($Q_n$) and heat fluxes ($H$ and $AE$) at pixel scale highlight the important role of aerodynamic resistance parameterization in flux estimation, which is directly related to the estimation of roughness height parameters ($z_{200}$, $z_{50}$, $d_0$). Further research on the uncertainty analysis of roughness estimation at coarse scale resolutions is required to better characterise the degree of this influence.

It should be emphasized that the effect of aggregation of input forcing on the roughness height parameters (and subsequently to the evapotranspiration estimation) described here, will be general to those methods based upon the form of Monin–Obukhov Similarity theory equations as employed in SEBS. While they may also be pertinent to other flux estimation techniques involving these parameters, research by Allen et al. (2007b) and Long et al. (2011) has shown that roughness parameters play an insignificant role in METRIC and SEBAL (both are a form of energy balance approach). The main reason for this lies in the structure (formulation and parameterization) of METRIC and SEBAL, as these models use modified forms of flux-gradient functions (with simplifications and empiricism), resulting in their different response to the scaling of roughness parameters.

### 4.2. Effects of flux aggregation approach

Aggregation of fluxes from fine to coarse resolutions is a common practice in regional to global climate model evaluation (Jiménez et al., 2011; Mueller et al., 2011) and in the assessment of coarser scale flux products such as those derived from geostationary satellites (Brunsell & Anderson, 2011). As such, it is important to understand how aggregated fluxes differ from the fluxes at the native resolution in terms of their statistical structure, magnitude, and spatial distribution. There are a number of commonly used spatial interpolation methods that can be employed for such aggregation and it is likewise important to identify their effects on preserving the spatial characteristics of the original fine resolution fluxes. To better understand whether the choice of aggregation technique has an effect on evapotranspiration estimates, ET (or the latent heat flux, $AE$) was calculated using the original Landsat data at 120 m resolution and then aggregated to 960 m resolution using the simple averaging, nearest neighbour, bilinear, and bicubic interpolation methods, all of which are common approaches in spatial interpolation.

Fig. 9 presents the response of evapotranspiration aggregation using these different techniques, showing that the spatial details present in the fine resolution evapotranspiration maps decrease dramatically by the 960 m resolution. The nearest neighbour (NN) approach causes sharp discrepancies in the flux values, while the SA, BL, and BC produce a smoother transition between pixel responses. In contrast to the NN aggregation, the visual difference

![Fig. 7](image-url) The relative error ($E_r$) at the pixel scale for $T_s$ (°C), $u_*$ (m s$^{-1}$), $r_o$ (s m$^{-1}$), $H$ (W m$^{-2}$), $AE$ (W m$^{-2}$) and $Q_n$ (W m$^{-2}$) across aggregated resolutions from 240 to 960 m, increasing by 120 m increments. Triangles represent the mean relative errors and lines above and below identify the 75th and 25th percentiles respectively for each aggregated resolution.
between SA, BL and BC is not significant, with all exhibiting similar spatial patterns. In terms of statistical metrics, all aggregation methods yield a similar image scale mean (mean of all pixels of the image) compared to the fine resolution evapotranspiration images. However, for image scale standard deviation (as a measure of spatial variability), NN aggregated images show an improved match against their corresponding fine resolution image, which is statistically significant and in agreement with previous research results (e.g. Hong et al., 2009; Sridhar et al., 2003), but is not significant from a hydrological perspective (as shown below).

In order to evaluate the uncertainties associated with flux aggregation using these different approaches, values of evapotranspiration in W/m² are converted to volumetric evapotranspiration, and errors due to the aggregation are calculated for each coarse resolution pixel as \( ET_{960} - ET_{120} \). As can be seen in Fig. 10, the SA approach produces no errors, but underestimation and overestimation errors in volumetric evapotranspiration are evident in NN method maps for all days. In contrast to NN, volumetric evapotranspiration errors for BL and BC are lower.

When undertaking regional scale analysis or water balance estimation, the pixel scale volumetric evapotranspiration errors can accumulate and potentially cause large mass imbalances in hydrological studies. To evaluate this, total volumetric evapotranspiration errors have been calculated for the study area as shown on each image of Fig. 10. It is clear that the SA method has the best performance in preserving the mass balance. The NN method results in a significant underestimation of evapotranspiration at the image scale for all Landsat images. Although BL and BC produced lower errors than the NN approach, they still result in an imbalance and hence are not suitable for flux aggregation. From a hydrological perspective, the simple averaging approach is the preferred technique for flux aggregation.

5. Summary and conclusion

Understanding the effects of aggregation on the estimation of hydrological variables is of considerable importance, especially in
relation to the accurate retrieval of land surface fluxes from remote sensing observations. The availability of remote sensing images from fortnightly to sub-daily temporal resolutions and from metres to kilometre spatial resolutions, provides a great opportunity for operational assessment and management of water resources. As fine resolution imagery has shorter temporal resolutions (e.g. fortnightly) and often limited availability due to atmospheric influences, coarser resolution images from MODIS type sensors provide greater utility to the water resource community. Therefore, it is crucial to understand the implications of coarse resolution retrieval of heat fluxes relative to higher resolution responses.

To examine the influences of spatial scale on remotely sensed land surface heat flux estimation, an evaluation of the aggregation effects on a temporal sequence of high resolution Landsat 5 TM images was performed. The scaling effect on simulations was examined by a) aggregating the key input forcing of surface temperature and vegetation, and b) assessing the influence of the flux aggregation approach on flux retrieval. It was determined that the influence of input forcing aggregation resulted in the underestimation of evapotranspiration at the satellite image scale, with up to 15% lower retrievals than it occurred at the original high resolution Landsat image. It was reasoned that the most likely explanation for this response was an increase in the aerodynamic resistance, originating from a change in the roughness height estimation across aggregated resolutions. However, comparison with similar studies suggested that the significance of input aggregation on roughness parameterization (and subsequently on evapotranspiration estimation) may be specific to the SEBS model.

Further work examining other model structures and types is clearly required, given the influence of these parameterisations shown here. Results also show that in aggregating fine resolution fluxes to coarser scales, a simple averaging scheme outperforms other common approaches by preserving both the spatial distribution of evapotranspiration and the magnitude of volumetric evapotranspiration at the pixel and image scales. In contrast, a nearest neighbour method for flux aggregation can cause large errors.

While this study was limited to the spatial resolution of MODIS thermal data, coarser resolution (but high temporal response) geostationary satellite data could also be investigated, and the study area expanded to include basin and regional scale responses. Such an analysis would provide a more comprehensive spatio-temporal scaling scheme for heat flux simulations than that undertaken here (although it should be noted that this study is one of a few that attempts to examine the temporal scaling response by including multiple Landsat images across a changing land surface condition). Another issue that requires further consideration, relates to the effects of aggregation on the estimation of vegetation structure parameters (e.g. leaf area index) and subsequently on roughness parameterization (e.g. 20m, 200m). Aggregation of such data could have considerable impact on the so-called T—I VI family of evapotranspiration models (Carlson, 2007; Long et al., 2011; Petropoulos et al., 2009). Spatial scaling has the capacity to alter the geometry of the scatterplot space between the land surface temperature and vegetation indices, which would affect the geometry of dry and wet edges and consequently the resulting evapotranspiration. Likewise, by expanding this analysis to encompass a greater range of surface types, conditions, and resolutions, a generalization of the results from this study to other hydrometeorological conditions and ecosystems may be made.

Acknowledgement

Funding of this research is provided jointly by an Australian Research Council (ARC) Linkage project (LP0989441) and scholarship support for Ali Ershadi from the National Centre for Groundwater Research and Training (NGCRT) in Australia. Authors appreciate free access to Landsat and MODIS images used in this research from the USGS (through EarthExplorer — earthexplorer.usgs.gov) and NASA (through LAADS-Web — ladsweb.nascom.nasa.gov). Mr. Robert Pipunic from the University of Melbourne is thanked for providing the required meteorological data used in this study.

References


Fig. 10. Errors in estimation of the total mass of evapotranspiration (in m3) due to the aggregation of the latent heat flux for different spatial interpolation methods. Negative values mean aggregated pixels are less than the original fine resolution values. Numbers above the images for NN, BL, and BC are spatial sum.


