

Validation of surface water maps in selected Australian floodplains derived from Landsat imagery using hydrodynamic modelling

Catherine Ticehurst^a, Dushmanta Dutta^a, Fazlul Karim^a and Jai Vaze^a

^a CSIRO, Land and Water Business Unit, Black Mountain Laboratories, Australian Capital Territory
Email: catherine.ticehurst@csiro.au

Abstract: Mapping spatial inundation dynamics during floods is important for environmental management and disaster monitoring. Remote sensing technologies provide an affordable means for the routine monitoring of flood size and dynamics with reasonable spatial and temporal coverage. Landsat imagery (30-m horizontal resolution) can provide the appropriate spatial detail for many water mapping applications, although its temporal frequency is every 16 days. The public release of Geoscience Australia's Australian Geoscience Datacube (AGDC) offers opportunities for the consistent, repeatable production of surface water maps. The AGDC is based on the entire Landsat archive for Australia, following orthorectification and correction to surface reflectance with associated cloud masks. However, validation of these Landsat derived surface water maps is essential to provide confidence in them.

Hydrodynamic (HD) modelling tools are widely used for floodplain inundation modelling. While these tools require large amounts of data for model configuration and calibration, and can take a long time to process, they are highly effective in producing detailed characteristics of floodplain inundation with high accuracy at sub-daily time-steps throughout a flood event. For arid floodplains, antecedent soil moisture conditions and groundwater infiltration may need to be incorporated in such modelling.

This study compares two methods used for mapping surface water with Landsat in an Australian catchment and compares them to output derived from a well calibrated/validated, high resolution two-dimensional (2D) HD model. Using the 2D HD model to validate surface water extent derived from remote sensing has a number of advantages compared to the more common methods, which use higher resolution optical imagery or the same Landsat data to manually create a water/non-water dataset. The HD model provides an independent source of information, is available at a high spatial resolution, can be extracted to occur at a similar time as the image acquisition used to derive the water map, and can have full spatial coverage of the floodplain of interest. It is hoped that independent validation methods can help to provide confidence to the end user in the adoption of Landsat derived water maps.

Keywords: *Flood mapping, remote sensing, floodplain inundation, hydrodynamic modelling*

1. INTRODUCTION

Mapping spatial inundation dynamics during floods is important for environmental management and disaster monitoring. Remote sensing data can provide regional coverage at a spatial scale suited to the flood size and dynamics, with the spatial detail of Landsat imagery (30-m horizontal resolution) being appropriate for many water mapping applications (e.g. Fisher *et al.*, 2016). The release of the Landsat archive, making it freely available, has provided an opportunity for high spatial and temporal resolution analysis at a scale not seen before (including global; Pekel *et al.*, 2016). In Australia, this has included the public release of the Australian Geoscience Datacube (AGDC), which is based on the entire Landsat archive for Australia following orthorectification and correction to surface reflectance (Mueller *et al.*, 2016).

Generating maps of surface water from Landsat data can involve calculation of simple indices such as the Normalized Difference Water Index (NDWI), and modifications thereof (Fisher *et al.*, 2016; Xu, 2006). Other research efforts have incorporated Landsat bands and/or indices into their more complex algorithms to improve results (Mueller *et al.*, 2016). Validation of these methods is often based on the identification and/or manual digitisation of water in higher-resolution remote sensing imagery (e.g. Fisher *et al.*, 2016), or relying on the same remote sensing dataset to manually create a water/non-water pixel dataset (e.g. Tulbure *et al.*, 2016). While these methods are, in most cases, the most practical and cost-effective method for validating automated methods of mapping water using remote sensing imagery, since they eliminate the need for on-the-ground validation, it can be time consuming and subject to user error.

Hydrodynamic modelling tools are widely used for floodplain inundation modelling (Dutta *et al.*, 2016). These tools require large amounts of data for model configuration and calibration and take a long time to run – even over relatively small areas – so are really only practical to use over limited extents and dates. However they are highly effective in producing detailed characteristics of floodplain inundation with high accuracy. In recent years, some of the hydrodynamic models (such as MIKEFlood) have incorporated the capability to handle open water evaporation and infiltration rate, which are required for accurate simulation of inundation characteristics in large arid floodplains. This research compares two methods (modified NDWI (mNDWI; Xu, 2006) and Water Observations from Space (WOfS; Mueller *et al.*, 2016)) for mapping surface water along the Murray Darling floodplain in Australia, and compares it against output derived from a well calibrated/validated, high resolution two-dimensional hydrodynamic (HD) model. For this purpose, it is important to calibrate the HD model using independent datasets such as gauged data and field survey. Using the HD model to validate the surface water extent derived from remote sensing has a number of advantages over the more common methods of using higher resolution optical imagery, or using the same Landsat data to manually create a water/non-water pixel dataset: it provides an independent source of information; is available at a high spatial resolution; can be derived at a similar time to image acquisition; and can have full spatial coverage of the area of interest.

2. STUDY AREA

Two large floodplains, the Lower Balonne and Middle Darling floodplains, located in the northern regions of the Murray Darling Basin (MDB) were selected for implementation of this study. These two areas were selected due to the availability of high resolution LiDAR data to undertake hydrodynamic modelling. The Lower Balonne floodplain is one of the most complex floodplains with the flood hydraulics of the system resembling that of a delta with multiple channels. After the town of St George, the Balonne River divides into five separate channels. The Culgoa and Narran Rivers are the main channels, conveying 35% and 28% of the long-term mean annual flow at St George respectively; while the Ballandool and Bokhara Rivers and Birrie Creek flow only during higher discharge periods (Thoms *et al.*, 2002; MDBA, 2012). There are a number of significant wetlands located in the system. The area is frequently flooded with flows of up to 30,000 ML/d dispersing into many small flood channels. At higher flows, water spreads out over the floodplain and a significant portion does not return to the river as a result of evaporation and infiltration (MDBA, 2012). The floodplain in the Middle Darling System is confined mainly to the Darling River from Bourke to Tilpa and after that the Darling River divides to Wilcannia main channel and Talyawalka Creek and becomes a complex system connecting with a series of small and large lakes. For the purpose of this study, the Lower Balonne floodplain system (LBS) covers the floodplain between St George in Queensland and the Barwon River in northern New South Wales, and the Middle Darling floodplain system (MDS) covers the floodplain from Bourke up to Wilcannia (Figure 1). The study area was divided into three zones: zone 1: Lower Balonne System (LBS), zone 2: Darling Bourke (DB) and zone 3: Middle Darling System, as shown in Figure 2, based on the availability and quality of LiDAR data (Dutta *et al.*, 2016).

3. HYDRODYNAMIC MODEL

The two dimensional hydrodynamic model MIKE21 was used to undertake floodplain inundation modelling in the three modelling zones. MIKE21 is a computational hydrodynamic model that simulates the water level variation and flow in response to a variety of forcing functions in floodplains, lakes, estuaries, bays and coastal areas. The model solves the St-Venant’s equations of continuity and momentum. The MIKE21 model was set up for a number of historical flood events for the purpose of calibration and validation. The calibration parameter was mainly the Manning’s roughness coefficient, which was initially derived based on the land-use map. The model was manually calibrated against the observed water levels at different gauges for the selected historical flood events. In addition, the model was refined a number of times to accurately represent simulated inundation extents based on the information of historical flood events obtained from local communities in different locations within the floodplain through a number of consultations at different stages of the modelling (Overton *et al.*, 2016).

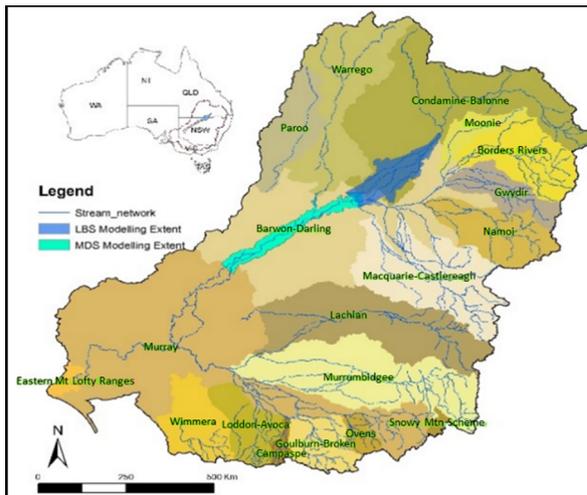


Figure 1. The extents of floodplain inundation modelling in LBS and MDS within the MDB.

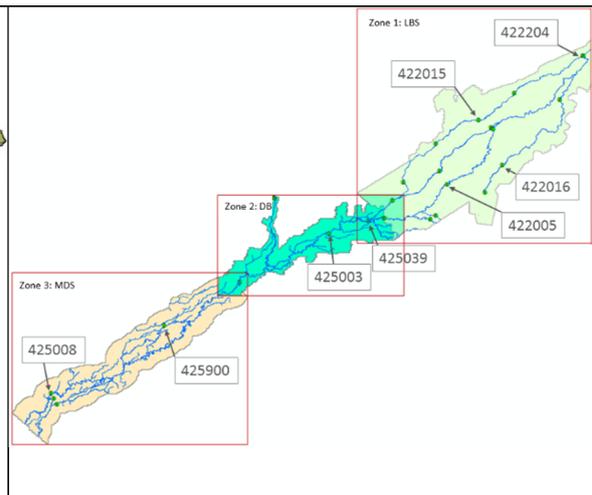


Figure 2. Number of modelling zones covering the two floodplains (relevant flow gauges shown).

Two dimensional Hydrodynamic models are data intensive. A large amount of temporal and spatial data are required for setting up a hydrodynamic model for inundation simulation. Data needed to setup the MIKE21 model include: water level and streamflow data; surface runoff; topography data (LiDAR data of 1-m resolution); stream network and cross-sections; surface roughness data and initial surface water levels. The details of the data collation and preparation for the model setup are elaborated in Dutta *et al.* (2016).

Four historical flood events were selected for the model calibration in the LBS and two events were selected for the DB and MDS. These were chosen because they are large flood events occurring within the lifetime of the Landsat data archive. The list of the selected events for the three modelling zones are presented in Table 1. The MIKE21 model was setup for these events for the three zones using the processed temporal and spatial datasets. The modelling time-step was five seconds.

Table 1. Flood events used for model calibration and validation.

Modelling zones	Year of selected flood event	Period of inundation modelling	Used for
LBS	1995-1996	27/12/1995-30/01/1996	Validation
	2008	19/01/2008-23/02/2008	Validation
	2010-2011	30/12/2010-31/01/2011	Calibration
DB	2012	28/01/2012-02/03/2012	Validation
	2010-2011	10/12/2010-28/02/2011	Calibration
MDS	2012	15/02/2012-22/03/2012	Validation
	2010-2011	20/12/2010-28/02/2011	Calibration
	2012	22/02/2012-31/03/2012	Validation

4. MAPPING WATER USING THE LANDSAT AGDC

4.1. Landsat water maps

Landsat data from Geoscience Australia’s Australian Geoscience Datacube (AGDC) (based on the Australian Reference Grid 25 product) was used to derive water maps. These data have been normalized for sun angle

effects (i.e. Bidirectional Reflectance Distribution Function (BRDF) corrected) to be spatially and temporally consistent and provided at a 25-m pixel size, and contain a mask of cloud and cloud shadow (Mueller *et al.*, 2016). The mNDWI and WOfS water maps are derived from the AGDC. The modified Normalized Difference Water Index (mNDWI) uses: $(\text{green-MIR})/(\text{green}+\text{MIR})$ (Xu, 2006). A threshold above -0.3 was used to mask the mNDWI image as water similar to Sims *et al.* (2014), who found this threshold was best suited to their application (based on multi-seasonal visual interpretation) in the Murray Darling Basin. The Water Observations from Space dataset is generated by Geoscience Australia and available to the public. WOfS uses a decision tree approach based on a selection of spectral bands and indices from the entire Landsat archive for Australia. It also provides a layer of confidence probability allowing users to assess the reliability of each pixel (Mueller *et al.*, 2016). The extensive WOfS accuracy analysis shows an accuracy of 93% in correctly identifying water.

4.2. Comparing Landsat and HD model water maps

Landsat imagery was identified for the three floodplain zones during flood events corresponding with the HD model runs. One clear, cloud-free Landsat scene was selected for each zone for a flood event. The selected Landsat overpass dates and Australian Eastern Standard Times are 22 January 2011 at 9.50am, 28 December 2010 at 9.57am and 20 January 2011 at 10.03am for the LBS, DB and MDS study sites respectively. The water maps generated from the HD model (at a 6-hourly time-step) were selected with the closest time to the Landsat overpass: 11am for LBS, and midday for DB and MDS as modelled on those same dates.

Spatial comparisons were made between the HD model and Landsat water maps (mNDWI and WOfS) for the three zones. The accuracy assessment (using the HD model as “ground truth”) was based on the Kappa statistic (Landis and Koch, 1977) due to the large number of non-water pixels within the floodplains in the images.

5. RESULTS

5.1. MIKE21 Hydrodynamic Modelling

Figure 3 presents the comparison of the simulated daily water levels against the observed daily water level at a number of key gauges in the LBS during the periods of simulations of two selected flood events for HD modelling (the yellow line shows the date of the Landsat scene). As can be seen from the figures for different gauges, the model performed well in water level simulations in both calibration and validation events. In particular, the model performance was reasonably satisfactory during the period of the high flood events of 2010-2011 and 2012 for most of the gauges (and very well for three of the four gauges at the time the Landsat scene was acquired). Figure 4 presents the comparison of the simulated and observed daily water levels at two gauges (423003 and 4250039) located in DB for the periods of two selected flood events of 2010-2011 and 2012. The agreements are highly satisfactory for both the events at the two gauges. Comparison of the simulated and observed water levels for two gauges located in MDS (425900 and 425008) for the two modelled events are shown in Figure 5. The simulated and observed data showed good agreements at the two gauges for both events. The values of linear correlation coefficients are above 0.6 for all events.



Figure 3. Comparison of simulated and observed daily water levels at a number of selected gauges in the LBS during the periods of the selected historical flood events (2011 and 2012 shown here). (The yellow line shows the date of Landsat scene).

5.2. Comparison of Landsat mNDWI and WofS with HD model

When the mNDWI and HD model water maps (resampled to the Landsat pixel size) are compared, the Kappa statistic indicates substantial agreement (i.e. between 0.6 and 0.8 (Landis and Koch, 1977)) for the LBS, with moderate agreement (between 0.4 and 0.6) for the DB and MDS study sites (Table 2). Comparison of the WofS and the HD model water maps show substantial agreement for the LBS and DB, but poor agreement (i.e. less than 0.4) for the MDS, according to the Kappa statistic. The mNDWI and WofS water maps show similar Kappa statistics for the LBS and DB, however the mNDWI is substantially better for the MDS than the WofS water map.

A comparison of the water maps are shown in Figure 6, where overall the agreement between the Landsat water maps (mNDWI and WofS) and HD model water maps is generally good (as indicated in blue). Figure 6 shows that the Landsat water maps are identifying water in irrigated cropping areas as well as small dams (as indicated in orange), where the HD model did not account for this. The HD model is mapping water in parts of the floodplain caused by overbank flow from the rivers, especially visible in the MDS study site, where the mNDWI – and more notably the WofS – are not identifying water (as indicated in

maroon). These areas are very flat and are visibly prone to flooding in the Landsat imagery.

Table 2. Comparison of mNDWI and WofS water maps with the HD model for the LBS, MDS and DB floodplains based on the Kappa statistic.

Floodplain	mNDWI	WofS
Lower Balonne System	0.66	0.64
Middle Darling System	0.57	0.39
Darling Bourke System	0.58	0.61

6. DISCUSSION AND CONCLUSIONS

The HD model was able to reproduce observed water levels extracted from streamflow gauges in most of the simulations tested in this study with a reasonable degree of accuracy. Comparison of the mNDWI and HD model water maps show that there is moderate to substantial agreement, while comparison of the WofS and HD model water maps show two of substantial agreement and one of poor agreement (according to the Kappa statistic). Reasons for discrepancies between these water maps could be due to a number of reasons:

- Although the difference between the Landsat overpass and HD model output is only one hour for the LBS and two hours for the MDS and DB floodplains, the LBS showed the best agreement, and the MDS showed the least accurate agreement between the Landsat and HD water maps. Altering the HD model output to directly coincide with Landsat overpass times may improve results, but examination of the HD model outputs during the MDS flood event shows the improvement would only be minimal. Simulated and observed daily water levels at the two gauges in the MDS compared well at the times the HD water maps were extracted, giving confidence in the water maps derived from the HD model, although no gauge was located in the area where the difference was greatest.
- The mNDWI threshold used to identify water was set at -0.3 , however optimisation of this threshold for different environments, such as the OTSU algorithm which uses a dynamic threshold method (as described in Du *et al.*, 2016), may help improve results – including those differences seen in the MDS.
- There are discrepancies between the HD model and Landsat water maps in irrigated agriculture and small permanent water bodies (e.g. dams) since they are not part of the HD modelling process, but are automatically detected in the Landsat water maps. Manual masking of these water features, or use of a land-use map if appropriate, would improve the statistical comparison results.
- Landsat water maps will map shallow water disconnected from the river, and possibly saturated soil as a result of local rainfall events, as water, whereas the HD model simulates inundation caused by overbank flow from the river network.
- The WofS water maps have a more complex algorithm for mapping surface water compared to the simple mNDWI. However the results here show that the individual flood maps within the WofS dataset tend to underestimate water extent compared to the mNDWI. This has been seen in previous unpublished research, where water has been identified in similar areas for both water maps. However the area of each water feature is not as large in the WofS data when compared to the mNDWI water maps as well as water that is visually identified in higher resolution imagery. Despite this result, the WofS statistical summaries, based on the entire Landsat archive, provides a valuable source of information on water dynamics and persistence within Australia.
- Although remotely sensed imagery provides an invaluable resource for flood mapping, it does have some limitations where the HD modelling is necessary, such as changes to the floodplain landscape (e.g. levees), diversion channels, environmental flow releases and scenario modelling. The HD model also has its

limitations in its dependency on the quality of DEM – which needs to reflect the actual landscape at the time of modelling. It is also unable to identify off-channel water features, even though they may be important in some environmental water applications.

The results show that the HD model can be used to provide one independent method for validating water maps generated from remote sensing, which can assist in providing confidence to end users in the use of Landsat-derived water maps. However, there are some areas of discrepancy such as irrigated agriculture and dams (both in the Landsat imagery), as well as areas where flooding varied between the HD model and Landsat water maps. Future work will further examine these differences and expand into other Australian catchments, as well as explore additional Landsat water mapping methods for comparison with HD models.

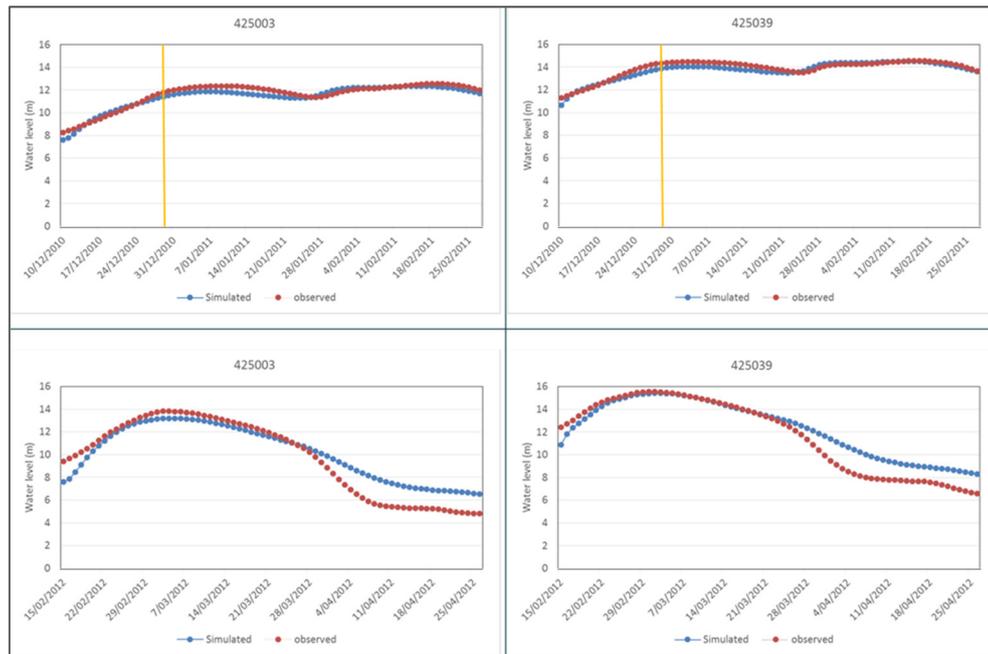


Figure 4. Comparison of simulated and observed daily water levels at two selected gauges in the DB during the periods of the selected historical flood events. (The yellow line shows the date of Landsat scene).

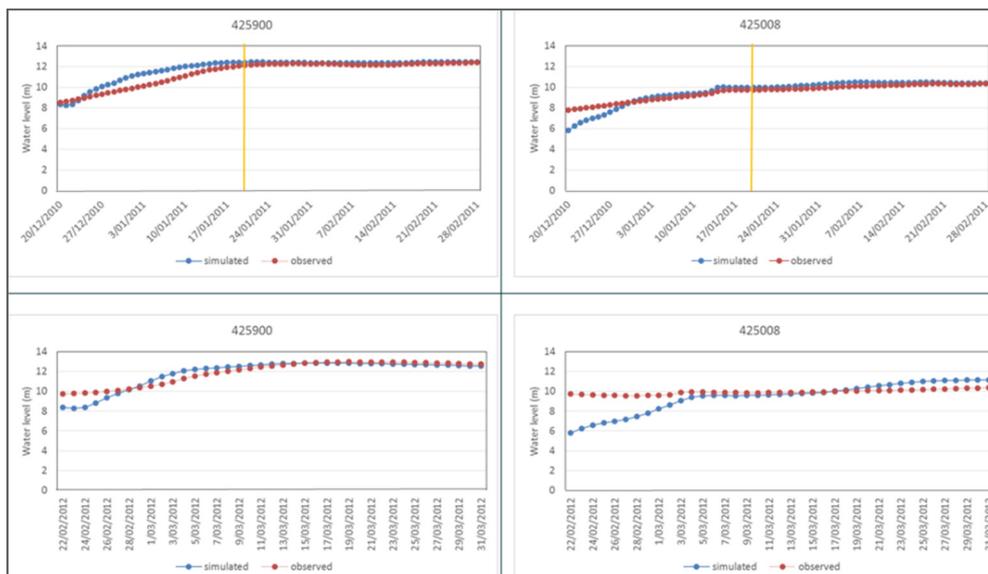


Figure 5. Comparison of simulated and observed daily water levels at two selected gauges in the MDS during the periods of the selected historical flood events. (The yellow line shows the date of Landsat scene).

ACKNOWLEDGMENTS

This work was carried out in CSIRO Land and Water.

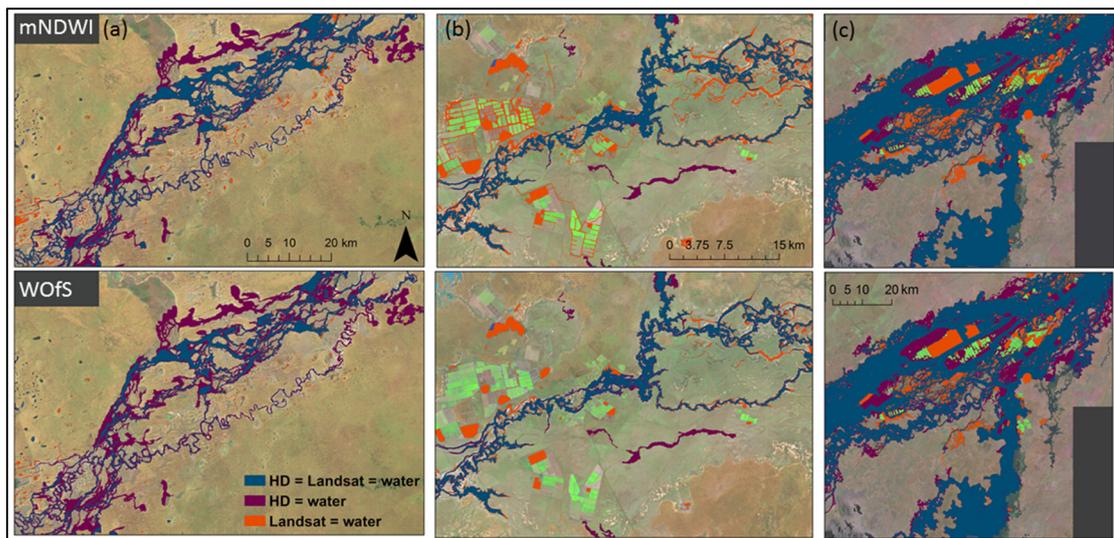


Figure 6. Spatial comparison of Landsat (mNDWI and WofS) water maps with the HD model for the (a) MDS on 20 January 2011, (b) DB on the 28 December 2010, and (c) LBS on the 22 January 2011.

REFERENCES

- Du, Y., Zhang, Y., Ling, F., Wang, Q., Li, W., Li X. (2016). Water bodies' mapping from Sentinel-2 imagery with modified Normalized Difference Water Index at 10-m spatial resolution produced by sharpening the SWIR band. *Remote Sensing*, 8, 354.
- Dutta, D., Vaze, J., Karim, F., Kim, S., Mateo, C., Ticehurst, C., Teng, J., Marvanek, S., Gallant, J., Austin, J. (2016). Floodplain Inundation Mapping and Modelling in the Northern Regions, the Murray Darling Basin. CSIRO Land and Water, Australia.
- Fisher, A., Flood, N., Danaher, T. (2016). Comparing Landsat water index methods for automated water classification in eastern Australia. *Remote Sensing of Environment*, 175, 167-182.
- Landis, J.R., and Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174.
- MDBA (2012). Assessment of environmental water requirements for the proposed Basin Plan: Lower Balonne Floodplain, MDBA Publication No: 24/12, Murray–Darling Basin Authority for and on behalf of the Commonwealth of Australia, 2012.
- Mueller, N., Lewis, A., Roberts, D., Ring, S., Melrose, R., Sixsmith, J., Lymburner L, McIntyre, A., Tan, P., Curnow, S., Ip, A. (2015). Water observations from space: Mapping surface water from 25years of Landsat imagery across Australia. *Remote Sensing of Environment*, 174, 341-352.
- Pekel, J., Cottam, A., Gorelick, N., and Belward, A.S. (2016) High-resolution mapping of Global Surface Water and its long-term changes. *Nature*, 540, 418-422.
- Overton IC, Doody TD and Kim S (2016). Data gathering and community engagement for validation of floodplain Inundation Modelling for the Lower Balonne and Darling Rivers, CSIRO Land and Water.
- Sims, N.C., Warren, G., Overton, I.C., Austin, J., Gallant, J., King, D. J., Merrin, L.E., Donohue, R., McVicar, T.R., Hodgen, M.J., Penton D.J., Chen, Y., Huang, C. & Cuddy, S. (2014). RiM-FIM Floodplain Inundation Modelling for the Edward-Wakool, Lower Murrumbidgee and Lower Darling River Systems. Report prepared for the Murray-Darling Basin Authority. CSIRO Water for a Healthy Country Flagship, Canberra.
- Thoms, M., Quinn, G., Butcher, R., Phillips, B., Wilson, G., Brock, M., Gawne, B. (2002). Scoping study for the Narran Lakes and Lower Balonne floodplain management study (R2011), Cooperative Research Centre for Freshwater Ecology, Canberra.
- Tulbure, M. G., Broich, M., Stehman, S. V., & Kommareddy, A. (2016). Surface water extent dynamics from three decades of seasonally continuous Landsat time series at subcontinental scale in a semi-arid region. *Remote Sensing of Environment*, 178, 142-157.
- Xu H (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27(14), 3025-3033.