



## Development of an automated method for mapping fire history captured in Landsat TM and ETM + time series across Queensland, Australia

Nicholas R. Goodwin <sup>\*</sup>, Lisa J. Collett

Department of Science, Information Technology, Innovation & the Arts, Remote Sensing Centre, Ecosciences Precinct, GPO Box 2454, Brisbane, QLD 4001, Australia



### ARTICLE INFO

#### Article history:

Received 14 October 2013

Received in revised form 24 March 2014

Accepted 25 March 2014

Available online 22 April 2014

#### Keywords:

Burnt area mapping

Fire history

Landsat

Time series

### ABSTRACT

Remote sensing can quantify past and present fire activity at spatial scales useful for a range of fire and vegetation management applications. In this study, we present a new automated approach to classifying burnt areas across the state of Queensland, Australia. The method is applied to complete time series of Landsat TM/ETM + imagery rather than single images and considers spectral (band 4, B4, and bands 4 + 5, B45), thermal, temporal and contextual information within a hierarchical framework. To maximise the available observations and the burnt area detected, we used imagery containing up to 60% cloud that was screened during pre-processing. Median filters were applied to smooth the time series and multi-date change detection used to locate negative outliers (large declines in reflectance relative to the median-smoothed time series). Watershed region growing was used to segment and map a larger spatial extent of the change while minimising commission errors. These segmented change objects were attributed as either burnt or unburnt using their thermal, reflective and contextual characteristics in a classification tree. Thermal information was found to be more important than reflective indices in the change attribution. Algorithm calibration used training data from ten Path/Rows located strategically across Queensland with four images sampled per path row ( $n = 40$ ). Thresholds were optimised to maximise the burnt area detected while limiting under/over-growing of burnt area. Validation data covered a range of burnt areas from ten independent Path/Rows with ten images sampled across a range of burnt area fractions per Path/Row ( $n = 100$ ). The results for burnt area mapping demonstrated an average producer's accuracy of 85% (range of 28 to 100% for individual images) and average user's accuracy of 71% (range of 4 to 99% for individual images). A morphological dilation of one pixel restricted to locations exhibiting a decline in B45 over time, increased the producer's accuracy by 4% but reduced the user's accuracy by 8%. The total accuracy for the burnt area classification was greater than 99%, however this is more a reflection of the small fraction of landscape represented by burnt area rather than the ability to detect burnt area. Areas frequently misclassified were related to areas of high spectral/land use change which included areas of cropping, frequently inundated land, and moisture/ground cover variations over dark soils. In this study, we applied a crop and water mask to minimise commission errors. Significantly, the results of this study demonstrate that an automated time series method for mapping burnt areas can be successfully applied across a diversity of land cover types. The method may be applied in similar savanna dominated environments but is likely to require modification to be applicable in other landscapes.

© 2014 Elsevier Inc. All rights reserved.

### 1. Introduction

Understanding and characterising fire history are important for improving our knowledge and management of fire, climate, land-use and vegetation interactions. In Australia, for example, fire has been a key process shaping vegetation dynamics and ecosystem function (Russell-Smith et al., 2003). Furthermore, it has been demonstrated that significant changes to a region's fire regime (i.e. spatial extent, timing, patchiness, frequency and intensity of fire events) over time

can result in a loss of biodiversity (Keith, Williams, & Woinarski, 2002), changes to species composition (Orr, McKeon, & Day, 1991) and vegetation structure (Burrows et al., 2002; Crowley & Garnett, 1998; Vigilante & Bowman, 2004), as well as having wide ranging implications for carbon dynamics (Henry, Danaher, McKeon, & Burrows, 2002; Williams, Griffiths, & Allen, 2002). Wildfires also have a range of socioeconomic impacts including damage to property, infrastructure and livestock, pollution of air and water, increased risk of soil erosion and run-off and in extreme cases may result in the loss of human life. Consequently, there is a requirement to quantify past and present fire patterns. This will assist the development of appropriate fire management practices and benefit a range of conservation and resource management objectives, as well as ongoing scientific research.

\* Corresponding author.

E-mail address: nicholas.goodwin@science.dsiti.qld.gov.au (N.R. Goodwin).

The Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper (ETM+) archive provides a frequent sample of land surface reflectance from ~1984 to present at a spatial resolution (~30 m) useful for characterising burnt area at local and regional scales. Until recent times, the use of Landsat for mapping historical patterns of burning has been limited, principally by imagery costs, to localised studies involving largely manual or semi-automated mapping (Edwards et al., 2001; Felderhof & Gillieson, 2006; Russell-Smith, Ryan, & Durieu, 1997; Russell-Smith et al., 2003). Large-scale, satellite-based burnt area mapping has been mostly derived from coarse spatial resolution imagery from sensors such as the Moderate Resolution Imaging Spectroradiometer (MODIS) (Justice et al., 2002; Roy et al., 2008), available since December 1999 or the Advanced Very High Resolution Sensor (AVHRR), (Craig et al., 2002). With the release of the Landsat archive by the United States Geological Survey (USGS), and advances in digital data storage and computing technology, mapping at the Landsat scale over large regions back to the mid-1980s is now possible.

There has been considerable research undertaken to develop spectral indices from Landsat to discriminate burnt vegetation from other land surfaces. These include the Normalised Burn Ratio (NBR, normalised difference of TM bands 4 and 7) and its multi-temporal derivatives (Boer et al., 2008; Key & Benson, 1999; Miller & Thode, 2007; Miller & Yool, 2002), the ratio of TM bands 4 and 5, the Leaf Area Index (Boer et al., 2008) and the Burnt Area Index (BAI) (Bastarrika, Chuvieco, & Martín, 2011; Masek et al., 2008; Stroppiana et al., 2012). However, many of these indices have been developed for forested areas and are not suitable for all environments, in particular, savannas. The widely adopted NBR, for example, is less sensitive to fire related change in savanna environments (Disney et al., 2011; Maier, 2010). This is because there is a general decrease in reflectance in both spectral bands in response to fire rather than the characteristic decrease in Landsat TM band 4 and increase in band 7 (Key & Benson, 1999) which is seen following fires in forested areas. This was demonstrated by a series of 3D model simulations developed by Disney et al. (2011) in a 2-layer savanna system that concluded that burn events are likely to cause a decline in reflectance rather than a diagnostic change in the spectral shape.

Landscapes are dynamic and are often characterised by high spatial and temporal heterogeneity in land surface reflectance. A range of factors may influence land surface reflectance including changes in vegetation structure, species composition and dynamics, exposed soil fractions and reflectance characteristics, soil moisture and inundation, land-use, cloud and cloud shadow, and burn characteristics (fraction of white and black ash relative to photosynthetic and non-photosynthetic vegetation as well as soil fractions). In general, large changes in vegetation cover and high fractions of char due to fire, followed by a moderate to long period of recovery (>1 yr), are more likely to achieve high detection rates. In contrast, fires of low severity with small fractions of char and patches of unburnt vegetation can be more challenging to detect in remotely sensed imagery. Post-fire climatic effects such as rainfall and wind can also impact the ability to detect burnt areas. Given the wide range of burn conditions there is generally a trade-off between omission (under-estimation) and commission (over-estimation) rates often determined by localised calibration and specific product needs. This has made the development of automated approaches difficult as methods may not be transferable (without re-calibration) to other locations or potentially over time.

To improve the balance between omission and commission errors, Bastarrika et al. (2011) implemented an automated two-phase approach to mapping burnt areas using Landsat TM/ETM+ imagery. The first phase identifies 'core burned' pixels within a burnt patch, using a series of single date and multi-temporal decision rules based on known metrics. The second phase uses a region growing approach to map pixels in the immediate neighbourhood of the 'core' burnt pixels, which may exhibit less apparent changes due to fire. The two-phase method was calibrated with a large multi-region calibration set and

found to be robust when applied across Portugal and California, USA in 2003 ( $\kappa = 0.85$ ). Stroppiana et al. (2012) also developed an approach using multiple metrics and region growing to detect burnt areas across the Mediterranean with encouraging results for a validated Landsat scene (omission and commission errors of 3% and 21%, respectively). Furthermore, these findings are consistent with other studies that coupled image differencing and region grow operations to screen cloud and cloud shadow (Derrien & Le Gleau, 2010; Goodwin, Collett, Denham, Flood, & Tindall, 2013).

The use of multi-temporal information to detect burnt areas in Landsat imagery has largely been limited to two date comparisons. In particular, differencing NBR derived from pre- and post-fire image dates has been widely used (Key & Benson, 1999) and is central to the Monitoring Trends in Burn Severity (MTBS) project that is mapping historic burnt areas across the United States (Eidenshink et al., 2007). Yet, considerable advances have been made in characterising trends, disturbances and/or temporal anomalies using the time series information contained in the Landsat archive. Kennedy, Cohen, and Schroeder (2007) segmented spectral trajectories to automatically detect trends and disturbances such as forest clearing or partial harvests in reflectance time series. Huang et al. (2009) developed the vegetation change tracker (VCT) method to automatically detect disturbances in forest environments. Röder et al. (2008) applied spectral mixture analysis to examine trends and changes in the green cover fraction post-fire. Goodwin et al. (2013) used an automated algorithm to screen cloud and cloud shadow using complete time series of Landsat imagery. There is potential for time series approaches to improve the characterisation of burnt areas and the automation of image processing.

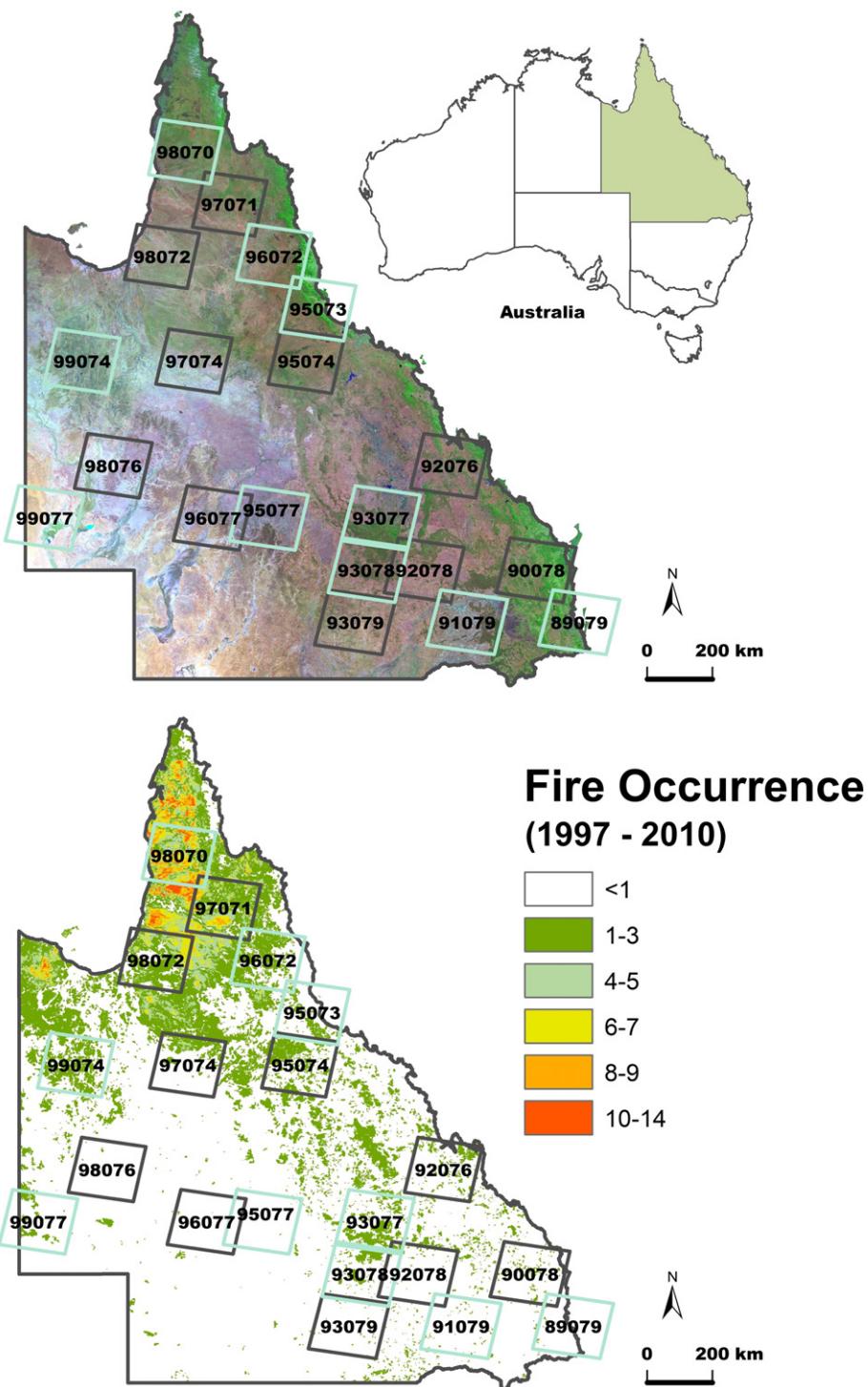
In this study, we describe and validate a method for operational burnt area mapping in Queensland, Australia. The method uses Landsat reflectance time series together with contextual information contained in single-date reflective and thermal bands. The method builds on our recent cloud and cloud shadow screening framework (Goodwin et al., 2013) which includes: i) time series change detection to identify outliers relative to the time series which may be due to burnt vegetation; ii) region growing to map the spatial extent of the detected change, and iii) an object-scale classification tree to distinguish burnt area from non-fire related change regions. The method has been developed with the objective of deriving fire history using the complete Landsat TM/ETM+ archive as well as operational burnt area mapping into the future, across the state of Queensland, an area covered by 100 Landsat WRS footprints (at present >60,000 images). The benefits and limitations of an automated approach are also discussed.

## 2. Methods

### 2.1. Study area

Queensland is the second largest State of Australia with an area of 1.7 million km<sup>2</sup> and a broad range of climate zones, topography, vegetation communities, geological landforms and soils (Fig. 1). Approximately 80% of Queensland comprises rangeland savannas and grasslands which are extensively grazed by livestock, with less than 2% of the state under cropping (EHP, 2011). Tropical and subtropical rainforest occurs along the coast, with tracts of dry forest and woodlands occurring in the central and south eastern regions. Queensland's rainfall is highly variable and strongly seasonal with the majority of rain falling in the summer months (December–February). There is significant inter-annual and inter-decadal variability due in part to the El Niño Southern Oscillation, the Inter-decadal Pacific Oscillation (Klingaman, 2012). There is also a strong north–south rainfall gradient with average annual rainfall ranging from 3000 mm in the north-eastern wet tropical region, to less than 200 mm in the south-west semi-arid grasslands (BOM, 2013).

Fire (and its exclusion) is a driving factor in ecosystem, function and productivity in northern Australian landscapes (Dyer & Smith, 2003;



**Fig. 1.** The study region showing the twenty Landsat Path/Rows sampled across Queensland for calibration (shown in light blue) and validation (shown in grey) overlaid upon: (a) Landsat TM composite of Queensland, Australia (RGB: 542) and (b) estimated fire occurrence derived from NOAA AVHRR imagery (1997–2010) (source: NAFI, 2014; [www.firenorth.org.au/nafi2/](http://www.firenorth.org.au/nafi2/)).

Russell-Smith, Whitehead, Cooke, & Yates, 2009; Russell-Smith et al., 2003). However, there are large regional variations in fire regimes and fire management, primarily driven by rainfall patterns, but also land use, resource management practices and cultural attitudes towards the prescriptive use of fire. Across the tropical savannas of north Queensland, monsoonal summer rainfall drives an annual accumulation of large quantities of fine fuel ( $2\text{--}5 \text{ t ha}^{-1}$  grass and tree litter annually; Williams et al., 2002). In the absence of proactive management, fires in these regions are spatially extensive, strongly seasonal (predominantly

late dry season; August–November) and very frequent, where grazing is limited (Maier & Russell-Smith, 2012; Russell-Smith et al., 2003). Prescribed burning in this region is undertaken in the early dry season (April–June) to reduce fuel, increase landscape heterogeneity, and achieve a range of conservation and resource management objectives (Dyer & Smith, 2003; Russell-Smith et al., 2003). Early dry season fires are generally smaller and patchier due to higher fuel moisture and humidity. In contrast, fire frequencies in the drier rangelands which cover most of the state, are much lower and highly correlated with

inter-annual rainfall variability (Allan & Southgate, 2002; Felderhof & Gillieson, 2006).

Twenty regions, represented in each case by a Landsat World Reference (WRS) Path/Row, were selected to capture the diversity in fire regime, land-use, biogeographic and climatic variation. To show the distribution of land surface types and fire regimes sampled, these twenty Path/Rows are overlaid on a statewide Landsat composite image and map of fire occurrence derived from AVHRR imagery between 1997 and 2010 (Fig. 1a and b).

## 2.2. Image pre-processing

Time series of all available ortho-corrected Landsat TM and ETM+ images (1986–2013), including SLC-off and cloud affected imagery, were acquired from the USGS for ten calibration and ten validation Path/Rows (Fig. 1). This incorporated between 5 and 41 images per year with an average of 21 images per year (TM and ETM+), while the total number per Path/Row ranged between 428 and 608 image dates. Imagery was converted from radiance to a standardised surface reflectance following Flood, Danaher, Gill, and Gillingham (2013). This correction method has been parameterised for eastern Australia and corrects for atmospheric, topographic and bi-directional effects. The Second Simulation of the Satellite Signal in the Solar Spectrum (known as 6S; Vermote, Tanre, Deuze, Herman, & Morcrette, 1997) is used to correct for atmospheric effects, while the Bi-Directional Reflectance Distribution Function (BRDF) effects were modelled using overlapping image pairs to produce standardised reflectance (i.e. reflectance illuminated and observed from a single direction). Atmospheric effects due to high aerosol loadings are however, likely to remain (Flood et al., 2013) as no operational method is at present available for accurately estimating aerosol loading/haze/smoke over much of Australia (Gillingham, Flood, Gill, & Mitchell, 2012; Mitchell, Qin, & Campbell, 2009). The thermal infrared band 6 (10.40–12.50 μm) was converted to radiance and then top-of-atmosphere brightness temperature ( $B_T$ ; Chander, Markham, & Helder, 2009) and resampled to 30 m × 30 m pixels to match reflectance. No atmospheric correction for the thermal band was applied. Hereafter, “reflectance” refers to surface reflectance.

Imagery containing up to 60% cloud cover were included in the time series as frequently the only observation of a burnt area is captured in cloud affected imagery. This also maximises the number of land surface observations included in the time series analysis. Across our sampled Path/Rows, over 70% of the imagery contained less than 20% cloud cover, while 16% and 12% of imagery contained cloud cover in the ranges 20–40% and 40–60%, respectively. Cloud and cloud shadow affected pixels were screened following Goodwin et al. (2013). This method involves locating outliers in the time series of reflectance along with region growing and object-based classification.

## 2.3. Framework for burnt area mapping

Mapping burnt area requires processing an entire Path/Row at a time with imagery ordered in temporal sequence. This involves three key processing stages detailed in the following sections.

### 2.3.1. Detection of time series outliers

The first step in our burnt area detection is to identify pixels which are negative outliers relative to their observed reflectance over time. This approach was taken as we found that the most characteristic feature of a burn event across a large range of vegetation types is a post-fire decline in reflectance. This is consistent with earlier studies (Disney et al., 2011; Maier, 2010; Stroppiana, Pinnock, Pereira, & Grégoire, 2002). Post-fire declines in reflectance were evident across all reflective bands but were greatest in Landsat TM bands 4 ( $B_4$ , near infrared, 0.76–0.90 μm) and 5 ( $B_5$ , mid-infrared, 1.55–1.75 μm) due to their large dynamic range and sensitivity to fire-related change

(i.e. loss of photosynthetic vegetation and increased char/ash). However, the relative post-fire behaviour of  $B_4$  and  $B_5$  varies across different environments. For example, there may be little or no observed decline in  $B_4$  after a savanna fire, where the largest decline occurs in  $B_5$ , and conversely, little or no observed change in  $B_5$  observed after a forest fire. Normalised difference or direct band ratios using  $B_4$ ,  $B_5$  or other bands, were found to be ineffective as they rely on a consistent change in spectral shape in response to fire, e.g. a large decline or increase in one band relative to another. As such, we used the summed reflectance in both bands,  $B_4 + B_5$  ( $B_{45}$ ) to identify potentially burnt pixel outliers in reflectance. To further separate fire from non-fire  $B_{45}$  outliers, the time series of  $B_4$  was also considered. This was used to eliminate pixels showing a  $B_4$  increase (notwithstanding the  $B_{45}$  decline), associated with changes in vegetation vigour or moisture content. The time series of  $B_5$  alone was found to add little to the time series detection of potentially burnt pixel outliers given the additional computational overhead.

To detect negative outliers relative to the time series, a reference of ‘no change’ is required. We applied two running median filters to both time series of  $B_4$  and  $B_{45}$ :

- (i) a preceding median using the 7 nearest image dates captured before the target image ( $B_{45PM}$ ;  $B_{4PM}$ ; Eqs. 1 and 2). This filter is best suited for pixels exhibiting long-term non-seasonal trends. We found a sample of 7 image dates provided a good balance between characterising the variation in unburnt reflectance and limiting the inclusion of pixels affected by past burn events. Across our calibration/validation data set, the sample periods for the preceding median ranged from 56 to 308 days with an average period of 125 days.

$$B_{45PM} = \text{median}(B_{45_{t-7}}, B_{45_{t-6}}, \dots, B_{45_{t-1}}) \quad (1)$$

$$B_{4PM} = \text{median}(B_{4_{t-7}}, B_{4_{t-6}}, \dots, B_{4_{t-1}}) \quad (2)$$

- (ii) a seasonal median ( $B_{4SM}$ ;  $B_{45SM}$ ) derived from images captured under comparable growing conditions (inferred from solar geometry), and from within a particular time interval around the target image. This is better suited for pixels exhibiting highly periodic, seasonal trends. Image sampling is based on a progressive search, targeting imagery captured under solar azimuth and zenith angles within  $+/- 3$  to  $15^\circ$  ( $1^\circ$  search increments) and  $+/- 1$  to 5 years (1 year search increments) of the target date. The incremental search continues with both geometry and date criteria incremented at each iteration of the search until at least 4 image dates have been sampled. Images captured up to 60 days after the target date were excluded to limit inclusion of burnt pixels which may persist in subsequent images after the first capture of a burn event. All other images meeting the date and geometry criteria are included in the median calculation. A second pass of the median filter is used to eliminate spurious results and further smooth the time series. Experimenting without the 60 day exclusion period had very little effect (<1%) on detection rates, however, this may need further investigation for adoption in regions outside the study area where fire scars persist for longer in the landscape.

Despite efforts to limit their inclusion, it is conceivable that burnt pixels may contaminate preceding or seasonal medians due to highly frequent fire events, data gaps (e.g. cloud) or burnt areas which persist in imagery for long periods of time. However, in both the preceding and seasonal sampling strategies, the median is robust against selection of newly burnt pixels, which tend to represent localised time series minima rather than median values. It follows that even in cases where the median selects a pixel capturing some degree of post-fire change in reflectance, this would not necessarily preclude the detection of a newly burnt pixel.

For each pixel in a given image, the  $B45$  reference ( $B45_{ref}$ ) was selected from one of the two medians ( $B45_{SM}$ ,  $B45_{PM}$ ), depending on how well it characterised the observed reflectance temporal curve. This was quantified by calculating the average positive residuals in  $B45$  over time ( $t$ ) from each of the medians, as follows:

We define the set of time series residuals,  $R_{PM}$ , from  $B45_{PM}$ :

$$R_{PM} = (B45_t - B45_{PM}) \quad (3)$$

$R_{PM}^P = R_{PM} | R_{PM} > 0$  is the set of positive residuals

$$\bar{R}_{PM}^P = \text{average of } R_{PM}^P$$

We define the set of time series residuals,  $R_{SM}$ , from  $B45_{SM}$ :

$$R_{SM} = (B45_t - B45_{SM}) \quad (4)$$

$R_{SM}^P = R_{SM} | R_{SM} > 0$  is the set of positive residuals

$$\bar{R}_{SM}^P = \text{average of } R_{SM}^P$$

In these calculations only positive residuals were considered to limit the inclusion of large negative residuals resulting from changes due to fire or missed cloud shadow in the time series. It was assumed that the median recording the lowest average residual,  $\min(\bar{R}_{SM}^P, \bar{R}_{PM}^P)$ , over time was most representative of the observed temporal reflectance curve and was selected as  $B45_{ref}$  for that pixel.  $B4_{ref}$  was then selected using the same median filter. Where pixels exhibited strong seasonality in reflectance over time,  $B45_{SM}$  was more likely to produce the lowest average residual whereas pixels exhibiting non-seasonal variation may be better represented by  $B45_{PM}$ .

Visual inspection of the average residuals showed that the spatial distribution of pixels characterised by high average residuals was not random. In general, land cover types such as crops, ephemeral water bodies and wetlands were well delineated in the average residuals due to their comparatively high residuals. These features can exhibit frequent changes in reflectance of the same or greater magnitude as those due to fire, which can result in significant increases of commission rates in burnt area classifications (Bastarrika et al., 2011). To limit the impact of these areas, a set of state-wide crop and water masks were derived from the most recent Queensland Land Use Mapping Project data (<http://dds.information.qld.gov.au/DDS/SearchResult.aspx>).

Difference images,  $resid_{B45t}$  and  $resid_{B4t}$  (Eqs. 5 and 6) were derived for each date of imagery in a Path/Row image stack:

$$resid_{B45t} = B45_{ref} - B45_t \quad (5)$$

$$resid_{B4t} = B4_{ref} - B4_t. \quad (6)$$

The set of negative time series outliers, corresponding to potential core burnt area pixels was then defined by:

$$(resid_{B45t} > Th_S) \text{ AND } (resid_{B4t} > Th_{B4}) \quad (7)$$

where:  $Th_S$  ( $B45$  seed threshold) and  $Th_{B4}$  ( $B4$  change threshold) values were determined by optimisation (refer Section 2.4).

Examples of reflectance curves ( $B4$  and  $B45$ ), medians ( $B45_{SM}$ ,  $B45_{PM}$ ,  $B4_{SM}$ ,  $B4_{PM}$ ), residuals ( $resid_{B45}$ ,  $resid_{B4}$ ) are shown in Fig. 2 for two pixel time series (Path/Row 089/079). Change thresholds delineating negative outliers,  $Th_S$ , and  $Th_{B4}$ , are also shown. In the first example the seasonal median was selected as the 'no change' reference for a pasture pixel due to the seasonal variation in reflectance (Fig. 2a and b), and in the second, the preceding median was selected for a eucalypt woodland pixel characterised by non-seasonal effects (Fig. 2c and d). Three observations can be made: (i) gaps due to cloud and other data loss are evident, leading to difficulties in quantifying recovery times for surface changes; (ii) the temporal behaviour in response to burnt area and cloud shadow can be similar in  $resid_{B45}$  and  $resid_{B4}$  (Fig. 2b), both in terms of the magnitude of change, and the recovery time. In

these cases, further information is required for correct attribution; and (iii) discrimination between burnt and unburnt change is possible where there is a large change in  $B45$  but only a small or negligible change or an increase in  $B4$ , as seen in Fig. 2d. In this example the burnt and unburnt changes (which are due to changes in soil moisture and cover) are separable in  $resid_{B4t}$ , as the unburnt changes do not exceed the threshold,  $Th_{B4}$ .

### 2.3.2. Region growing of change regions

Spatial clusters of negative outliers (Eq. 7) were used as 'seeds' in a watershed region growing operation to map the full extent of the change region. In a watershed filter, the difference image,  $resid_{B45t}$  may conceptually be seen as a topographic relief, where the magnitude of change may be interpreted as its elevation. A drop of water will flow down the relief until reaching a local minimum, or basin, with watersheds, defining the boundaries between adjacent catchment basins (Vincent & Soille, 1991). Goodwin et al. (2013) used a flood-filling watershed filter from the VIGRA Computer Vision Library (University of Heidelberg, 2012; <http://hci.iwr.uni-heidelberg.de/vigra>) for mapping clouds and their shadows from Landsat time series. This approach used a fixed interval of change, which was determined by optimisation, to delineate regions of cloud/shadow. We have adopted the same VIGRA watershed filter, however, instead of using a fixed interval, two sets of seeds were used: the 'change' seeds which were the set of outliers from Eq. (7); and a set of unburnt seeds (containing no burnt pixels; Eq. 8):

$$resid_{B45t} < = Th_B \quad (8)$$

where:  $Th_B$  is the  $B45$  unburnt class threshold determined by optimisation (refer Section 2.4).

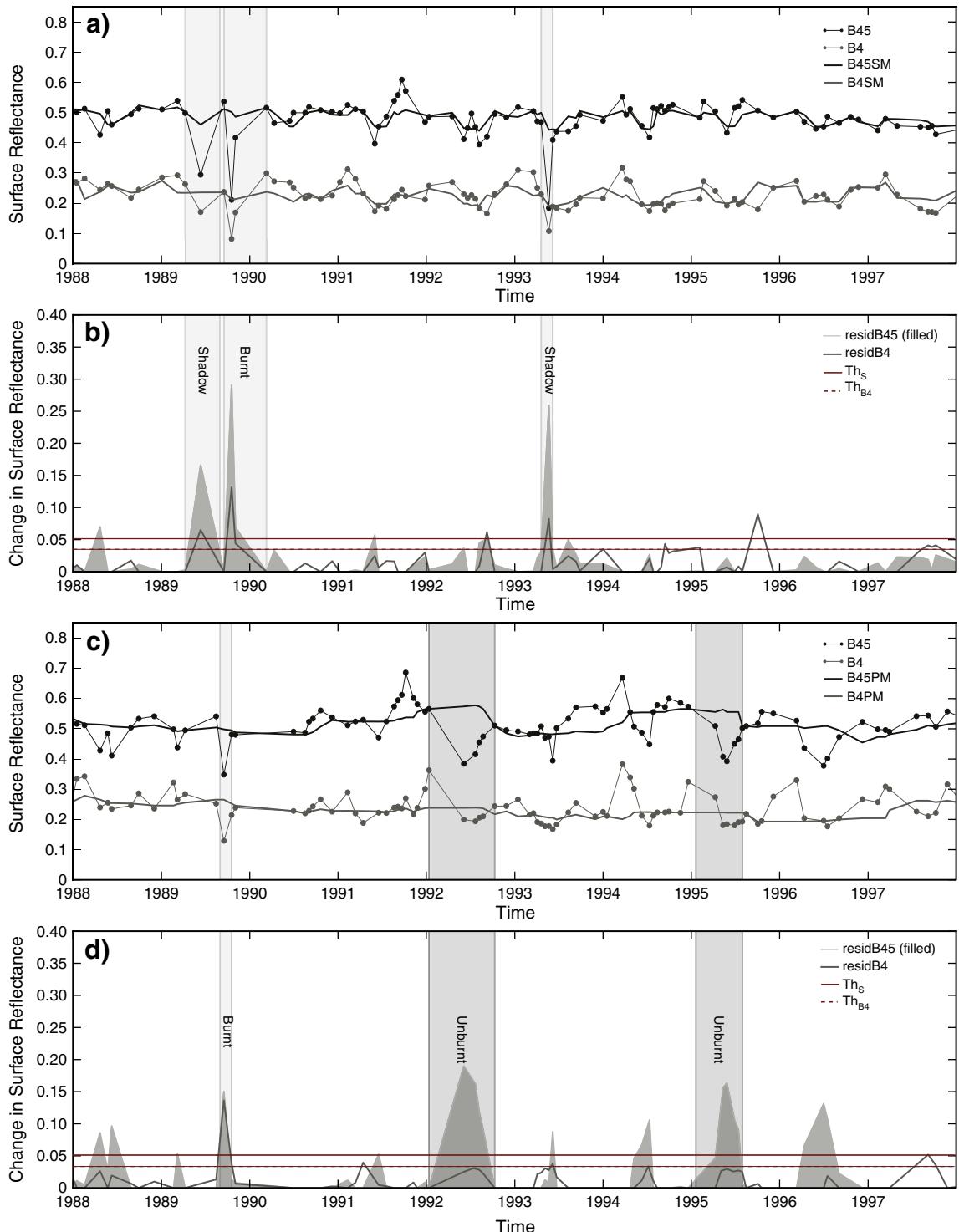
We found that the two seed watershed reduced the occurrence of over- or under-segmentation compared to defining a fixed interval of change. Spatially contiguous clusters of outliers were filtered to eliminate clusters smaller than 5 pixels. All remaining clusters of outliers were used as change seeds. To run the watershed filter,  $resid_{B45t}$  (Eq. 5) was inverted so that large magnitudes of change were converted to local minima. The set of change seeds then define local minima, and the unburnt seeds, elevated points on the relief. The filter is run using eight-neighbour connectivity, and the gradient of magnitude of change as the cost function to determine the watersheds (segment regions of change and no change).

### 2.3.3. Attribution of change regions

The attribution of change involves an object-scale, rather than a per-pixel, classification tree. Firstly, each of the watershed change regions were uniquely labelled to make a set of change objects. For Landsat-7 SLC-off imagery, a morphological closing filter was applied prior to labelling, to reduce the fragmentation of change objects due to SLC-off data. This was only applied to assist the attribution process, and immediately reversed after its completion.

All of the pixels in a given change object were assigned to a single output class: one of burnt or unburnt change, using a hierarchical sequence of decisions. The classification tree was based on the following spectral and thermal characteristics of the change object:

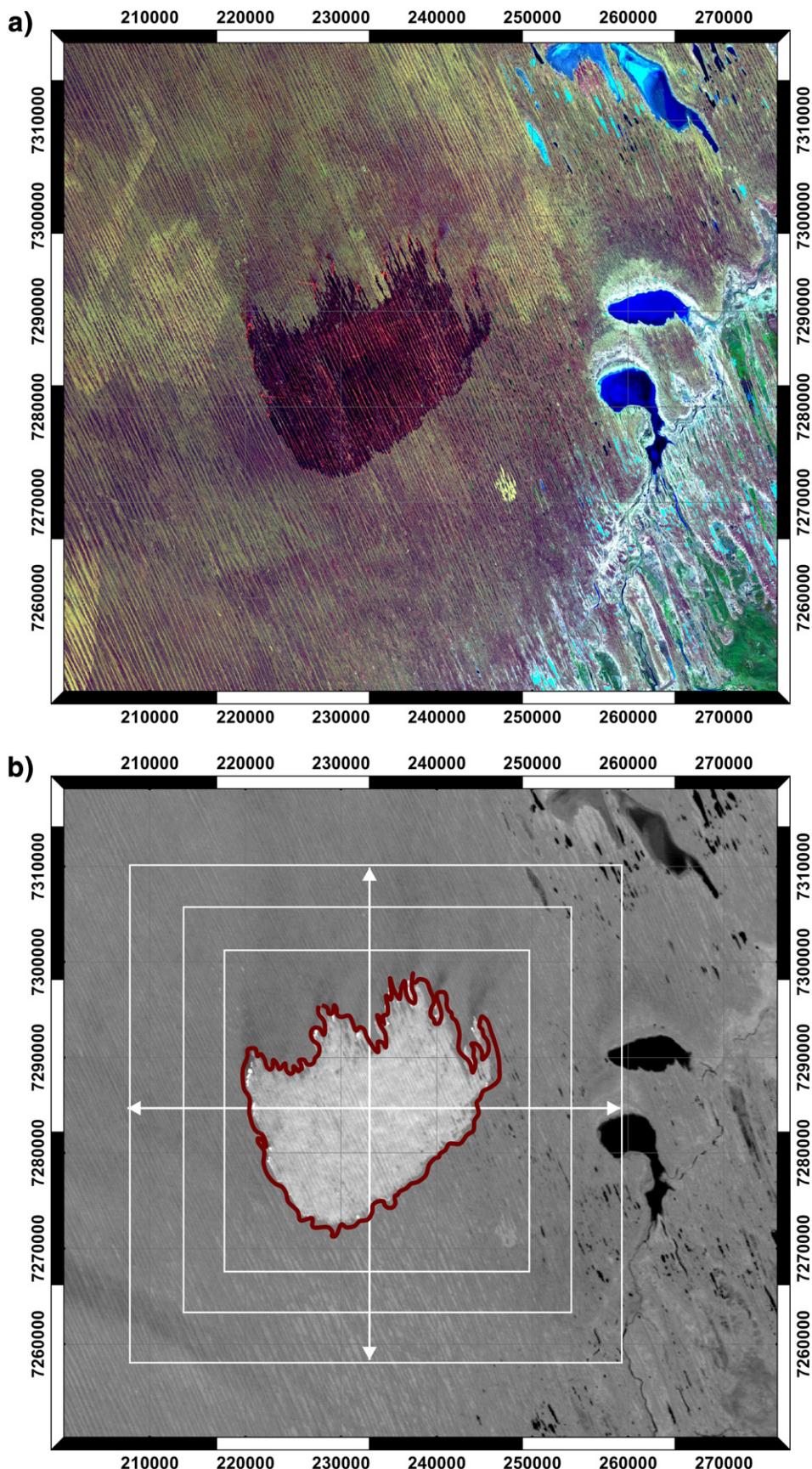
- i. The difference in median  $BT$  between the change object and a sample of unburnt pixels selected from the immediate neighbourhood of the change object. The median of this sample is assumed to represent the median  $BT$  of the surrounding unburnt vegetation. This test assumes that burnt pixels are likely to emit more thermal energy (i.e. be hotter) than neighbouring unburnt pixels as a result of reduced foliage and the presence of charcoal. Conversely, change objects caused by cloud/topographic shadow, and/or water inundation are likely to be cooler than, or the same temperature as surrounding pixels. Cloud, cloud shadow, and water masks were applied to  $BT$  to



**Fig. 2.** Example B4 and B45 reflectance curves for two different pixels: (a) where the seasonal median was used as the 'no change' reference for a pasture location; and (c) where the preceding median was used for a eucalypt woodland location. The magnitude of change in B4 and B45,  $\text{resid}_{B45t}$  and  $\text{resid}_{B4t}$  are also shown for the two pixels in (b) and (d). The red solid and dashed lines represent the optimised thresholds on  $\text{resid}_{B45t}$  and  $\text{resid}_{B4t}$ ,  $\text{Th}_S$  and  $\text{Th}_{B4}$  required for a pixel to be detected as a negative outlier (potentially burnt). Examples of burnt, cloud shadow and unburnt change are shown, highlighting both similarities and differences in B4 and B45 temporal characteristics in response to different surface changes. Note: the time period shown is not the complete time series but has been subsetted to highlight detail of reflectance time series.

minimise the risk of wet, shaded, or cloud contaminated pixels biasing the median neighbourhood BT. All change objects mapped by the watershed region growing filter (Section 2.2) were also eliminated from the BT sampling for the same reason. Systematic sampling of the neighbourhood was undertaken by selecting all valid pixels

within a rectangular window around the change object centre, and incrementally growing the window until a valid sample of neighbourhood pixels was found. For each increment, the window is expanded by 10 pixels in each direction until a valid sample of between 50 and 500 "unburnt" pixels is found (refer Fig. 3). Sampling



**Fig. 3.** Demonstration of the median brightness temperature difference test ( $BT_{neighborhoodDiff}$ ) used in the classification tree, showing a fire scar in Spinifex grasslands, far western Queensland dune-fields: (a) TM RGB 541, Path/Row: 97/77 and date: 6th of November 2011; and (b) the corresponding brightness temperature (from band 6). The relative temperature difference between the hotter (brighter) fire scar and surrounding unburnt vegetation is evident. The white squares indicate the domain of the growing neighbourhood sampling window, and the red line indicates the mapping derived from change detection and region growing.

was also constrained by the percentage cloud cover in the neighbourhood window to minimise the risk of undetected cloud influencing the median neighbourhood brightness temperature. If there was greater than 75% cloud cover (determined from cloud mask) the window was incrementally grown beyond these limits.

- ii. the change object's median  $B7$ ;
- iii. the change object's minimum band 7 reflectance,  $B7$ ;
- iv. the change object's mean and maximum  $B45$  value.

The construction of the decision tree classification and selection of these object indices is detailed in [Section 2.4.2](#).

#### 2.4. Optimisation

Image classifications generate both commission and omission errors with the desired outcome usually dependent on a trade-off between the two, determined by the requirements of a particular application. In this study, optimisation was performed in three separate stages with equal weighting given to omission and commission errors. The first step was to determine an initial 'best set' of thresholds for change detection ( $Th_S$ ,  $Th_B$  and  $Th_{B4}$ ). These results incorporate both fire and non-fire related changes. The second step was the construction of the 'best' classification tree for attribution of detected change. The final step involved a re-optimisation of change detection thresholds with the final classification tree included, as interactions between change detection and attribution stages will influence the balance of omission and commission rates and therefore the final selection of thresholds.

##### 2.4.1. Optimisation of change detection

The optimisation of thresholds required a representative sample of burnt pixels across Queensland. To achieve this, ten Path/Rows were strategically selected to capture the variation in fire and land surface characteristics ([Fig. 1](#)). For each Path/Row, 4 image dates were selected. Using a list of image dates from a random number generator, the first two images visually identified as containing burnt area and two without were selected for training purposes (total number of training images = 40). For each training image, 400 training points were selected in a split sampling process. This included: i) 200 random points; and ii) 200 stratified random points from above the 98th percentile of  $resid_{B45t}$  (total number of random points for all Path/Rows and image dates = 16,000). For each point, an image analyst recorded whether the pixel contained burnt vegetation (due to a recent or older fire), cloud, cloud shadow, smoke, water or crop. The remaining points were attributed as 'unburnt'. We considered a pixel to be burnt if there was any evidence of charcoal/ash, presence of a visible active fire front, or other changes in vegetation such as removal of foliage or scorch which could be attributed to a fire by the analyst. An older burn is a change due to fire captured in the image(s) immediately preceding the target date, while a recent burn recorded the first occurrence.

The optimisation was performed by iterating over all combinations of the thresholds  $Th_S$ ,  $Th_B$  and  $Th_{B4}$  within suitable ranges, initially without, and then including the attribution classification tree. The range for testing each given threshold was determined through experimentation. In the final optimisation, the sensitivity, stability and behaviour of the mapping performance over these ranges were assessed through examining the variation in total error, commission and omission rates with different thresholds ([Table 1](#)).

#### 2.4.2. Development of the classification tree

Change attribution was developed using a classification tree ([Breiman, Friedman, Olshen, & Stone, 1984](#)). This involves recursive splitting of the training data into increasingly pure "nodes" of burnt and unburnt observations based on a hierarchical sequence of rules. The classification tree is applied to change objects rather than single pixels and a set of training objects were required. The set of change layers corresponding to the dates and Landsat WRS Path/Rows from [Section 2.4.1](#) were produced using the initial 'best' set of thresholds (determined without attribution). The stratified random points (from [Section 2.4.1](#)) were then used to target change objects. Manual editing of the training objects ensured no gross mapping errors or mislabelled change and supplemented the training data when there were too few objects sampled.

A range of spectral and thermal indices known, or experimentally determined, to be good predictors of burnt/unburnt vegetation were calculated for each pixel in each training object. Then a number of statistical metrics were used to summarise the pixel distributions of each index in the object. These include: mean, standard deviation and percentiles. Additionally, the difference between object and neighbourhood median brightness temperatures ([Section 2.3.3](#)), and object and global median brightness temperatures were computed for each object. The set of predictors used in each object's classification is summarised in [Table 2](#). The Gini Importance measure from an exploratory Random Forests analysis was used to quantify the importance of the classification predictors and eliminate those of less explanatory value ([Breiman, 2001](#)).

The classification tree was constructed using the Recursive Partitioning (Rpart) module ([Themeau, Atkinson, & Ripley, 2012](#)) in the R statistical software. The Gini Impurity Index was used to determine the best split at each node. Theoretically, the tree would be grown until all training data is 100% correctly classified. However, this can result in a very large tree in which the training data is over fitted, leading to poor performance when applied to other data. The Rpart complexity and maximum depth parameters were used to initially constrain the tree size. Pruning and implicit cross validation were then used to combine some of the lower nodes to achieve the desired balance of omission and commission errors. Default prior probabilities (0.4 and 0.6 for burnt and unburnt objects) based on sampling proportions were used. Commission and omission errors were treated equally (no weightings were used).

#### 2.4.3. Validation

Selecting a validation dataset that includes a range of burn characteristics is challenging given that fire can be a geographically small and infrequent feature in a Landsat time series. To assess both commission and omission errors across the entire time series, it is also advantageous to sample images containing a range of fire activity as well as images containing no fire. Using the ten validation Path/Rows ([Fig. 1](#)), we adopted a two stage sampling approach. First, a range of classified burnt area fractions (i.e. fraction of the image classified as burnt) were selected within a given Path/Row. All image dates were ranked in ascending order based on the fraction of classified burnt area. The images corresponding to the 1st, 20th, 40th, 60th, 70th, 80th, 90th, 93rd, 96th and 99th percentiles were then selected as validation dates. Secondly, for each selected image 5000 random points were generated and stored in a vector file (10 Path/Rows  $\times$  10 images  $\times$  5000 points = 500,000

**Table 1**

Parameters used in the optimisation. Note: ranges were determined experimentally.

Index	Range	Increment	Number of iterations
$B4$ difference	0.000 - 0.080	0.005	17
$B45$ difference seed threshold	0.050 - 0.120	0.005	15
$B45$ difference region grow threshold	0.010 - 0.070	0.005	13

**Table 2**

The predictors used in the classification tree.

Predictor	Spectral index	Statistical metrics
$BT_{Ptile0}, BT_{Ptile10}, \dots, BT_{Ptile100}, BT_{Mean}, BT_{Std}$	Brightness temperature ( $BT$ )	Percentiles: 0,10, ..., 100, mean, std. deviation
$B1_{Ptile0}, B1_{Ptile10}, \dots, B1_{Ptile100}, B1_{Mean}, B1_{Std}$	TM/ETM + reflective bands 1–5 and 7 ( $B1-B5, B7$ )	Percentiles: 0,10, ..., 100, mean, std. deviation
...		
$B7_{Ptile0}, B7_{Ptile10}, \dots, B7_{Ptile100}, B7_{Mean}, B7_{Std}$	Band 4 + band 5 ( $B45$ )	Percentiles: 0,10, ..., 100, mean, std. deviation
$B45_{Ptile0}, B45_{Ptile10}, \dots, B45_{Ptile100}, B45_{Mean}, B45_{Std}$	Normalised burn ratio ( $NBR$ ; Key & Benson, 1999)	Percentiles: 0,10, ..., 100, mean, std. deviation
$NBR_{Ptile0}, NBR_{Ptile10}, \dots, NBR_{Ptile100}, NBR_{Mean}, NBR_{Std}$	Difference between object and neighbourhood $BT$	Median
$BT_{nhoodDiff}$	Difference between object and global $BT$	Median
$BT_{globalDiff}$		

random points). An image analyst then visually identified and recorded the points that intersected a new burn (first occurrence in the time-series) or an old burn (second or higher occurrence in the time-series) as was done for the optimisation data. Ambiguous observations were also recorded where it was difficult to determine whether it was burnt or not burnt, or where smoke/cloud/shadow obscured a clear view of the surface. These were discarded from the analysis. For all but one image date (Path/Row: 90/78 and date: 19th of December 2001), this represented <1% of the sample observations. Images earlier in the temporal sequence were inspected to determine the timing of the fire event. All other points were considered to represent unburnt land surface reflectance. This data set was used in confusion matrices to assess the performance of the method. On-screen validation may sometimes miss burnt areas with low fractions of charcoal/ash as they can appear indistinct from surrounding land surfaces. However, in the absence of an independent set of field observations providing adequate spatial and temporal sampling, this approach was considered the best available.

### 3. Results

#### 3.1. Burnt area optimisation

The classification predictor importance measured by the Gini Index from exploratory Random Forests found the most important predictor was the difference between object and neighbourhood median  $BT$  ( $BT_{nhoodDiff}$ ). The use of  $BT_{nhoodDiff}$  resulted in a mean decrease in node impurity (measured by the Gini index) of greater than 4 times the

next most important predictor, which was median  $BT$  ( $BT_{Ptile50}$ ). Interestingly, the three most important predictors were all based on brightness temperature, while the reflective indices contributed relatively little improvement to node purity.

The final 8-node Rpart tree is shown in Fig. 4. The first split alone using  $BT_{nhoodDiff}$  indicates 84% of burnt and 90% of unburnt training objects were correctly attributed (commission rates of 23% for burnt and 11% for unburnt attribution). Only three other predictors were selected in the final tree. Analysis of predictor surrogacy in the Rpart tree showed that in the case of the two reflective indices ( $B45_{Mean}$  and  $B7_{Ptile0}$ ), other predictors, including  $NBR_{Ptile0}$ ,  $B45_{Ptile50}$ , and  $B4_{Ptile0}$ , may have been substituted with little impact on performance.

Finally, we explored the sensitivity of thresholds,  $Th_S$ ,  $Th_B$  and  $Th_{B4}$ , for mapping burnt area. This showed that five combinations of thresholds produced the same lowest error value of 0.041, and a further 90 produced an error within 0.5% of the lowest error. All omission rates were <10% and the commission rate varied between 11 and 18%. Given the application, we considered these errors to be acceptable.

Within these error ranges  $Th_S$  ( $B45$  seed threshold),  $Th_B$  ( $B45$  unburnt threshold) and  $Th_{B4}$  ( $B4$  change threshold) thresholds varied between 0.050–0.080, 0.005–0.025, and 0.000–0.040, respectively, which indicates that performance is not highly sensitive to threshold selection. For example,  $Th_B$  of 0.005–0.015 produced similar results for any given set of  $Th_S$  and  $Th_{B4}$ . There was however, a strong interdependency between  $Th_S$  and  $Th_{B4}$ , where high  $Th_S$  values were often associated with comparatively low  $Th_{B4}$  and vice versa. This is because higher seed values limit the number of false burns and a lower threshold on  $B4$  change allows a greater spatial extent of the detected burns to be

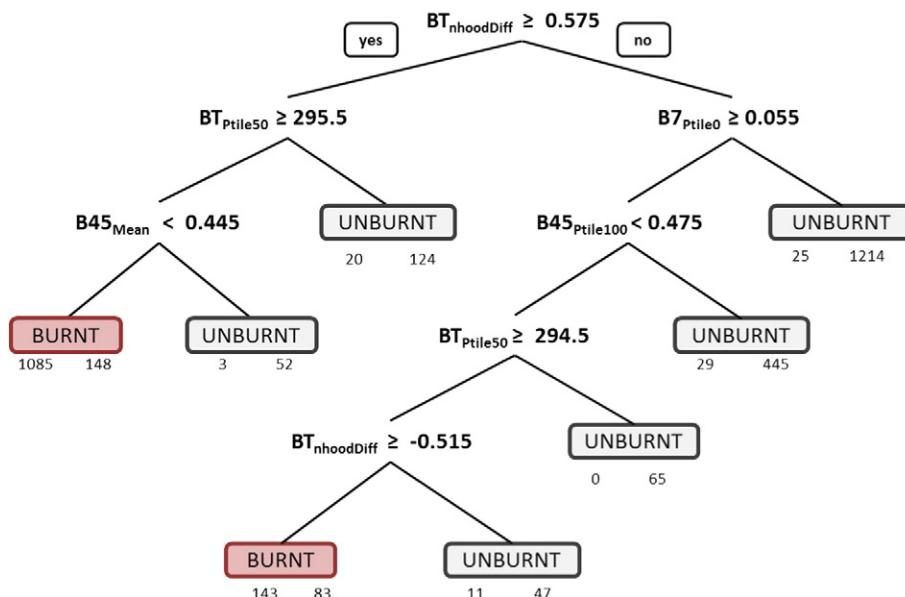


Fig. 4. Rpart classification tree used for attributing change objects.

mapped, while minimising commission errors. However, this can be at the expense of more subtle burnt areas that require lower  $Th_s$  for detection. Visual inspection of the resulting classifications found that the more extreme/polarised combinations, although producing similar results statistically, showed greater variability in commission/omission error rates in different land cover types. Based on the sensitivity analysis we chose  $Th_s = 0.050$ ,  $Th_B = 0.015$  and  $Th_{B4} = 0.030$  which resulted in omission, commission and total error rates of 8.6%, 15.7% and 4.1%, respectively. These optimisation results are unlikely to represent true accuracies due to the biased sampling of change areas for training purposes.

### 3.2. Burnt area validation

The validation data consists of 100 images sampled from 10 locations across Queensland. These sampled image dates however, show a bias towards the dry season (~May to October) particularly for images containing classified burnt area (Fig. 5). This is to be expected as this period covers the main burning season over much of Queensland, and the chance of acquiring imagery with low to moderate cloud cover is also higher in the dry season.

The validation results for our automated time series approach are shown in a contingency table (Table 3). These demonstrate that a high proportion of burnt area observations were correctly classified with an average producer's accuracy of 85% (range of 28 to 100% for individual images). An average user's accuracy of 71% (range of 4 to 99% for individual images) also indicates that the method has moderate commission errors. Although, the unburnt and total classification accuracy approached 100%, these results were largely driven by the unburnt class as the burnt area represents a small fraction of total area sampled (<1%) and only have a small influence on the total accuracy assessment. The producer's accuracy increased by 4% while the user's accuracy declined by 8% when a morphological dilation or buffer of 1 pixel was applied to the burnt class for pixels exhibiting a  $B45$  decline. The use of a pixel buffer to detect changes at the more subtle edges of fire scars may benefit some applications which seek to maximise the area of disturbance mapped, such as screening disturbances from time series. However, it may be problematic for other applications, such as studies looking at the fire patchiness and biodiversity changes in relation to fire.

An example of a burnt area classification sequence is shown in Fig. 6. This demonstrates that burnt areas are generally well delineated in the  $B45$  difference layer,  $resid_{B45t}$ , relative to unburnt vegetation. The spatial variability in the magnitude of  $resid_{B45t}$ , which most likely relates to both fire severity and the relative timing of fire and image acquisition, is also evident (Fig. 6b).

We examined the relationship between burnt area mapping accuracy and the fraction of a scene classified as burnt to determine whether the stratified sampling approach (by burnt class fraction) had any bearing on the validation results. The producer's and user's accuracies plotted as a function of the burnt class fraction per image are shown in Fig. 7. This demonstrates that similar producer's accuracies >75% were achieved across a range of burnt area fractions. Two images with >16 burnt validation points recorded producers accuracies below 50%. Further inspection revealed that in the first of these, burnt areas were undetected due to rapid greening of vegetation possibly in response to a rainfall event prior to the image capture (Path/Row: 97/74, 5th of December, 2012). In the second image, elevated smoke and atmospheric haze contaminated the attribution, and to a lesser extent the segmentation of burnt areas (Path/Row: 97/71, 4th of July, 2011). Images with fewer than 16 burnt validation points showed more variable results. In these cases even small fluctuations in the numbers of burnt omissions can strongly influence the producer's accuracy. The user's accuracies were more variable than the producer's accuracies across images with different fractions classified as burnt area. In general, where burnt area fractions are high (i.e. >3%) the user's accuracy is high. More variable user's accuracies were recorded for lower fractions of burnt area. Likewise for the producer's accuracy, when the burnt class fraction is low the user's accuracy is less reliable as even small numbers of falsely classified burnt pixels can have a big influence on the calculation.

The residuals in  $B4$  and  $B45$  for the validation data are shown in Fig. 8. Observations which failed to be classified as burnt generally have low  $B4$  and  $B45$  residuals which mostly fall below  $Th_s$  and  $Th_B$  thresholds. This indicates that the majority of burnt area omissions were due to insufficient magnitudes of change rather than incorrect attribution to burnt/unburnt change. Pixels misclassified as burnt occurred over a similar range to those correctly classified but were more clustered at lower residual levels (i.e.  $B4$  and  $B45$  residuals around 0.05 and 0.08, respectively).

Two additional factors possibly influencing the results include: (1) data gaps due to Landsat 7 Scan Line Correction (SLC) malfunction for imagery from 2003 onwards and (2) the fraction of cloud/cloud shadow cover. Visual inspection of classifications for SLC-off imagery showed that burnt areas were not always mapped to their full extent due to data gaps which prevented region grow operations. This was particularly evident towards the east/west edges of the images where the gap area is largest. These observations are supported by the results with producer's and user's accuracies lower for Landsat 7 ETM + SLC-off than the other imagery (producer's accuracy of 90% versus 68%; user's accuracy of 67% versus 29%; SLC-on versus SLC-off). However, as only 13 images out of the 100 sampled were SLC-off, more sampling would be required to confirm these results. The fraction of cloud/cloud shadow cover did not have a large impact on the results ( $r^2 < 0.02$ ). However, we have observed several examples where cloud shadow has occurred over part of a burnt area and the burnt area has been mis-attributed, or burnt pixels were removed during cloud shadow screening prior to burnt area detection. For individual images this may have a large impact on the classification accuracy.

Composites of burnt area from multiple dates for a given area can improve the detection of burnt area (the producer's accuracy). This is because burnt areas may be captured on subsequent image dates despite not being correctly mapped on the date of first capture. An example of a burnt area which was misclassified in the first capture date but then correctly mapped in the subsequent image acquisition is shown in Fig. 9. Conversely, the user's accuracy is likely to decline as

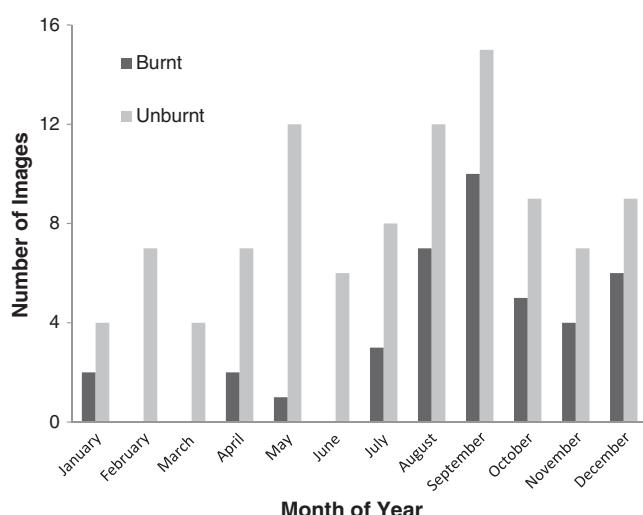


Fig. 5. Bar graph showing the distribution of burnt and unburnt images sampled throughout the year ( $n = 100$ ) for burnt area validation. There is a general bias towards the dry season (May–October) due to reduced cloud cover and increased fire activity.

**Table 3**

Contingency table showing the results with (and without) a pixel buffer.

	Burnt (reference)	Unburnt (reference)	Total	User's accuracy	Total accuracy			
Burnt (classified)	2424	(2545 <sup>a</sup> )	986	(1515 <sup>a</sup> )	3410	(4060 <sup>a</sup> )	71%	(63% <sup>a</sup> )
Unburnt (classified)	444	(323 <sup>a</sup> )	488,129	(487,600 <sup>a</sup> )	488,573	(487,923 <sup>a</sup> )	100%	(100% <sup>a</sup> )
Total	2868	(2868 <sup>a</sup> )	489,115	(489,115 <sup>a</sup> )	491,983	(491,983 <sup>a</sup> )		
Producer's accuracy	85%	(89%)	100%	(100%)				
Total accuracy							100%	(100% <sup>a</sup> )

<sup>a</sup> Pixel buffer of 1 pixel where  $B45$  change exceeded 0.

falsely mapped burnt pixels from different images in the sequence may accumulate in the composite.

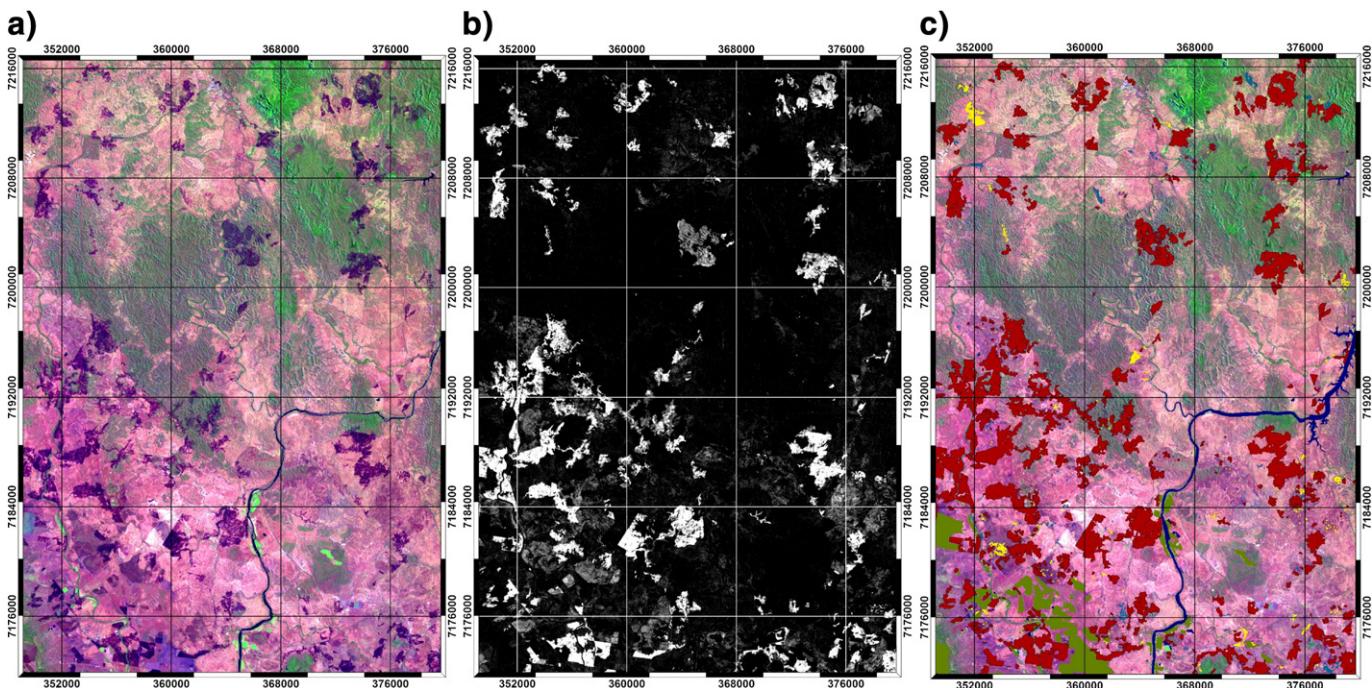
#### 4. Discussion

Combining time series change detection, region growing and object based attribution has proved to be an effective means of constraining both omission and commission rates when classifying burnt areas across a diversity of land surfaces and fire regimes in Queensland, Australia. These findings are consistent with earlier change detection and region growing studies (Bastarrika et al., 2011; Goodwin et al., 2013; Stroppiana et al., 2012). However, we have found that the use of time series analysis to characterise temporal trends in reflectance has facilitated the automation of our algorithm due to improved discrimination of fire related changes from seasonal and non-seasonal influences on vegetation.

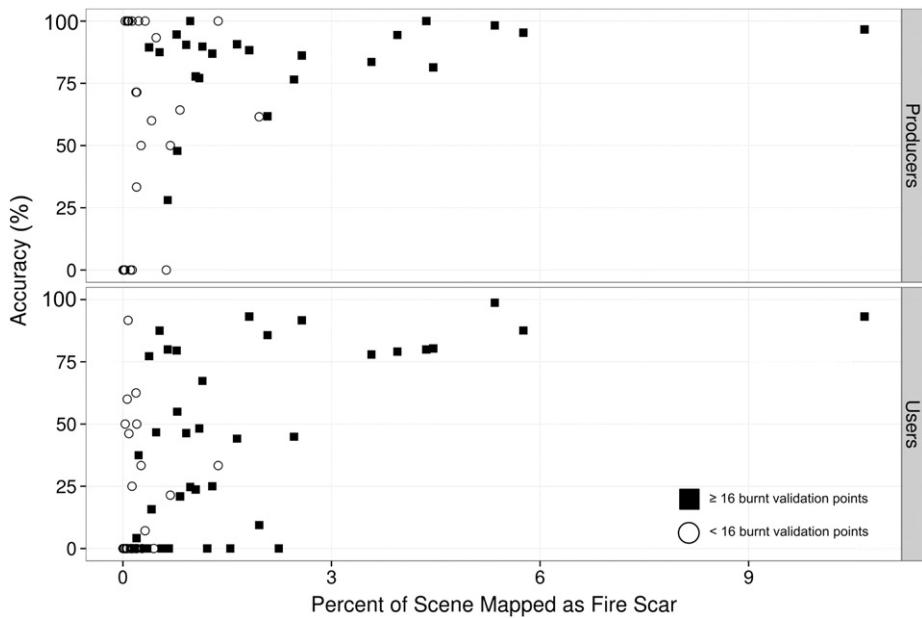
An important finding in this study was the usefulness of Landsat thermal band 6 in the separation of burnt and unburnt change objects in the classification tree. Earlier research by Lopez-García and Caselles (1991) suggested that Landsat band 6 thermal information could be useful in separating burnt areas from forest given that burnt areas are hotter than reference unburnt forests. Holden, Smith, Morgan, Rollins, and Gessler (2005) also developed a modified NBR using band 6 which showed encouraging results for a localised study. However, to

our knowledge, thermal information has not been widely used in the classification of burnt areas in Landsat imagery. After testing a range of indices, our work demonstrated that the two most important predictors in the classification tree were derived from brightness temperature (thermal infrared band 6) with the contextual neighbourhood difference test ( $BT_{nhoodDiff}$ ) proving to be crucial for classification accuracy and node purity. However, indices incorporating  $BT$  were not found to be useful in multi-temporal change detection as the high variability in  $BT$  (diurnal and seasonal) may obscure changes due to fire.

The two-seed watershed filter proved to be a robust algorithm for mapping the spatial extent of change regions, and showed an improvement on using a defined lower limit of change which was applied in cloud and shadow screening (refer Goodwin et al., 2013). There are however, future opportunities for developing better segmentation approaches, particularly given the increased adoption of segmentation techniques by the remote sensing community (Benz, Hofmann, Willhauck, Lingenfelder, & Heynen, 2004; Blaschke, 2010). Over- and under-mapping of burnt areas was evident, although not the primary source of error in mapping. Under-mapping was partially influenced by the threshold on  $B4$ , but also by the performance of the watershed filter, particularly in regions of denser woody vegetation. Over-mapping was sometimes evident where burnt areas were adjacent to other areas where large declines in reflectance were evident e.g. cloud shadow or cropping.



**Fig. 6.** Illustration of a burnt area classification and intermediate change layer. (a) Landsat image: RGB: 542, Path/Row: 90/78, and date: 29th of September 1996 (b)  $B45$  median residual ( $resid_{B45}$ ) where black indicated no or negative change and brighter colours indicates larger magnitudes of change and (c) Landsat image with burnt area classification overlaid (red = fire, green = crop mask, yellow = unburnt change, blue = water mask).



**Fig. 7.** Producer's and user's classification accuracies as a function of the scene percentage burnt. Note: points with fewer than 16 observations were marked as open circles.

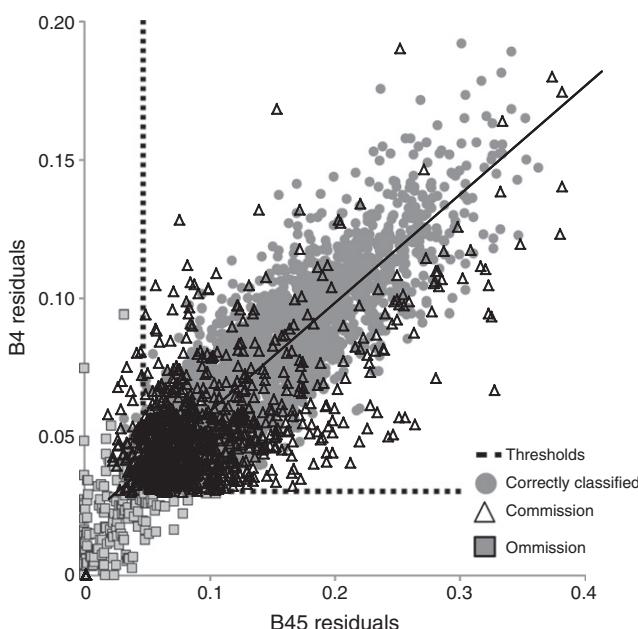
Discriminating the type of disturbance is a significant challenge in applying a change detection algorithm across a range of land cover types. In this study, we found that the Rpart classification tree provided accurate attribution of burnt areas at the object scale. It also proved to be an informative means of exploring the data, and controlling the balance of omission and commission rates. Yet, there is further potential to improve the classification and implement an updated version. The main problems encountered were due to misclassification where mixed features were represented within a change object e.g. where burnt area had grown into an adjacent cloud shadow or crop. There is potential

to use modal statistics as an indicator of two or more different surface changes occurring in a given object.

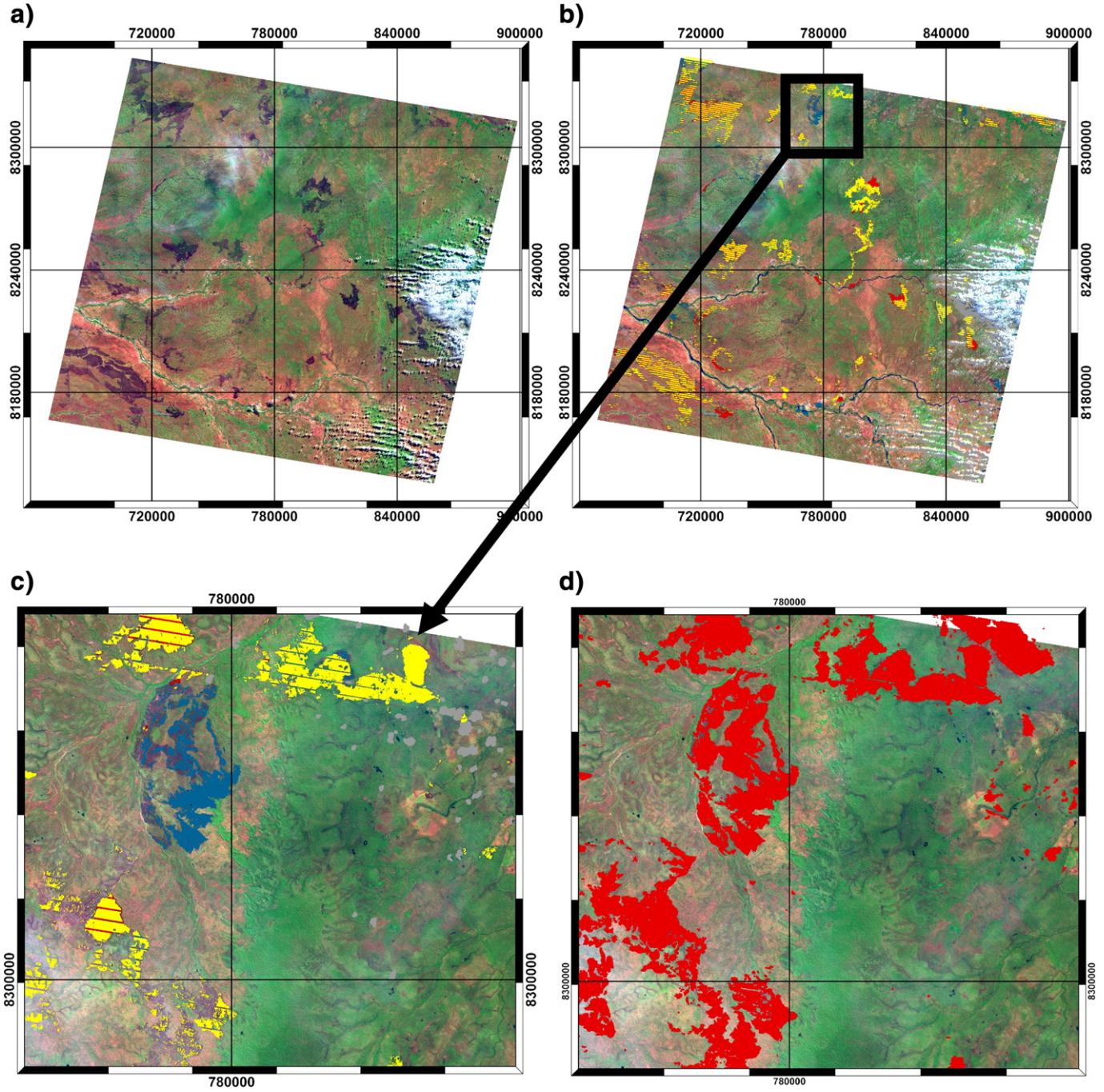
#### 4.1. Limitations

The results of this study are encouraging yet, burnt area maps derived from Landsat TM/ETM + imagery are unlikely to provide a complete record of fire activity. This is due to data gaps with sensor revisit times between 8 and 16 days which are further limited by cloud and cloud shadow obstruction. In dynamic environments such as savannas, burnt areas can remain undetected as ash/char can be blown or washed away over short periods of time (~weeks) in addition to the burnt area signals being masked by new growth of near ground vegetation or flush in the tree canopy. In extreme cases, we have noticed that burnt areas can fade in 16 days and appear indistinct from surrounding areas due to rapid greening of the burnt area.

The features/classes most frequently mis-classified as burnt area were cloud shadow, residual cropping/agriculture (not removed by the crop mask), and moisture/ground cover changes over dark soils. These features show similar spectral characteristics and temporal behaviour of  $B4$  and  $B45$  (i.e. negative outliers), often mimicking a burn event. Missed cloud in pre-processing was an issue for attribution as clouds are typically much cooler than the land surface biasing the neighbourhood brightness temperature difference test. Further improvements to cloud screening will alleviate this problem. Pixels containing agriculture/cropping are difficult to distinguish from burnt areas. Although a crop mask was applied, up-to-date maps are unrealistic due to frequent land-use change over a state covering 1.7 million km<sup>2</sup>. The challenge in cropping areas is that the surface is frequently harvested or ploughed for new crops which appear as negative outliers, particularly when the soil is dark. A further complication is that crop stubble may be burnt which can be difficult to determine, even visually, unless active fires are present. Attribution tests also encounter problems as fallow crops may be hotter than surrounding growing crops which are often irrigated. Ground cover and moisture changes in areas with dark soils were also a frequent source of commission error. These regions are often characterised by dynamic temporal reflectance curves showing significant inter- and intra-year variability in  $B4$  and



**Fig. 8.** Relationship between  $B4$  and  $B45$  difference layers for observations correctly detected, unclassified, and falsely classified. Note: these points were extracted following attribution.



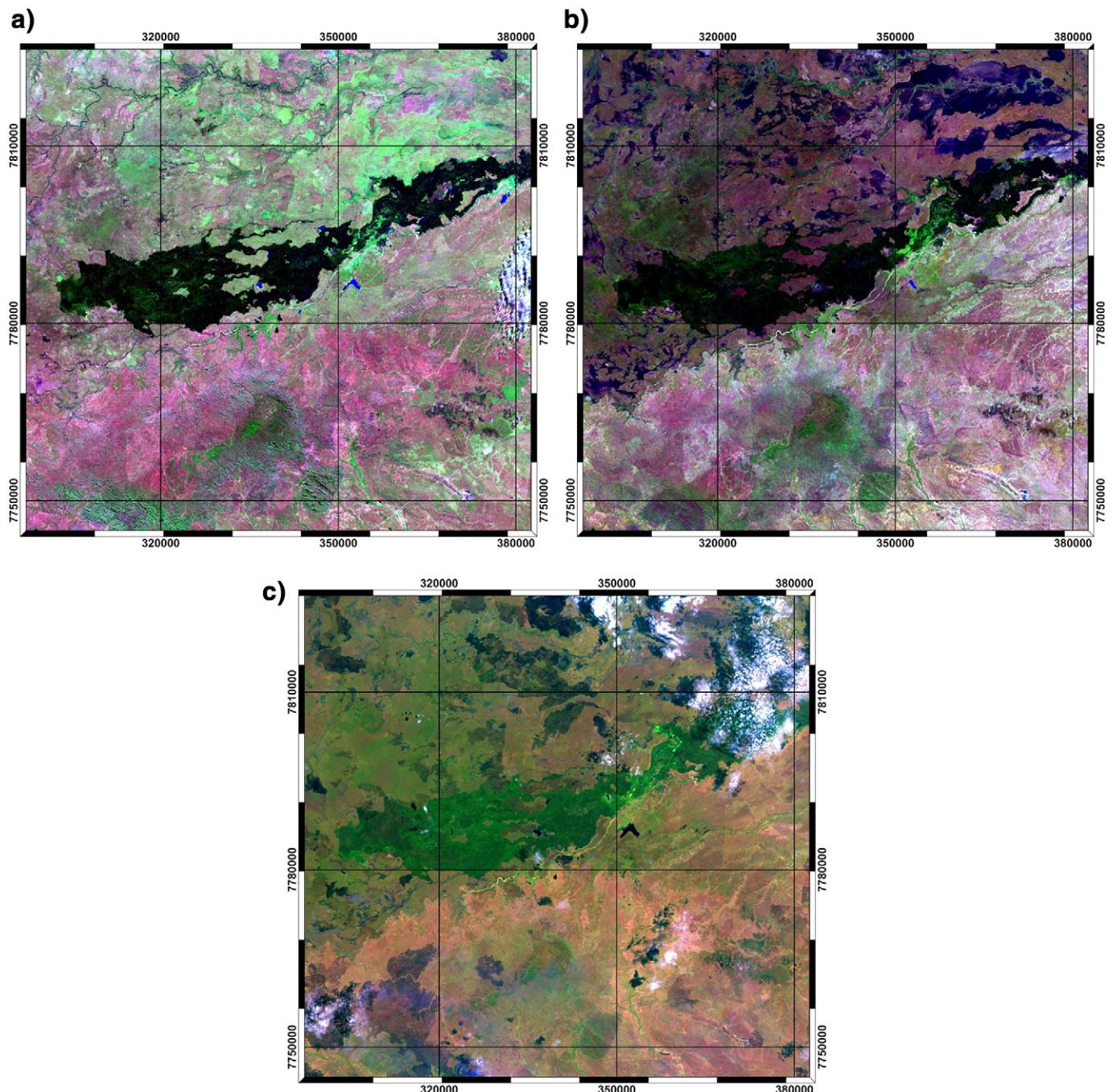
**Fig. 9.** Burnt area classification example: (a) TM image RGB: 542, Path/Row: 97/71, and date: 4th of July 2011 (b) image with burnt area classification in red and the preceding burnt area classification shown in yellow (date: 26th of June 2011), (c) illustration of mis-attributed burnt area in blue and (d) annual burn composite for 2011 shown in red indicating subsequent mapping of mis-attributed burnt area.

B45. An area of dark soil is shown in Fig. 10 illustrating the large variability in ground cover and difficulties in differentiating burnt areas from dark soil related changes.

The assumption that burnt areas decline in reflectance over time may not always be true. For example, sometimes only low fractions of charcoal/ash fractions are evident at image capture with a higher fraction of exposed soil. This can lead to an increase in B4 and B45 rather than a decrease. Specifically, there was visual evidence of omission errors in burnt area maps for Spinifex grasslands, primarily found in south-west Queensland. These were due to B4 and B45 increasing in response to fire due to the exposure of bright sandy soils after the char disappears (which can be very rapid). Ironically, the effects of fire

in Spinifex may be evident in satellite imagery for years but the charcoal typically remains only for short periods of time (i.e. days/weeks).

Furthermore, we found the validation of burnt area within a time series to be a difficult task. Typically, burnt areas represent a small fraction of the landscape and can experience considerable inter and intra-year variability. This means a random selection of image dates is not likely to capture a range of burn characteristics in all Path/Rows. Differences in fire characteristics also make it challenging to compare the results of different techniques, particularly where only a few fires/images have been sampled, as higher accuracies might be achieved due to the burn characteristics rather than the technique itself. Ideally, methods should be compared using the same dataset to avoid not only issues



**Fig. 10.** Example of land cover dynamics over a region of dark soils highlighting the difficulties in separating dark soils and burnt areas: (a) high fraction of green ground cover (date: 15th June 1998), (b) low ground cover and exposed dark soil on northern half of image (date: 2nd of October 1997), and (c) burnt areas on dark soils (date: 5th of December 1991). Note: RGB: 542 and Path/Row: 97/71 were constant.

relating to fire behaviour but also the diversity of land surface types that can vary spatially and temporally. There are also issues with using on-screen validation data in lieu of field data. Notably, burnt areas might be evident in the field but not visually evident in Landsat imagery, and burnt areas missed because of Landsat time series gaps will not be quantified. This may lead to higher classification accuracies than the truth. Nevertheless, field data collected in a consistent and repeatable fashion that is representative of the spatial and temporal range was not available for this study. This is a challenging issue in general for historic mapping from time series, particularly given the large and diverse study region involved. Future field-based analysis may be possible but will

be limited both spatially and temporally based on opportunistic data collection/collation.

## 5. Conclusions

In summary, the results of this study support the application of Landsat TM time series for classifying burnt areas. Our results averaged over ten Path/Rows showed that over 80% of burnt areas were detected with less than 30% commission error. Given the diversity of land surfaces and burn characteristics as well as the inclusion of cloud

contaminated imagery, we believe this to be an acceptable result for mapping burnt areas over large areas.

Our method has been developed for Queensland, Australia which is a savanna dominated environment and requires calibration to optimise region growing and the trade-off between omission/commission errors. However, the global algorithm is then able to be applied in an automated fashion. Even though the metrics used in this study may not be optimal in all environments, the framework that considers temporal, spectral, and contextual information should be transferrable to other environments.

Furthermore, the development of an automated time series approach for burnt area mapping has allowed the production of objective and repeatable maps of burnt area for the entire archive of Landsat imagery across Queensland, Australia (at present >60,000 images have been processed). This information will benefit not only our understanding of past and present fire regimes, but also allow natural resource managers to make more informed decisions.

## Acknowledgements

The authors acknowledge the support provided for this project by the: Science Division of the Queensland Department of Science, Information Technology, Innovation and the Arts; Reef Water Quality Science Program, Queensland Department of Environment and Heritage Protection; and State Rural Leasehold Land Strategy, Queensland Department of Natural Resources and Mines. The authors also wish to acknowledge the ongoing support and advice from Ken Brook, and the contributions of staff in the Queensland Remote Sensing Centre and the Joint Remote Sensing Research Program. Imagery used in this project was accessed from the United States Geological Survey's Landsat archive. Their efforts in making this archive available are gratefully acknowledged. Finally, we wish to thank the anonymous reviewers as well as Tony Gill (Office of Environment and Heritage, NSW Government) and Dan Tindall (DSITIA, Queensland Government) for their constructive feedback on the manuscript.

## References

- Allan, G., & Southgate, R. (2002). Fire regimes in the spinifex landscapes of Australia. In R. Bradstock, J. Williams, & A. Gill (Eds.), *Flammable Australia: The fire regimes and biodiversity of a continent* (pp. 145–176). Cambridge: Cambridge University Press.
- Bastarrika, A., Chuvieco, E., & Martín, M. P. (2011). Mapping burned areas from Landsat TM/ETM+ data with a two-phase algorithm: Balancing omission and commission errors. *Remote Sensing of Environment*, 115(4), 1003–1012. <http://dx.doi.org/10.1016/j.rse.2010.12.005>.
- Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., & Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3–4), 239–258. <http://dx.doi.org/10.1016/j.isprsjprs.2003.10.002>.
- Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1), 2–16. <http://dx.doi.org/10.1016/j.isprsjprs.2009.06.004>.
- Boer, M., Macfarlane, C., Norris, J., Sadler, R., Wallace, J., & Grierson, P. (2008). Mapping burned areas and burn severity patterns in SW Australian eucalypt forest using remotely-sensed changes in leaf area index. *Remote Sensing of Environment*, 112(12), 4358–4369. <http://dx.doi.org/10.1016/j.rse.2008.08.005>.
- BOM (2013). *Average annual, seasonal and monthly rainfall: 2011*. Australian Government, Bureau of Meteorology (Available at [www.bom.gov.au/jsp/ncc/climate\\_averages/rainfall/index.jsp](http://www.bom.gov.au/jsp/ncc/climate_averages/rainfall/index.jsp)).
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. United States of America: Chapman and Hall.
- Burrows, W. H., Henry, B. K., Back, P. V., Hoffmann, M. B., Tait, L. J., Anderson, E. R., et al. (2002). Growth and carbon stock change in eucalypt woodlands in northeast Australia: Ecological and greenhouse sink implications. *Global Change Biology*, 8(8), 769–784.
- Chander, G., Markham, B.L., & Helder, D. L. (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, 113(5), 893–903. <http://dx.doi.org/10.1016/j.rse.2009.01.007>.
- Craig, R., Heath, B., Raisbeck-Brown, N., Steber, M., Marsden, J., & Smith, R. (2002). The distribution, extent and seasonality of large fires in Australia, April 1998–March 2000, as mapped from NOAA-AVHRR imagery. *Australian fire regimes: Contemporary patterns (April 1998–March 2000) and changes since European settlement, Australia, State Environ. Second Tech. Pap. Ser.(Biodiversity)* (pp. 2) (Retrieved from <http://laptop.deh.gov.au/soe/2001/publications/technical/fire/pubs/part1.pdf>).
- Crowley, G. M., & Garnett, S. T. (1998). Vegetation change in the grasslands and grassy woodlands of east-central Cape York Peninsula, Australia. *Pacific Conservation Biology*, 4(2), 132–148.
- Derrien, M., & Le Gleau, H. (2010). Improvement of cloud detection near sunrise and sunset by temporal-differencing and region-growing techniques with real-time SEVIRI. *International Journal of Remote Sensing*, 31(7), 1765–1780. <http://dx.doi.org/10.1080/01431160902926632>.
- Disney, M. I., Lewis, P., Gomez-Dans, J., Roy, D., Wooster, M. J., & Lajas, D. (2011). 3D radiative transfer modelling of fire impacts on a two-layer savanna system. *Remote Sensing of Environment*, 115(8), 1866–1881. <http://dx.doi.org/10.1016/j.rse.2011.03.010>.
- Dyer, R., & Smith, M. S. (2003). Ecological and economic assessment of prescribed burning impacts in semi-arid pastoral lands of northern Australia. *International Journal of Wildland Fire*, 12(4), 403–413.
- Edwards, A.C., Hauser, P., Anderson, M., McCartney, J., Armstrong, M., Thackway, R., et al. (2001). A tale of two parks: Contemporary fire regimes of Litchfield and Nitmiluk National Parks, monsoonal northern Australia. *International Journal of Wildland Fire*, 10, 79–89.
- EHP (2011). *State of environment Queensland 2011*. Brisbane: Queensland Government Environment and Heritage Protection.
- Eidenshink, J., Schwandt, B., Brewer, K., Zhu, Z.-L., Quayle, B., & Howard, S. (2007). A project for monitoring trends in burn severity. *The Journal of the Association for Fire Ecology*, 3(01), 3.
- Felderhof, L., & Gillieson, D. (2006). Comparison of fire patterns and fire frequency in two tropical savanna bioregions. *Austral Ecology*, 31, 736–746.
- Flood, N., Danaher, T., Gill, T., & Gillingham, S. (2013). An operational scheme for deriving standardised surface reflectance from Landsat TM/ETM+ and SPOT HRG imagery for Eastern Australia. *Remote Sensing*, 5(1), 83–109. <http://dx.doi.org/10.3390/rs5010083>.
- Gillingham, S. S., Flood, N., Gill, T. K., & Mitchell, R. M. (2012). Limitations of the dense dark vegetation method for aerosol retrieval under Australian conditions. *Remote Sensing Letters*, 3(1), 67–76. <http://dx.doi.org/10.1080/01431161.2010.533298>.
- Goodwin, N. R., Collett, L. J., Denham, R. J., Flood, N., & Tindall, D. (2013). Cloud and cloud shadow screening across Queensland, Australia: An automated method for Landsat TM/ETM+ time series. *Remote Sensing of Environment*, 134, 50–65. <http://dx.doi.org/10.1016/j.rse.2013.02.019>.
- Henry, B. K., Danaher, T., McKeon, G. M., & Burrows, W. H. (2002). A review of the potential role of greenhouse gas abatement in native vegetation management in Queensland's rangelands. *The Rangeland Journal*, 24(1), 112–132.
- Holden, Z. A., Smith, A.M. S., Morgan, P., Rollins, M. G., & Gessler, P. E. (2005). Evaluation of novel thermally enhanced spectral indices for mapping fire perimeters and comparisons with fire atlas data. *International Journal of Remote Sensing*, 26(21), 4801–4808. <http://dx.doi.org/10.1080/01431160500239008>.
- Huang, C., Goward, S. N., Schleeweis, K., Thomas, N., Masek, J. G., & Zhu, Z. (2009). Dynamics of national forests assessed using the Landsat record: Case studies in eastern United States. *Remote Sensing of Environment*, 113(7), 1430–1442. <http://dx.doi.org/10.1016/j.rse.2008.06.016>.
- Justice, C. O., Giglio, L., Korontzi, S., Owens, J., Morisette, J. T., Roy, D., et al. (2002). The MODIS fire products. *Remote Sensing of Environment*, 83, 244–262.
- Keith, D. A., Williams, J. E., & Woinarski, J. C. Z. (2002). Fire management and biodiversity conservation: Key approaches and principles. In R. A. Bradstock, J. E. Williams, & A.M. Gill (Eds.), *The fire regimes and biodiversity of a continent. Flammable Australia*, Vol. 1. (pp. 401–428). United Kingdom: Cambridge University Press.
- Kennedy, R. E., Cohen, W. B., & Schroeder, T. A. (2007). Trajectory-based change detection for automated characterization of forest disturbance dynamics. *Remote Sensing of Environment*, 110(3), 370–386. <http://dx.doi.org/10.1016/j.rse.2007.03.010>.
- Key, C. H., & Benson, N. C. (1999). Measuring and remote sensing of burn severity. In L. F. Neuenschwander, & K. C. Ryan (Eds.), *Proc. Joint Fire Science Conference and Workshop, vol. II, Boise, ID, 15–17 June 1999*: University of Idaho and International Association of Wildland Fire (284 pp.).
- Klingaman, N.P. (2012). *A literature survey of key rainfall drivers in Queensland, Australia: Rainfall variability and change*. QCCCE Research Report: Rainfall in Queensland. Part 1. Brisbane, Australia: Department of Environmental and Resource Management, Queensland Government.
- Lopez-Garcia, M. J. L., & Caselles, V. (1991). Mapping burns and natural reforestation using thematic Mapper data. *Geocarto International*, 6(1), 31–37. <http://dx.doi.org/10.1080/10106049109354290>.
- Maier, S. W. (2010). Changes in surface reflectance from wildfires on the Australian continent measured by MODIS. *International Journal of Remote Sensing*, 31(12), 3161–3176. <http://dx.doi.org/10.1080/01431160903154408>.
- Maier, S. W., & Russell-Smith, J. (2012). Measuring and monitoring of contemporary fire regimes in Australia using satellite remote sensing. *Flammable Australia: Fire regimes, biodiversity and ecosystems in a changing world*. Melbourne: CSIRO Publishing.
- Masek, J. G., Huang, C., Wolfe, R., Cohen, W., Hall, F., Kutler, J., et al. (2008). North American forest disturbance mapped from a decadal Landsat record. *Remote Sensing of Environment*, 112(6), 2914–2926. <http://dx.doi.org/10.1016/j.rse.2008.02.010>.
- Miller, J.D., & Thode, A. E. (2007). Quantifying burn severity in a heterogeneous landscape with a relative version of the delta Normalized Burn Ratio (dNBR). *Remote Sensing of Environment*, 109(1), 66–80. <http://dx.doi.org/10.1016/j.rse.2006.12.006>.
- Miller, J.D., & Yool, S. R. (2002). Mapping forest post-fire canopy consumption in several overstory types using multi-temporal Landsat TM and ETM data. *Remote Sensing of Environment*, 82(2), 481–496.
- Mitchell, R., Qin, Y., & Campbell, S. (2009). Dust sentinel pilot: Capacity of three polar orbiting satellite instruments to map dust aerosol over the Australian continent. *Technical Report. CSIRO Marine and Atmospheric Research*. Canberra: ACT.

- NAFI, North Australian Fire Information (2014). Available at: [www.firenorth.org.au/naf1/](http://www.firenorth.org.au/naf1/)
- Orr, D.M., McKeon, G. M., & Day, K. A. (1991). Burning and exclosure can rehabilitate degraded black speargrass (*Heteropogon contortus*) pastures. *Tropical Grasslands*, 25, 333–336.
- Röder, A., Hill, J., Duguy, B., Alloza, J., & Vallejo, R. (2008). Using long time series of Landsat data to monitor fire events and post-fire dynamics and identify driving factors. A case study in the Ayora region (eastern Spain). *Remote Sensing of Environment*, 112(1), 259–273. <http://dx.doi.org/10.1016/j.rse.2007.05.001>.
- Roy, D. P., Ju, J., Lewis, P., Schaaf, C., Gao, F., Hansen, M., et al. (2008). Multi-temporal MODIS–Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data. *Remote Sensing of Environment*, 112(6), 3112–3130. <http://dx.doi.org/10.1016/j.rse.2008.03.009>.
- Russell-Smith, J., Ryan, P. G., & Durieu, R. (1997). A Landsat MSS-derived fire history of Kakadu National Park. *Journal of Applied Ecology*, 34, 748–766.
- Russell-Smith, J., Whitehead, P. J., Cooke, P.M., & Yates, C. (2009). Challenges and opportunities for fire management in fire-prone northern Australia. *Culture, ecology and economy of fire management in north Australian savannas: Rekindling the wurrk tradition* (pp. 1–22). Collingwood: CSIRO Publishing.
- Russell-Smith, J., Yates, C., Edwards, A.C., Allan, G. E., Cook, G. D., Cooke, P., et al. (2003). Contemporary fire regimes of northern Australia, 1997–2001: Change since Aboriginal occupancy, challenges for sustainable management. *International Journal of Wildland Fire*, 12, 283–297.
- Stroppiana, D., Bordogna, G., Carrara, P., Boschetti, M., Boschetti, L., & Brivio, P. A. (2012). A method for extracting burned areas from Landsat TM/ETM+ images by soft aggregation of multiple spectral indices and a region growing algorithm. *ISPRS Journal of Photogrammetry and Remote Sensing*, 69, 88–102. <http://dx.doi.org/10.1016/j.isprsjprs.2012.03.001>.
- Stroppiana, D., Pinnock, S., Pereira, J., & Grégoire, J. M. (2002). Radiometric analysis of SPOT-VEGETATION images for burnt area detection in Northern Australia. *Remote Sensing of Environment*, 82(1), 21–37.
- Themeau, T., Atkinson, B., & Ripley, B. (2012). Rpart: Recursive partitioning. R package version 3.1-55. <http://CRAN.R-project.org/package=rpart>
- Vermote, E., Tanre, D., Deuze, J., Herman, M., & Morcette, J.J. (1997). Second simulation of the satellite signal in the solar spectrum, 6S: An overview. *IEEE Transactions on Geoscience and Remote Sensing*, 35, 675–686.
- Vigilante, T., & Bowman, D.M. J. S. (2004). Effects of fire history on the structure and floristic composition of woody vegetation around Kalumburu, North Kimberley, Australia: A landscape-scale natural experiment. *Australian Journal of Botany*, 52, 381–404.
- Vincent, L., & Soille, P. (1991). Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6), 583–598.
- Williams, R. J., Griffiths, A.D., & Allen, G. E. (2002). Fire regimes and biodiversity in the savannas of northern Australia. In R. A. Bradstock, J. E. Williams, & A.M. Gill (Eds.), *Flammable Australia* (pp. 289–293). Cambridge: Cambridge University Press.