Coondewanna Flats Natural History
December 2021

Prepared for
BHP Billiton Iron Ore

astron

Report Reference: 2481-21-EOSR-1RevB_230323

This page has been left blank intentionally.

Coondewanna Flats Natural History

Prepared for
BHP Billiton Iron Ore

Job Number: 2481-21

Reference: 2481-21-EOSR-1RevB_230323

Revision Status

Rev	Date	Description	Author(s)	Reviewer
A	11/04/2022	Draft Issued for Client Review	R. Archibald L. Robinson T. Kovacs-Ledo	J. Delfos
B	23/03/2023	Draft Issued for Client Review	R. Archibald L. Robinson	J. Delfos

Approval

Rev	Date	Issued to	Authorised by	
			Name	**Signature**
A	11/04/2022	K. Flowerdew	B. Lucas	
B	23/03/2023	K. Flowerdew	S. Easton	

Abbreviations

Abbreviation	Definition
BOM	Bureau of Meteorology
CI	Cordeiro Index
CPWRMP	Central Pilbara Water Resource Management Plan
DEA	Digital Earth Australia
FCIR	False Colour Infrared
GIS	Geographical Information System
ha	Hectares
KBDI	Keetch-Byram Drought Index
km²	Kilometres square
LiDAR	Light Detection and Ranging
m	Metres
MAR	Managed aquifer recharge
mm	Millimetres
MNDWI	Modified Normalised Difference Water Index
MSAVI	Modified Soil-Adjusted Vegetation Index
NAFI	North Australia and Rangelands Fire Information
NCI	National Computational Infrastructure
NIR	Near infrared
P	Priority
PEC	Priority Ecological Community
RF	Random Forest
SILO	Scientific Information for Land Owners
SLC	Scan Line Corrector
SWIR	Short-wave infrared
TCW	Tasselled Cap Wetness Index
WRI	Water Ratio Index
°C	Degree Celsius

Executive Summary

Lake Robinson is an ephemeral lake within Coondewanna Flats (the flats). Following periods of heavy rainfall, surface water pools on the lake and may persist for several months. There is uncertainty surrounding the nature of possible flooding events at other times, particularly when rainfall events and run-off have been significant but not extreme.

The vegetation on the flats has ecological significance, representing a Priority Ecological Community (PEC) *Coolibah-lignum flats: Eucalyptus victrix over lignum community in the Pilbara*. Surface flow and flooding are likely to be important in sustaining the structure and composition of the PEC. BHP's Mining Area C operations, immediately east of the flats, have the potential to impact hydrological regimes and the health of the PEC at the flats.

The purpose of this study was to use historical remote sensing imagery to better understand the history of flooding on the flats and the response of the vegetation to flood events. A study of all available cloud-free Landsat imagery from 1986 to 2021 was used to reconstruct the history of flooding and vegetation condition on the flats. The Modified Normalised Difference Water Index was used to detect flood events (defined as inundation that covers more than 10 ha) and the Modified Soil Adjusted Vegetation Index (MSAVI) was used to track vegetation condition. Data were compared against selected weather variables and a model to predict the presence-absence of flood events was devised using the Random Forest classification algorithm.

Based on Landsat image analysis, 11 flood events have occurred on Coondewanna Flats at an average frequency of one in every three years between 1986-97 and 2020-21. This frequency is higher than previous estimates of one in four years. Three small flood events, previously unreported, were detected to have occurred since the well-known flood event in early 2012.

The use of Landsat imagery to detect inundation enabled the distribution of water across the flats to be mapped for each flood event. These data showed during the maximum extent of large flood events, inundation persists across broader areas of the flats and not on Lake Robinson. For smaller flood events, inundation is more prevalent on Lake Robinson.

Modelling to predict flood events was attempted using variables relating to rainfall, evaporation, drought and time since previous flood events. The strongest predictor was 1-month rainfall total. This accords with the findings of earlier studies. However, these previous studies nominate 180 mm as the threshold 1-month rainfall total that initiates a flood event. The average 1-monthly total for the floods detected in this Landsat-based study was 131 mm.

While the use of Landsat imagery represents a useful technique for detecting historical flood events in remote areas, gaps in the available data due to cloud make it difficult to accurately identify the start of flooding. Furthermore, cloud may obscure some short-lived flood events. Data from an on-ground sensor network of surface flow gauges and cameras in suitable locations on the flats would enable the commencement of floods to be more accurately identified in future years. The use of cloud-penetrating radar data from the Sentinel satellite could also be utilised to detect patterns of flooding, noting that this satellite only became available from 2014.

The time series data for vegetation condition on the flats using the MSAVI index from Landsat data showed a strongly cyclical pattern whereby condition increases after heavy rainfall and flood events. Marked difference in vegetation dynamics following floods were evident between the two sub-types of the PEC on the flats. The results highlighted the importance of periodic floods in the ecology of this PEC.

Table of Contents

List of Figures

List of Tables

List of Appendices

1 Introduction

1.1 Background

Lake Robinson (the lake) is an ephemeral lake within Coondewanna Flats (the flats). The lake occurs in a topographic low at the central-north extent of the flats. Following periods of heavy rainfall, surface water pools on the lake and may persist for several months. Water is lost from the lake and the surrounding flats through evaporation or as recharge to soil and groundwater stores. The frequency of flooding and recharge events is estimated to be approximately once in every four years (AQ2 2015, BHP 2021).

Previous examinations of the flooding history of the flats have primarily focused on anecdotal observations, rainfall data and data from groundwater bores and surface flow gauging stations in adjacent catchments (RPS Aquaterra 2011, AQ2 2015, 2016). Studies by RPS in AQ2 (2015) identified the three wet seasons since 1990 where flooding was either observed (2011-12) or inferred (2002-03 and 2005-06) based on rainfall and runoff data. There appears to be some uncertainty surrounding the nature of possible flooding events at other times, particularly when rainfall events and run-off data have been significant but not extreme. A monthly rainfall total of 180 mm has been identified as a simple threshold predicting flooding on the flats (AQ2 2015).

The vegetation on the flats has ecological significance, representing a Priority Ecological Community (PEC) *Coolibah-lignum flats: Eucalyptus victrix over lignum community in the Pilbara*. Surface flow and flooding are likely to be important in sustaining the structure and composition of the PEC (Astron Environmental Services 2014, AQ2 2016). Surface water inflows replenish soil moisture reserves within plant root zones while flooding on the flats is linked to groundwater recharge events (AQ2 2015). Flooding and/or high soil moisture levels are also known to promote the regeneration of seedlings of riparian species in arid zone environments.

BHP's Mining Area C operations, immediately east of the flats, have the potential to impact hydrological regimes and the health of the PECs at the flats. The Central Pilbara Water Resource Management Plan (CPWRMP) (BHP 2021) summarises the potential impacts on hydrological regimes at the flats as:

- groundwater drawdown from mine pit dewatering
- groundwater mounding (rise) due to the reinjection of surplus water as part of a managed aquifer recharge (MAR) scheme
- changes in surface flows arising from disturbance within 7% of the catchment
- changes in the timing and duration of flows due to discharge from infiltration ponds, in particular an increase in the duration of flooding on the lake including flooding at atypical times outside of the wet season.

Potential impacts on the flora and vegetation within the PEC on the flats as a result of these hydrological changes encompass:

- declining health of phreatophytic species from groundwater drawdown – however, based on studies indicating that the vegetation on the flats is unlikely to be groundwater dependent, these risks are considered low
- alteration of growth patterns and community composition within the PEC from groundwater mounding.

Despite potential changes in flooding regimes at the lake, potential impacts to the PEC arising from changes of the magnitude predicted are considered negligible (BHP 2021).

To address the potential impacts on the PECs at the flats, Ministerial approval for Mining Area C includes Condition 6-1 (1):

> 'no reduction in the extent of each of the following components of the Coolibah-Lignum Flat Priority Ecological Community occurrence on the Coondewanna Flats:
>
> a) Coolibah woodlands over lignum over swamp wandiree, or
>
> b) Coolibah and mulga woodland over lignum and tussock grasses on clay plains, attributable to the revised proposal'.

To evaluate compliance against Condition 6-1 (1), the CPWRMP sets out a monitoring program for surface water and groundwater along with trigger and threshold criteria.

1.2 Scope of this Report

Given the potential for operations at Mining Area C to change the timing and duration of flooding on the lake, BHP has commissioned further study into the historical patterns of flooding, and vegetation response to flooding, at the lake. Records of flooding events and the response of vegetation to floods are lacking. As such, it is currently difficult to determine whether any changes in flooding frequency or duration are within the bounds of natural variability. Furthermore, if a change in natural flooding regimes was to occur, it is unclear what the implications of this would be for vegetation on the flats. A particular feature of this study, distinguishing it from previous studies, was to use historical remote sensing imagery to characterise the flooding history of the flats. Specific elements of the scope included the following:

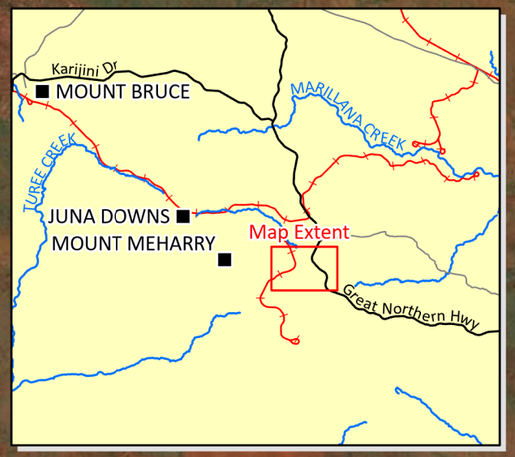
- a climate review to identify historical events that may have occurred at the flats
- analysis of remote sensing back captures to create a riparian vegetation health history compared to the above targeted key climatic events
- analysis of remote sensing back captures to create a flooding history and comparison to the climatic record.

The study results will be included in the future annual monitoring reports for the flats to provide additional background and context.

675000 680000 685000 690000

7460000

7455000



690
690
694
694
Great Northern Hwy
694
690
694

State Road

Local Road

Railway

Contours

Priority Ecological Community (PEC) *Coolibah-lignum flats: Eucalyptus victrix over lignum community*

Priority 3: Coolibah and mulga (Acacia aneura) woodland over lignum and tussock grasses on clay plains (Coondewanna Flats and Wanna Munna Flats)

Priority 1: Coolibah woodlands over lignum (Duma florulenta) over swamp wandiree (Lake Robinson is the only known occurrence)

BHP

Coondewanna Flats Remote Sensing History

Figure 1: Coondewanna Flats

Coordinate System: GDA 1994 MGA Zone 50

0 500 1,000 1,500 2,000 2,500 Metres

N

astron delivering environmental intelligence™

Author: R. Archibald | Drawn: L. Robinson | Date: 22-03-2023 | Figure Ref: 2481-21-BIDR-1RevB_230322_Fig01_Locn

2 Methods

2.1 Site Description

2.1.1 Climate

Analysis by (AQ2 2015) highlighted a significant shift towards wetter conditions in the Pilbara since the early 1960s. Given this, climate data dating back to 1961 were used so that recent trends can be compared to the contemporary climate. Data were sourced from the Scientific Information for Land Owners (SILO) database (Queensland Government 2022) for Coondewanna Flats (-23.0°S, 118.8°E), for the 1961 to 2021 period. SILO sources climatic data from the Bureau of Meteorology (BOM) and interpolates data between weather stations to provide a complete dataset for any location.

The climate of the flats is arid sub-tropical with two distinct seasons, a hot wet summer from October to April and a mild dry winter from May to September. Rainfall is variable with significant summer rainfall commonly associated with tropical lows or cyclones that move inland. Most rainfall occurs between January and March (Figure 2).

Average annual rainfall at the flats is 355 mm (1961 to 2021) which is well below average annual evaporation: 3,122 mm (1961 to 2021) (Queensland Government 2022). The evaporation-rainfall differential typically occurs across all months of the year (Figure 2). High evaporation rates are driven by hot temperatures with average maximum temperatures for December to February exceeding 40°C (Figure 3). High summer temperatures towards the end of the dry season result in very dry soil surface conditions, as reflected in the sharp upward trend in the Keetch-Byram Drought Index (KBDI) from August to September (Figure 4; see Section 2.3.3 for further discussion of the KBDI).

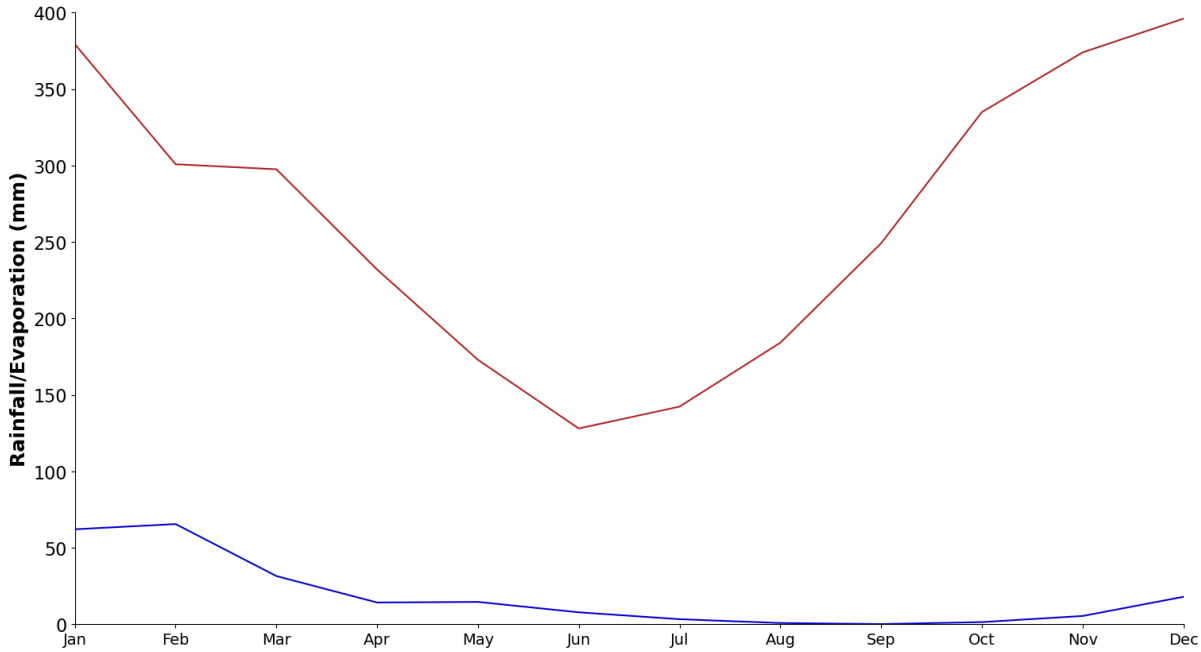


Figure 2: Average monthly rainfall and evaporation for Coondewanna Flats (1961 to 2021). Data are for -23.0°S, 118.8°E and were derived from the Scientific Information for Landowners (SILO) database (Queensland Government 2022).

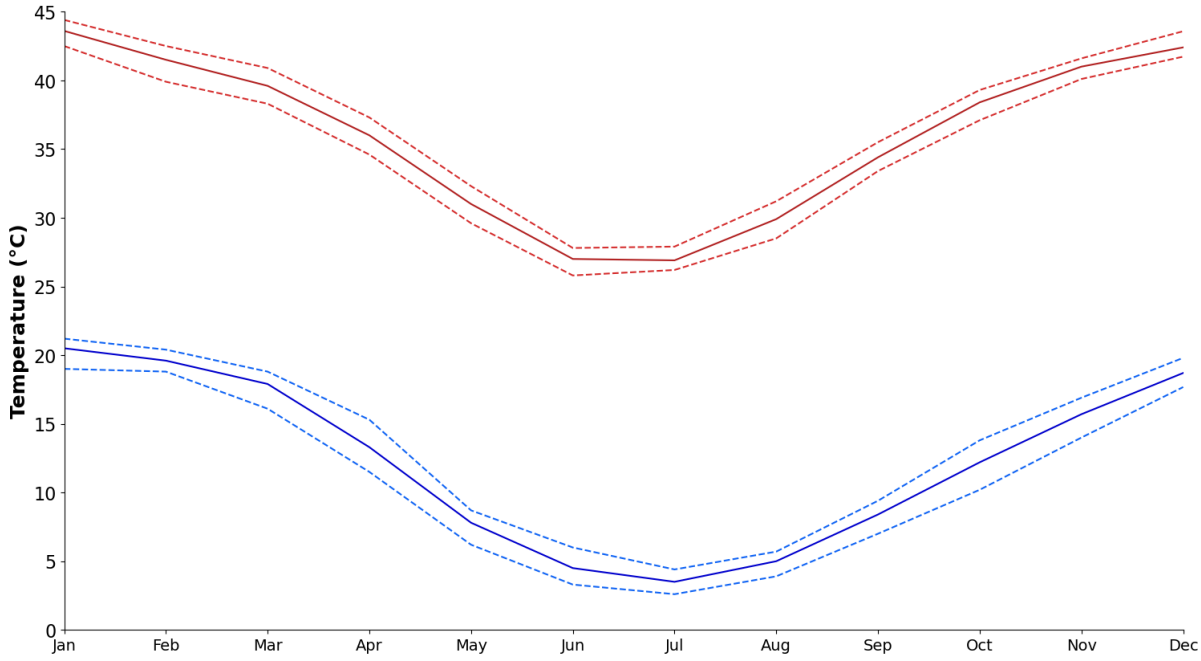


Figure 3: Average minimum and maximum temperature for Coondewanna Flats (1961 to 2021). Data are for -23.0°S, 118.8°E and were derived from the Scientific Information for Landowners (SILO) database (Queensland Government 2022).

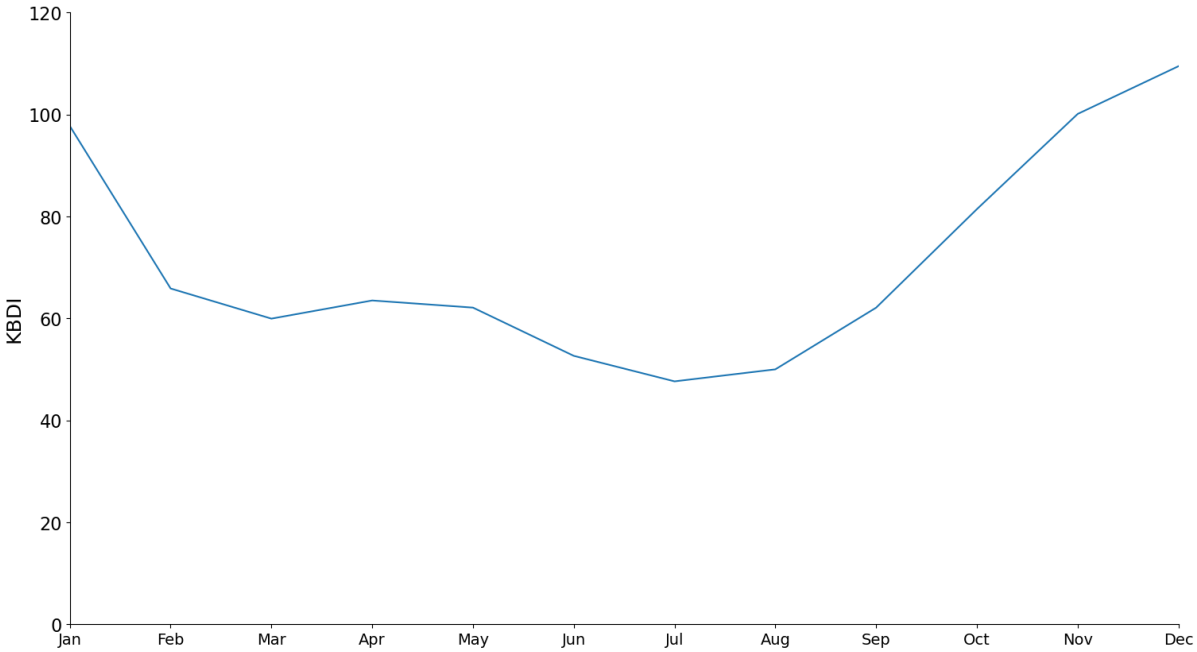


Figure 4: Average monthly Keetch-Byram Drought Index for Coondewanna Flats (1962 to 2021). Data are for -23.0°S, 118.8°E and were derived from the Scientific Information for Landowners (SILO) database (Queensland Government 2022).

2.1.2 Hydrology

Lake Robinson is situated within the northern section of the flats, although not necessarily the lowest lying area of the flats (Figure 5). Lake Robinson and the broader flats is a basin, receiving surface water inflows from the surrounding catchment of some 866 km^2 in area. The boundary of the Coondewanna Flats Basin was defined as the extent of the 690 m contour in a study by RPS Aquaterra (2011). The extent of flooding from 1 in 100-year floods was estimated in the study to be the 694 m contour.

Homestead Creek is the main drainage line that distributes flows onto Lake Robinson and the broader flats from northern sections of the catchment (AQ2 2016). Drainage across much of the catchment occurs over relatively flat areas with east-west ridge lines concentrating flows from the east and west (AQ2 2016). There are no official surface flow gauging stations within the catchment of the flats. Stations were established by BHP in 2011 and 2012, including on Homestead Creek and on Lake Robinson (AQ2 2016). A review by AQ2 (2016) of the data to 2014 was unable to conclude any relationship between flooding on Lake Robinson and the available data.

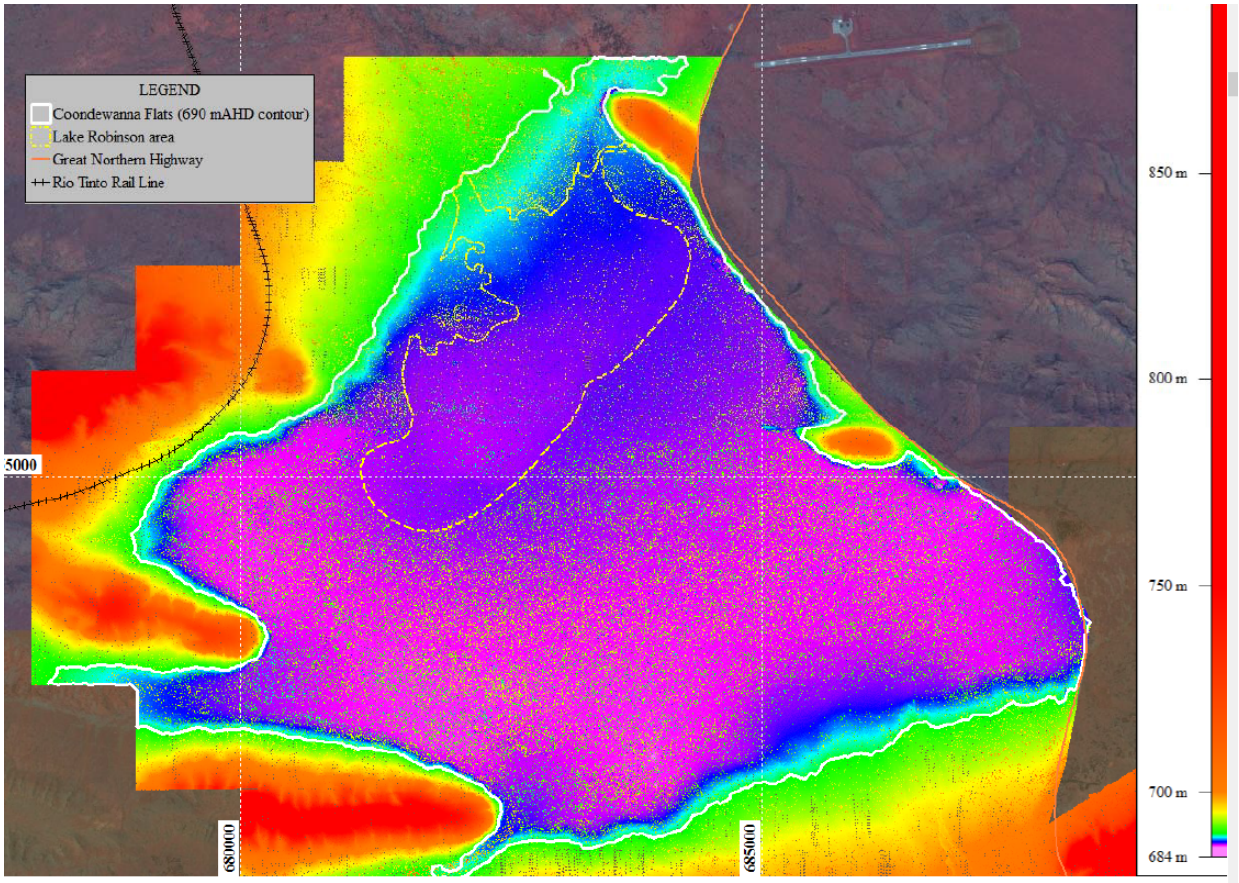


Figure 5: An elevation map of Lake Robinson and Coondewanna Flats based on LiDAR data and reproduced from AQ2 (2016).

2.1.3 Hydrogeology

The hydrogeology of the flats was reviewed by AQ2 (2016). The water table is situated at a depth of approximately 20 m below the surface where a layer of calcrete occurs and extends up to 16 m below the surface. The calcrete layer forms an unconfined aquifer. This layer is overlain by tertiary detritals of low to moderate permeability. Groundwater in the unconfined calcrete aquifer is recharged following heavy rainfall events of a magnitude that results in flooding on Lake Robinson. The water table can rise to within 15 m of the surface following heavy rainfall events and flooding and is most responsive in the north and western sections of the flats where inundation is more common. Surface inflow which does not result in groundwater recharge, will either recharge the unsaturated soil profile or evaporate from floodwaters.

2.1.4 Vegetation

Two sub-types of the PEC are recognised on the flats (Figure 1):

> *Priority 3: Coolibah and mulga (Acacia aneura) woodland over lignum and tussock grasses on clay plains (Coondewanna Flats and Wanna Munna Flats)*
>
> *Priority 1: Coolibah woodlands over lignum (Duma florulenta) over swamp wandiree (Lake Robinson is the only known occurrence).*

Eucalyptus victrix (coolibah) is a prominent tree species on the flats that is known to require groundwater and/or surface flows that increase soil moisture reserves (Pfautsch et al. 2011, Astron

Environmental Services 2014). The *E. victrix* trees on the flats are likely to be mostly dependent on the soil stores of moisture at depth, replenished by periodic inflows, rather than groundwater which occurs at approximately 20 m on average (Astron Environmental Services 2014, AQ2 2016). Prominent co-occurring trees species on the flats, *Acacia aptanuera* (mulga complexes), draw on water stored within surface soil layers, but are adapted to significant periods where these layers hold minimal stores of plant-available water.

Vegetation monitoring for BHP has been undertaken biannually at the flats since 2016, although no monitoring occurred in 2018 (Astron Environmental Services 2021). Prior to 2016, vegetation monitoring data were collected as part of studies into the ecological water requirements of the flats (Astron Environmental Services 2011, 2014). The program includes measurements of sample trees or plants across *E. victrix*, *A. aptanuera* and *Duma florulenta* and the analysis of high-resolution WorldView satellite imagery. No impacts to the vegetation that can be linked to BHP operations have been detected (Astron Environmental Services 2021). However, it has been noted that some *E. victrix* trees within Lake Robinson have declined and died within recent years, the causes of which are yet to be determined. Reductions in rainfall since 2016 and/or reductions in surface water inflows have been highlighted as possible factors (Astron Environmental Services 2022)

2.1.5 Landuse

The flats are located on Juna Downs pastoral station, which is owned and operated by a subsidiary of the Rio Tinto Group. BHP has current and future mining areas in proximity to the flats, including Mudlark and Tandanya to the west and South Flank and Mining Area C to the east. Potential impacts to the flats from these mining operations include groundwater drawdown, reductions in surface water inflows and groundwater rise from managed aquifer recharge activities. The BHP Juna Downs reinjection borefield north of the flats has been in operation since 2019 (BHP, unpublished communication). Abstraction from the BHP NPI borefield just west of the flats commenced in 2002 and has occurred sporadically since (BHP, unpublished communication).

2.2 Flooding History – Literature and High-resolution Imagery

A first-pass examination of the flooding history of the flats was compiled by examining available high-resolution imagery provided by BHP or held by Astron (Table 1), and by reviewing relevant reports. This information was used to provide further context for the study and to assist in evaluating the method of detecting flooding on the flats across the Landsat time series (see Sections 2.4, 2.5 and 2.5.2).

Table 1: High resolution imagery examined for the presence of flooding over Coondewanna Flats.

Source	Year	Month	Area (name of the image file)	Coverage across Coondewanna Flats
Aircraft	2004	August	Munjina and Governor	Complete*
	2008	April	Area C Infill	Partial
	2008	June	Area C Regional	Complete
	2009	January	Area C Jinayri Mine	Partial
	2010	December	South Flank	Partial
	2011	July	Area C Regional	Complete
	2012	June	South Flank	Partial
	2013^	Unknown	Central Pilbara~	Complete
	2014^	Unknown	Central Pilbara	Complete

Source	Year	Month	Area (name of the image file)	Coverage across Coondewanna Flats
	2015	May	Central Pilbara	Complete
WorldView Satellite^	2017	April	N/A	Complete
	2018	April	N/A	Complete
	2020	April	N/A	Complete
	2021	April	N/A	Complete

* Complete when both areas are combined together.

~ Area not specified but has the same extent as the Central Pilbara imagery.

^ Only those WorldView images captured at the end of the wet season (i.e. April and May) were examined.

2.3 Weather

Weather data were sourced from the SILO database (Queensland Government 2022) for Coondewanna Flats (-23.0°S, 118.8°E) for the 1961 to 2021 period. Annual data for variables were presented within 1 November to 31 October periods as this period encompasses a single wet season period (December to February). When a year is shown on a chart that depicts annual data, the second year within the 12-month period is quoted (e.g. 2021 is used for 2020-21) to simplify detail on the x-axis. When a year is mentioned in text, the two years are mentioned (e.g. 2020-21). A standard 1 January to 31 December period would encompass parts of two separate wet seasons which would complicate the interpretation of results.

2.3.1 Rainfall

Values for a range of rainfall variables were calculated. The following variables were calculated and presented in charts to describe rainfall since 1961:

- annual rainfall totals
- maximum rainfall for any calendar month within each year.

Additional variables were calculated for use as inputs for a flood prediction model. These variables were calculated across the 1987 to 2021 period, aligning with the available archive for consistent Landsat imagery suitable for flood detection. The variables were rainfall totals for periods prior to each day, the periods being one month, three months and 12 months.

2.3.2 Temperature

Annual maximum mean monthly temperature was plotted. Temperature trends provide additional description of conditions since 1961. Trends in maximum mean monthly temperature also have some relevance to flooding regimes and vegetation health given the influence of temperature on soil dryness and evaporative demand. However, in comparison to rainfall, any variation in trends in temperature are likely to have a smaller influence on flooding regimes and vegetation health.

2.3.3 Evaporation and Drought

One-month evaporation values prior to each day in the record from 1986-87 to 2020-21 were calculated to provide a potential input into the flooding prediction model.

A drought index was also calculated for the purposes of describing an element of the weather since 1961, as well as providing additional inputs for the flood prediction model. The KBDI, an index of surface soil-water deficit, was selected as the drought index (Keetch and Byram 1968). The KBDI is

commonly used to predict fires by estimating moisture levels in the surface soil and understorey layer (Department of Fire and Emergency Services 2014).

The KBDI calculates a daily index based on mean annual rainfall, maximum temperature for the day as well as the previous day's rainfall and KBDI value (Equation 1). The KBDI was calculated in R with a requirement of the code that the soil was maximally dry at the time of the first record (Williamson 2016, R Core Team 2022). Examination of rainfall records revealed that rainfall prior to the summer date of 1 January 1962 was minimal across a long period of time and therefore represented a suitable start point. As such, a KBDI of 200 mm was assigned to this date and the drought time series commenced from 1962 rather than 1961 as was the case for other weather variables.

$$KBDI = KBDI_{prev} + (0.001 \cdot \frac{(203.2 - KBDI_{prev}) \cdot (0.968 \cdot e^{(0.0875 \cdot Max\ Temp)+1.552} - 8.3}{10.88 \cdot e^{-0.001736 \cdot Average\ Rainfall} + 1}$$

Equation 1: The Keetch-Byram Drought Index (KBDI) expressed in mm where KBDIprev = previous record for KBDI. Max Temp = maximum temperature.

Annual average KBDI was plotted from 1962-63 to 2020-21. In addition, the one-monthly average KBDI was calculated and used as inputs for flood prediction modelling.

2.3.4 Cyclones

Cyclone occurrence by year was summarised using records from the Bureau of Meteorology (2023). These records are available for the period commencing from 1969-70. Only those cyclones that tracked within 200 km of the flats were considered to be occurrences.

2.3.5 Fire

The available fire history of the flats is presented to assist with the interpretation of vegetation trends. Fire regimes have an important influence on vegetation structure and composition. Fire can result in a significant reduction in vegetation cover for several years following an intense fire. Post-fire recovery of foliage can occur predominantly by resprouting in some species (e.g. *E. victrix*) or predominantly by seedling regeneration in others (e.g. *A. aptanuera*) with the timing and pace of recovery determined by the severity of the fire and rainfall in the period following the fire.

A history of fires at the flats since 2000 was sourced from the North Australia and Rangelands Fire Information (NAFI) website (NAFI 2022). The data are acquired from heat sensors for five satellites: Terra, Aqua, NOAA-19, Suomi NPP and the JPSS-1 (or NOAA-20) satellite. Available data from NAFI begins from 2000.

2.4 Satellite Imagery for Time Series Analysis

To construct a reliable history of flood and vegetation cover dynamics, a long time series of data from a consistent imaging sensor is required. Satellites in the Landsat mission were selected on the basis that imagery has been captured every 16 days since 1986 from sensors with comparable specifications, despite changes in the platform from Landsat 5, Landsat 7, Landsat 8 and the recently launched Landsat 9. Imagery has been captured from Landsat 5 to Landsat 9 across 7 consistent bands, including near infrared (NIR) bands, which are valuable for vegetation studies, and short-wave infrared (SWIR) bands which are valuable for surface water studies. Another advantage of Landsat is that imagery is available in an analysis-ready format (atmospherically and geometrically corrected) within the Digital Earth Australia (DEA) catalogue hosted on the National Computational Infrastructure (NCI) facility.

Landsat (25 m spatial resolution) imagery was acquired from the DEA catalogue. The NCI provides a mask layer with all returned imagery that provides an indication of each pixel's quality within an image. This layer can be used to mask invalid pixels such as pixels containing cloud. A total of 795 images were downloaded from the NCI for the period 1986 to 2021 with 665 of these from Landsat 5 or Landsat 7 images and 130 from Landsat 8 images. All images had at a maximum cloud cover of 20%.

From May 2003, the scan line corrector (SLC) failed on the Landsat-7 satellite causing large bands of no data to be captured in all images from this date. The regular data gaps influence about 25% of the scenes with increasing width towards the edges (Pringle et al. 2009). As the flats are located toward the middle of the Landsat swath (path: 112, row: 76) the influence of the missing data is very minor. In the case of water detection where the data are binary (a pixel is either water or it is not), missing data have been estimated by interpolation. However, for vegetation health data, interpolation was not considered appropriate and scan lines were simply treated as missing data.

2.5 Water Detection from Landsat Imagery

Many water detection indices exist, but often, these are developed for application to the detection of large waterbodies or inundation in moderate and high rainfall zones (Ogilvie et al. 2018). This highlights the need for water indices to be tested for suitability within a particular location before they are applied to a particular study. Testing needs to also consider that accurate detection of water within one image may not guarantee accurate detection for a different image captured at a different point in time for the same study area (Ogilvie et al. 2018). This can be due to factors such as cloud cover, shadows and the SLC fault (Ogilvie et al. 2018). Full and thorough testing in the case of the study of the flats was somewhat hampered by the ephemeral flooding regime on the flats, the absence of any permanent water bodies within the general area and the scarcity of high-resolution imagery coinciding with flooding on the flats.

Several indices were tested prior to selecting one index to be applied to water detection across the Landsat time series (Table 2). Of particular note, is the study by Herndon et al.(2020) which found the Normalised Difference Moisture Index, Tasselled Cap Wetness Index (TCW) and Automated Water Extraction Index to be most suitable in detected ephemeral water bodies in an arid environment. The study by Ogilvie et al. (2018) found the Modified Normalised Difference Water Index (MNDWI) to outperform other water indices, such as the Normalised Difference Water Index, Normalised Difference Vegetation Index and Normalised Difference Turbidity Index at detecting ephemeral water bodies in an arid environment.

Indices were tested by comparison of outputs against one confirmed flood in 2012, the presence of water and indications of recent minor flooding across a small area (< 10 ha) within the northern section of Lake Robinson in aerial imagery of 2008 and periods where the flats were known to be dry based on on-ground observation and high-resolution WorldView imagery.

Table 2: Water detection indices tested for the study on Coondewanna Flats and the threshold value that indicates the presence of water. Equations of all water detection methods are provided except for the Cordeiro Index in which a Python package designed specifically for this index was used (Cordeiro et al. 2021).

Water Detection Method	Threshold	Citation	Equation
Automated Water Extraction Index (AWEI)	> 0	Zhai et al. (2015)	BLUE + 2.5 × GREEN − 1.5 × (NIR+SWIR1) − 0.25 × SWIR2
Cordeiro Index (CI)	= 0	Cordeiro et al. (2021)	N/A
Fishers Water Index (WI)	> 0	Fisher et al. (2016)	1.7204 + 171(GREEN) + 3(RED) + 70(NIR) + 45(SWIR1) + 71(SWIR2)
Modified Normalised Difference Water Index (MNDWI)	> 0	Xu et al. (2006)	(GREEN − SWIR)/(GREEN + SWIR
Normalised Difference Moisture Index (NDMI)	> 0	Herndon et al. (2020)	(NIR − SWIR)/(NIR + SWIR)
Normalised Difference Pond Index (NDPI)	< 0	Subramaniam et al. (2011)	(SWIR − GREEN)/(SWIR + GREEN)
Normalised Difference Turbidity Index (NDTI)	> 0	Subramaniam et al. (2011)	(RED – GREEN)/(RED + GREEN)
Normalised Difference Vegetation Index (NDVI)	< 0	Rouse et al. (1973)	(NIR − RED)/(NIR + RED)
Normalised Difference Water Index (NDWI)	> 0	McFeeters (1996)	(GREEN − NIR)/(GREEN + NIR)
Simple Water Index (SWI)	> 0	Acharya et al. (2019)	1/ √(BLUE − SWIR1)
Tasselled Cap Wetness Index (TCW)	> 0	Baig et al. (2014)	0.1511 × BLUE + 0.1973 × GREEN + 0.3283 × RED +0.3407 × NIR − 0.7117 × SWIR1 − 0.4559 × SWIR2
Water Ratio Index (WRI)	> 0	Acharya et al. (2019)	(GREEN + RED)/(NIR+ SWIR)

Notes

- NIR = near infra-red
- SWIR = short-wave infra-red with 1 or 2 referring to the SWIR band number.

With the exception of the Cordeiro Index (CI), TCW and Water Ratio Index (WRI) (Acharya et al. 2019, Herndon et al. 2020, Cordeiro et al. 2021), all indices accurately detected water on the flats. From the literature, it is not known if the CI and the WRI have ever been successfully used in arid environments.

Across the literature, MNDW was highlighted as an index suited to more arid areas and shallow waterbodies (Sarp and Ozcelik 2017, Dzinotizei et al. 2018, Ogilvie et al. 2018) (Equation 2). In addition, MNDWI is purported to minimise noise from other features such as vegetation and soil (Xu 2006). Visual assessment, whereby outputs were compared to visible water in the imagery, indicated that MNDWI performed well in detecting the presence-absence of water on the flats as well as the extent of water when present. Values for MNDWI range from -1 to +1, where positive values are associated with water and negative values are associated with soil, vegetation and other surfaces.

$$MNDWI = \frac{GREEN - SWIR1}{GREEN + SWIR1}$$

Equation 2: The Modified Normalised Difference Water Index (MNDWI) based on Landsat spectral band data: GREEN = reflectance from the green band and SWIR1 = reflectance from short-wave infrared band 1.

2.5.1 Flood Event Detection

In order to distinguish significant inundation (flood events) from minor inundation (small and or patchy wet areas), a set of assumptions were devised around what constitutes a flood event. A flood was defined as event where greater than 10 ha of water in total was detected across the flats. Furthermore, sequential 'wet' records (i.e. records where the area classified as wet was greater than or equal to four Landsat pixels across the flats, i.e. 0.25 ha) after the flood event were also considered part of the flood event. A 'dry' record (i.e. records where the area classified as wet was less than four Landsat pixels across the flats, equal to 0.25 ha) marked the end a flood event.

These assumptions were necessarily subjective, based on Astron's experience, because there is no one definition of what constitutes a flood event. For example, consider a heavy rain event that involves surface ponding over an area for several hours, or an inflow that results in a small pool covering 1% of the area of interest: would these constitute floods, and for how long or across what size area would water need to be present to be classed as a flood event? The answers to these questions are not clear. The suitability of these assumptions were assessed in the visual assessment of outputs from the water detection indices against surface water present in imagery.

2.5.2 Flood Event Duration

A flood event was classified as all instances in which the area of water detected across the flats was at least 10 ha. The duration of each flood event was calculated as the number of days between the midpoints of records (cloud-free Landsat images) so long as the records were at least classified as wet (i.e. water detected across 0.25 ha or more). Dry records (i.e. water absent or detected across < 0.25 ha) reset this flood duration to zero (Table 3).

Table 3: Hypothetical examples of flood, wet and dry events based on the mid-points between wet and dry records from remote sensing image captures.

Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Image capture	**Yes**	No	No	**Yes**	No	**Yes**	No	No	**Yes**	No	No	**Yes**	No	No	**Yes**
Area of water detected (ha)	**0.2**	N/A	N/A	**12.1**	N/A	**1.4**	N/A	N/A	**0.1**	N/A	N/A	**0.7**	N/A	N/A	**0.15**
Record class	**DRY**	None	None	**WET**	None	**WET**	None	None	**DRY**	None	None	**WET**	None	**None**	**DRY**
Event	**Dry**		**Flood**					**Dry**			**Wet**			**Dry**	

Notes:

- Dry = water detected across an area < 0.25 ha
- Wet = water detected across an area ≥ 0.25 ha
- Flood event occurs when at least one wet record ≥ 10 ha

2.5.3 Flood Prediction Models

A model was developed to predict the occurrence of a flood event, or more specifically, the start of flood event. Predictor variables included the following:

- Time since flood – the number of days since the end of a previous flood event
- 12-month rainfall – the total rainfall 12 months prior
- 3-month rainfall – the total rainfall three months prior
- 1-month rainfall – the total rainfall one month prior
- 1-month evaporation – mean evaporation rate one month prior
- 1-month KBDI – mean KBDI value one month prior.

The inclusion of Time since flood necessitated the exclusion of data for the first flood event in the development of the model.

To enable the prediction of the commencement of discrete flood events, records in the data set had to first be classified as either:

1. Flood start (i.e. the record on day 4 using the example in Table 3)
2. Flood event, but not the flood start (i.e. the record on day 6 using the example in Table 3)
3. Dry record (i.e. the records on days 1, 9 and 15 using the example in Table 3).

Records that were 2) *flood event, but not the start*, were removed to ensure that only discrete, independent floods were predicted.

A Random Forest (RF) ensemble learning method was used to develop the flood occurrence model. The RF method is a supervised classification technique that combines a large number of individual decision tree models. By combining high-variance, low-bias decision trees, a robust low-variance and low-bias model can be developed (Breiman 2001). The sklearn package in Python was used to develop RF classification models.

To validate the model, 10% of the records consisting of Flood start and Dry records were removed from the training data used for creating the random forest model and instead were used to validate the model as testing data. As the flood events were removed from the input data into the model, only Flood start and Dry records were used. Therefore, a positive record is the start of a Flood and a negative record is a Dry record. True positives are determined from the number of instances in which the model has detected a Flood start, and the testing data record is also a Flood start. True negatives were determined the same way, except only Dry records are considered. False negatives are the number of instances in which the model detected a Dry record, but the record was actually a Flood start.

The accuracy of the model is an indication of how well the model is able to detect a Flood start or Dry record, with higher values indicating the model is able to correctly classify a Flood Start or Dry record. The F1 score is an indication of how often the model wrongly classifies Flood Starts or Dry records, with higher values indicating the model is not overclassifying either Flood Starts or Dry records.

For this analysis, a total of 300 different RF trees were considered for all predictor variables (as listed in Section 2.5.3). The accuracy and the F1 score are both quoted for the RF model. Precision, recall, accuracy and F1 score are shown from Equation 3 to Equation 6.

$$Precision = \frac{True\ Positives}{True\ Positives + True\ Negatives}$$

Equation 3: Precision of classification model.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

Equation 4: Recall of classification model.

$$Accuracy = \frac{True\ Positives + True\ Negatives}{True\ Positives + True\ Negatives + False\ Positives + \ False\ Negatives}$$

Equation 5: Accuracy of classification model.

$$F1\ Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

Equation 6: F1 Score of classification model.

Due to the RF algorithm being a classifier rather than a binary predictive model, significant variables are assigned importance values (between 0 and 1), with higher numbers signifying greater importance to the model results.

Models were run firstly with all predictor variables. Progressively, predictor variables of lesser importance were removed, and models rerun. Results were compared amongst all models with assessment of the most appropriate model made based on the following elements combined: accuracy, F1 score, and model simplicity.

2.5.4 Limitations in the Use of Landsat for Flood Regime Description

Some flood events may have been missed or the duration and extent of flood events may be inaccurate (most likely overestimated) due to discontinuity in the Landsat image time series. Landsat satellites can potentially capture imagery at a specified location every 16 days. Since the launch of Landsat 7 in 1999, more than one Landsat satellite has been in operation, and this has reduced the potential image capture frequency to < 16 days. In addition to satellite return interval, the presence of clouds in imagery can further reduce the frequency of effective image captures. This is particularly problematic for flood detection given that dense cloud and flood events co-occur.

The number of cloud free image captures for the flats during the study was 795. This represents an average of one cloud-free image per 16 days. The longest period across which no suitable imagery was available was 112 days which occurred between 24 May 1987 and 13 September 1987 as well as between 7 February 1995 and 30 May 1995. Other extended periods with no suitable imagery being available were 3 September 1995 to 22 November 1995 (80 days) and 13 February 2000 to 19 May 2000 (96 days). Since May 2000, consecutive images have been captured within in two months or less of one another. The majority of images (753 images) were captured within one month of the previous capture.

2.6 Vegetation Detection and Condition

Vegetation indices such as Modified Soil-Adjusted Vegetation Index (MSAVI) (Equation 7) enable vegetation to be distinguished from other features and for quantitative data on vegetation health to be extracted. The MSAVI is well suited to use in arid areas as bare soil accounts for a large proportion of the visible surface and MSAVI minimises the contribution of soil reflectance (Qi et al. 1994). Higher

MSAVI values are associated with the presence of live vegetation rather than dead vegetation, soil or other features. Furthermore, where vegetation is present, higher MSAVI values are associated with denser and healthier foliage. As such, MSAVI was selected as a measure of vegetation condition, i.e. the cover and health of vegetation.

$$MSAVI = \frac{2 \cdot NIR + 1 - \sqrt{(2 \cdot NIR + 1)^2 - 8 \cdot (NIR - RED)}}{2}$$

Equation 7: Modified Soil Adjusted Vegetation Index (MSAVI) formula. NIR = near infrared reflectance and RED = red reflectance.

Time series data for MSAVI were produced from Landsat data for 1986-87 to 2020-21. Charts were produced separately for the two sub-types of the PEC on the flats, with values for areas where *E. victrix* and *A. aptanuera* canopy is present produced separately from vegetation as a whole (inclusive of all trees and understorey species). A separate focus on the two tree species was undertaken because these trees are a defining feature of flats and are likely to be sensitive to alterations in surface flow regimes.

The analysis focused on the two tree species and used an existing layer of classified canopy polygons to determine whether *E. victrix* and/or *A. aptanuera* canopy was present or absent within each Landsat pixel. The existing layer was developed from high resolution WorldView imagery using segmentation and classification as part of the Coondewanna Flats vegetation monitoring program (Astron Environmental Services 2021). Where a Landsat pixel intersected an *E. victrix* polygon, the MSAVI value for that pixel was considered to be an MSAVI value for *E. victrix*. The same applied to *A. aptanuera*. A limitation of this approach is the influence of soil and vegetation other than the tree species of interest to the MSAVI value (i.e. spectral mixing), especially where segments for a species of interest comprise a small proportion of a single Landsat pixel. Additionally, the MSAVI value of any pixel containing both tree species will be assigned to both *E. victrix* and *A. aptanuera*. An additional limitation relates to the fact that the classification of *E. victrix* and/or *A. aptanuera* canopy was developed in 2021 but applied to all years within the Landsat time series under the assumption that the occurrence of these species in 2021 is the same across the time series back to 1987.

The mean MSAVI values from all available cloud-free images (1986-87 to 2020-21) for *E. victrix, A. aptanuera* and all vegetation were displayed for each PEC as a line chart with monthly rainfall totals. Separate charts were also produced for these variables with data summarised to mean annual values for MSAVI and mean total values for rainfall. Flood events were also identified on these charts. Trends in MSAVI were interpreted in relation to rainfall, floods, fire and other relevant data.

3 Results and Discussion

3.1 Weather History (1960-61 to 2020-21)

The annual rainfall record since 1960-61 is characterised by extreme variability with a 10-fold difference between the highest (960 mm in 1999-2000) and the lowest annual total (91 mm in 1976-77) (Figure 6). There has been an increasing trend for annual rainfall at the flats between 1960-61 and 2020-21. Annual totals for the first half of the 60-year period are mostly clearly below the long-term average (355 mm) whereas annual totals in the second half of the record are mostly around or above the long-term average.

A trend of increase across the 1960-61 to 2020-21 period is also evident in maximum monthly rainfall totals (Figure 7). Similar to the annual rainfall record, a seven-fold difference between extremes is also present within the maximum monthly values (315 mm in 1994-95 versus 43 mm in 1982-83).

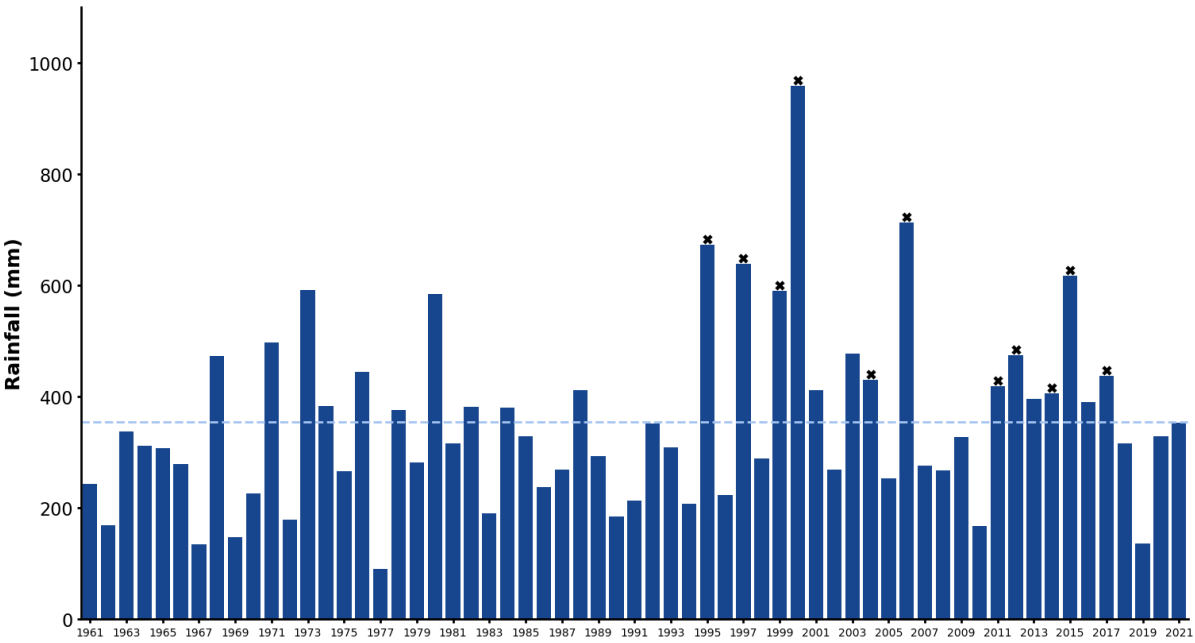


Figure 6: Annual rainfall (years represented from 1 November to 31 October) at Coondewanna Flats (-23.0°S, 118.8°E). Data is sourced from the Scientific Information for Land Owners (SILO) database (Queensland Government 2022). The blue dotted line denotes the annual average rainfall (1960-61 to 2020-21). The asterisks represent years of significant floods detected from the remote sensing analysis (1987 to 2021).

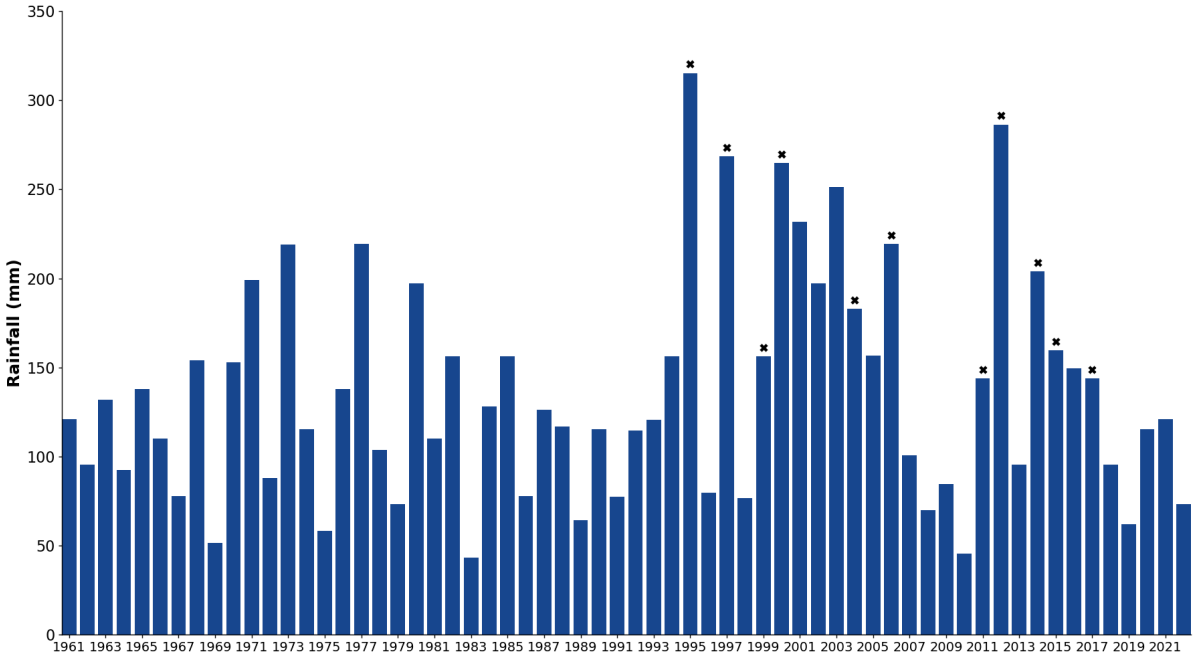


Figure 7: Maximum rainfall total recorded within a month (years represented from 1 November to 31 October) at Coondewanna Flats (-23.0°S, 118.8°E). Data is sourced from the Scientific Information for Land Owners (SILO) database (Queensland Government 2022). The asterisks represent years of significant floods detected from the remote sensing analysis (1987 to 2021).

A warming trend is clearly apparent in the 60 years to 2020-21 (Figure 8). From 1991-92, around half of all mean maximum monthly temperatures were higher than any recorded prior to this date. There is a degree of negative correlation between annual rainfall and mean maximum monthly temperature. A reduction in insolation due to persistent cloud coverage in wet years is one likely reason underlying this association.

The long-term trend in the mean annual value for the drought index (KBDI) is slightly negative (Figure 9). This reflects the increasing rainfall trend over the period as highlighted in Figure 6, and the fact that higher rainfall has not been offset by higher evaporation driven by rising temperatures (Figure 8). As with other weather variables, considerable interannual variation is present within the time series for mean annual KBDI.

The minimum mean monthly KBDI values were generally lower in the second half of the period from 1961-62 to 2020-21 (Figure 10). Low monthly values reflect periods of high moisture availability. The periods from 1996-97 to 2005-06 and 2010-11 to 2016-17 were associated with a succession of years where surface soil was likely to be at, or near, a state of saturation.

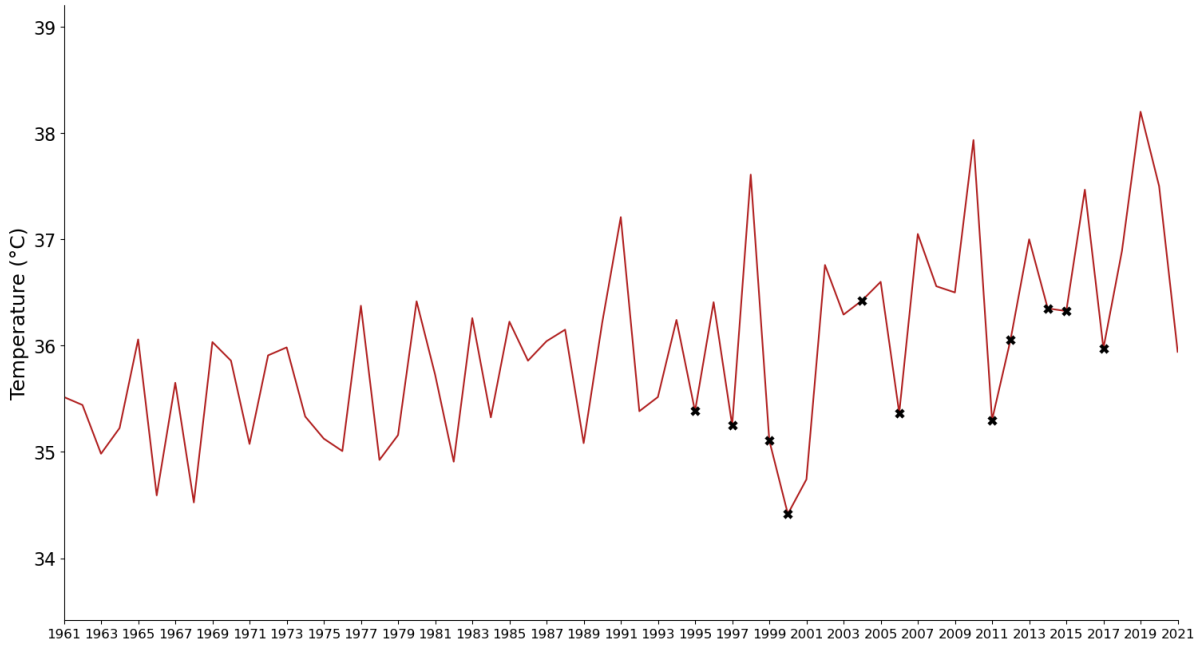


Figure 8: Mean maximum monthly temperature (years represented from 1 November to 31 October) at Coondewanna Flats (-23.0°S, 118.8°E). Data is sourced from the Scientific Information for Land Owners (SILO) database (Queensland Government 2022).The asterisks represent years of significant floods detected from the remote sensing analysis (1987 to 2021).

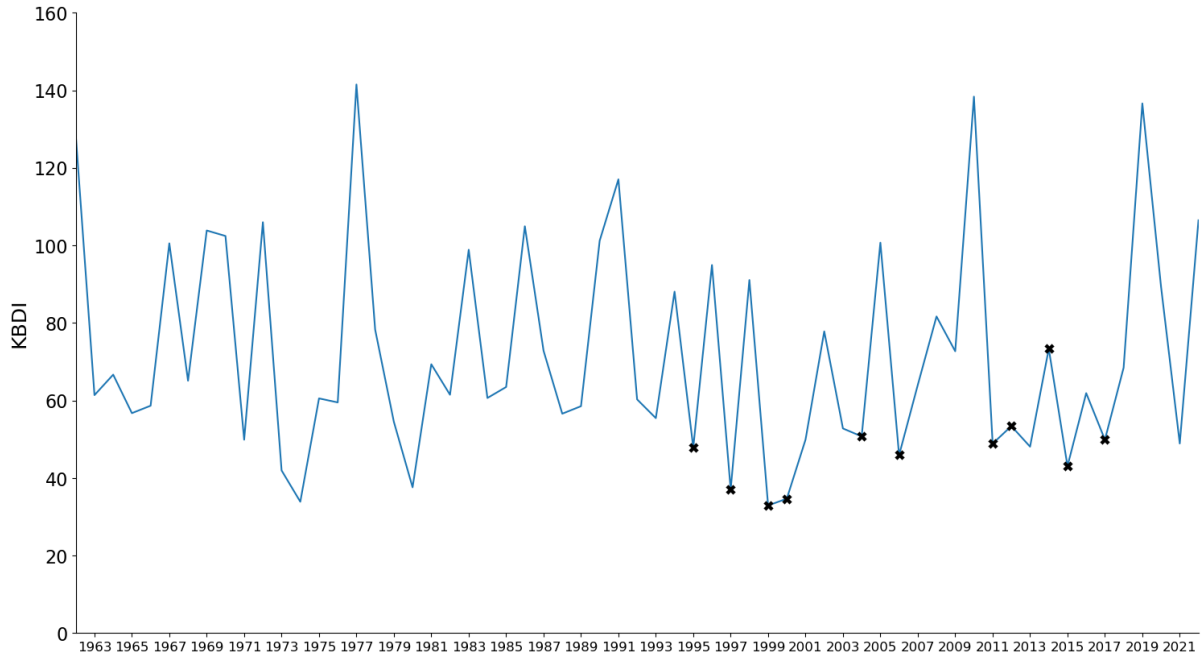


Figure 9: Mean annual value for the Keetch-Byram drought index (KBDI) in mm (years represented from 1 November to 31 October) at Coondewanna Flats (-23.0°S, 118.8°E). Data is sourced from Scientific Information for Land Owners (SILO) database (Queensland Government 2022).The asterisks represent years of significant floods detected from the remote sensing analysis (1987 to 2021).

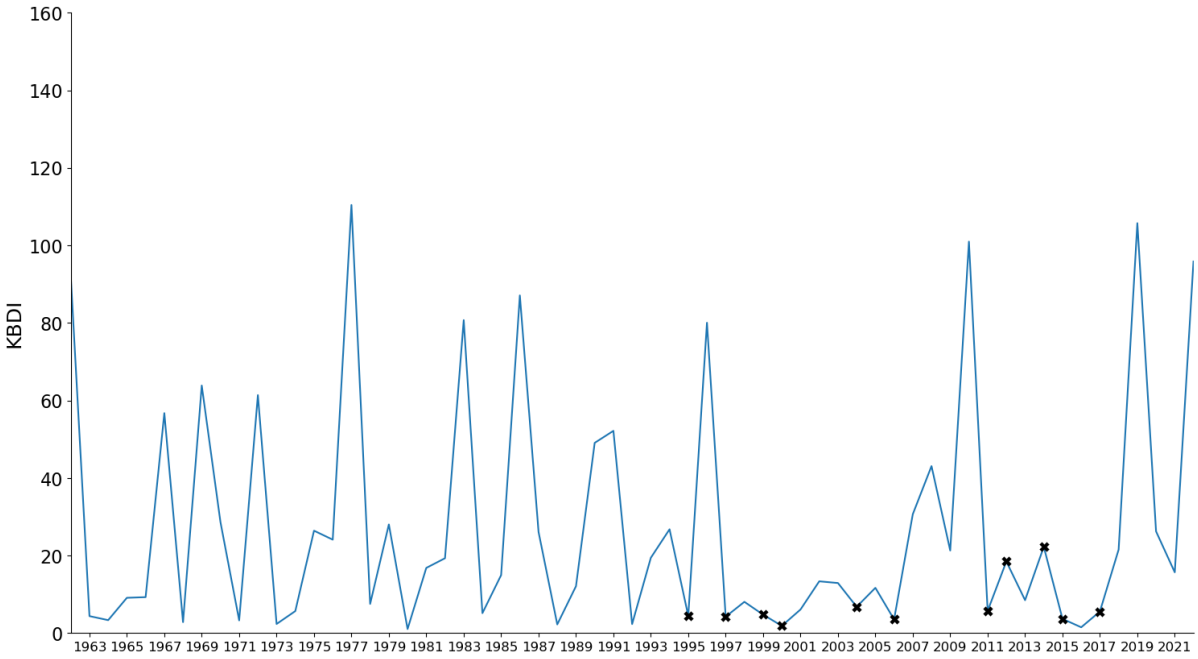


Figure 10: Minimum mean monthly value for the Keetch-Byram drought index (KBDI) in mm within a year (years represented from 1 November to 31 October) at Coondewanna Flats (-23.0°S, 118.8°E). Data is sourced from the Scientific Information for Land Owners (SILO) database (Queensland Government 2022). The asterisks represent years of significant floods detected from the remote sensing analysis (1987 to 2021).

3.2 Cyclones

There were 32 cyclones that passed withing 200 km of the flats in the 53 years from 1969-70 (Table 4). This equates to an average frequency of approximately one cyclone every two years. Multiple cyclones were recorded within five of the years and the maximum number of consecutive years across which no cyclones passed in the vicinity of the flats was four years; there were two such periods since 2006-2007. Most of the years where annual rainfall totals was well-above average (Figure 6), co-occurred with cyclone years. However, there were exceptions, such as 1994-95 where a flood and the third highest annual rainfall total since 1961 was recorded, yet no cyclone passed within 200 km of the flats.

Table 4: Records of cyclones passing within 200 km of Coondewanna Flats (Bureau of Meteorology 2023). The cyclone season spans the period from November to April.

Year	Names	No.	Year	Names	No.
1969-70	-	0	1996-97	Rachel	1
1970-71	Sheila	1	1997-98	-	0
1971-72	-	0	1998-99	-	0
1972-73	Kerry	1	1999-00	John	1
1973-74	-	0	2000-01	Wylva	1
1974-75	-	0	2001-02	Chris	1
1975-76	Joan	1	2002-03	01U	1
1976-77	-	0	2003-04	Monty	1
1977-78	Vern	1	2004-05	-	0
1978-79	Rosa	1	2005-06	Emma, Frank	2
1979-80	Amy, Dean, Enid	3	2006-07	George	1

Year	Names	No.	Year	Names	No.
1980-81	Neil	1	2007-08	-	0
1981-82	-	0	2008-09	-	0
1982-83	-	0	2009-10	-	0
1983-84	-	0	2010-11	-	0
1984-85	Emma, Frank, Gertie	3	2011-12	Heidi, Lua	2
1985-86	-	0	2012-13	-	0
1986-87	Connie	1	2013-14	Christine	1
1987-88	-	0	2014-15	-	0
1988-89	Ilona, Orson	2	2015-16	Stan	1
1989-90	-	0	2016-17	AU201617_22U	1
1990-91	-	0	2017-18	Joyce	1
1991-92	Ian	1	2018-19	-	0
1992-93	-	0	2019-20	-	0
1993-94	-	0	2020-21	-	0
1994-95	-	0	2021-22	-	0
1995-96	Oliva	1	**Total**		**32**

3.3 Fire History

From the available fire records, four fires have occurred at the flats between 1998-99 and 2020-21 (Figure 11). This represents a frequency of one fire every eight years. However, the average fire frequency at a local scale within the flats is much lower given that three of the four recorded fires burnt < 50% of the area of the flats. The most extensive fires on the flats, recorded in 1999-2000 and 2006-07, correspond to the two years with the highest annual rainfall totals between 1960-61 and 2020-21 (Figure 6). This highlights the role of extremely wet periods in promoting understorey growth which translates to higher and more extensive fuel loads.

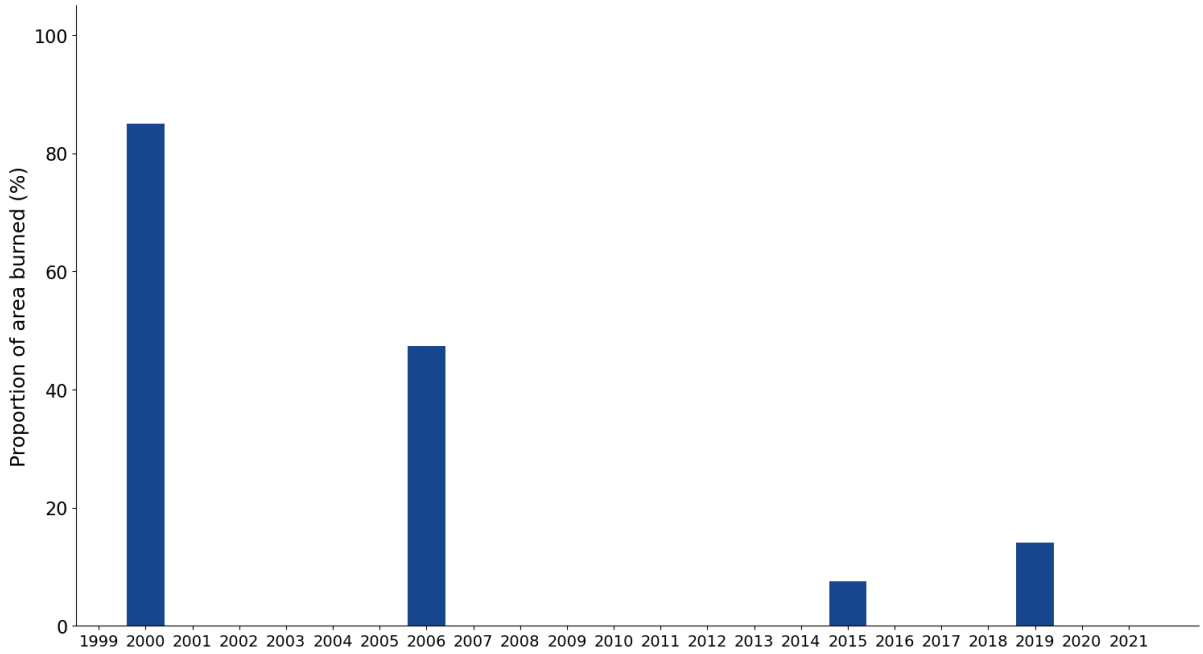


Figure 11: Fire history for Coondewanna Flats as a proportion of area (Priority Ecological Community boundary) burnt by year (years represented from 1 November to 31 October) based on Northern Australia Fire Information (2022). Data are available from 1999 only.

3.4 Flooding History

3.4.1 Literature Review and Visual Examination of High-Resolution Imagery

Limited available literature highlights two observed and two highly likely flood events since the 1970s (Table 5). Visual examination of available high-resolution imagery (Table 1) did not reveal any significant flood events at the time of image capture. However, the aerial image from April 2008 did show water within the main channel at the northern end of the flats with evidence of clearly defined wet soil extending across several hectares onto Lake Robinson. This wet area may be an indication of minor flooding prior to the date of image capture.

Table 5: Records of flooding or potential flooding on Coondewanna Flats (CF): observed or inferred.

Year (month if known)	Notes	Source
1975 (December)	Flooding observed for several months after passage of Cyclone Joan. Persisted for many months.	(RPS Aquaterra 2011, AQ2 2016)
2002-03	Potential flooding – inferred.	(AQ2 2015)
2006	Potential flooding – a year in which groundwater level rise was significant.	(AQ2 2015, BHP 2021)
2012 (January)	Flooding observed from January and persisted until beyond March. Note that BHP (2021) lists 2011 as the last year that flooding was observed – this is probably an error and 2012 is the correct year, unless 2011 refers to the 2011-12 wet season.	(AQ2 2016)

3.4.2 Landsat Image Analysis

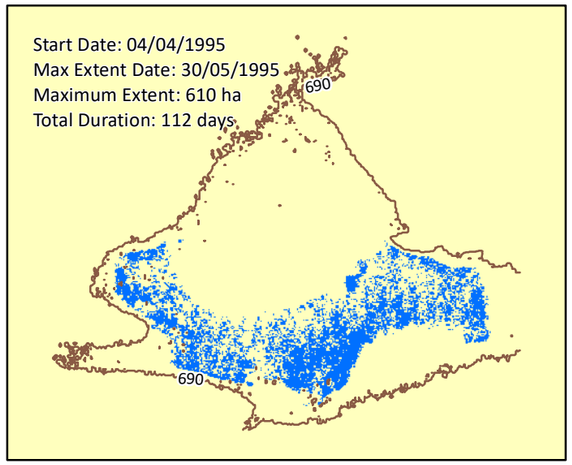
There were 11 flood events (i.e. floods greater than 10 ha in extent) detected on the flats between 1986-87 and 2020-21 using the MNDWI index applied to the available Landsat archive (Figure 12). One apparent flood event of 11 ha in September 2009 was excluded from the final list, given that this event was much smaller in extent and occurred outside the months within which other flood events occurred. Furthermore, rainfall preceding the apparent flood was low with monthly totals < 10 mm for April, May, July and August, with June receiving 30 mm. The identification that this event was a false positive provides some support for the initial 10 ha estimate of inundation as a threshold defining a flood event.

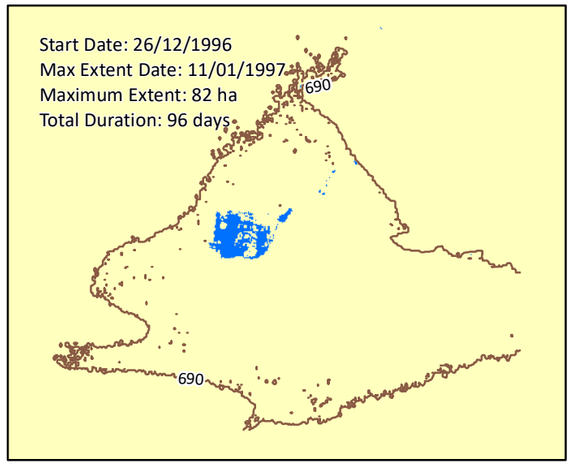
The total of 11 flood events across the period equates to a flood frequency of approximately once every three years. This is more frequent than the estimate of four years quoted in AQ2 (2015) and BHP (2021). The flood events detected in the Landsat analysis for 2013-2014 and 2014-2015 were not noted in AQ2 (2015, 2016) and the flood event for 2016-2017 was not noted in BHP (2021). Detection of smaller and short-lived floods is often more likely to be achieved with Landsat imagery (if cloud is not persistent) than from on-ground observation or inferences from a limited network of point-based surface flow or groundwater sensors. On-ground observations require an observer to be present on the flats at the right time and in the right location and then for the observation to be reported. Similarly, sensors must be positioned at locations that align with inundation for a flood record to be made; only a limited number of locations are instrumented on the flats. Despite the frequency of flood events at one in three years for the 1986-87 to 2020-21 period, the frequency is likely to be lower across the 1960-61 to 2020-21 period given that the period prior to 1986-87 was drier with fewer large rainfall events (Figure 6 and Figure 7).

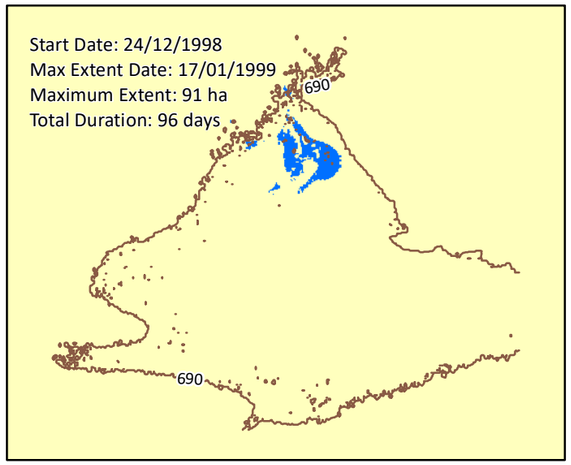
The anecdotal record of a flood in 2002-03 reported by AQ2 (2016) was not detected by the Landsat analysis. A lower-than-average number of cloud-free images were available for this period and this could account for the discrepancy. In the absence of any on-ground data it is difficult to draw any firm conclusions as to whether flooding occurred in this year or not.

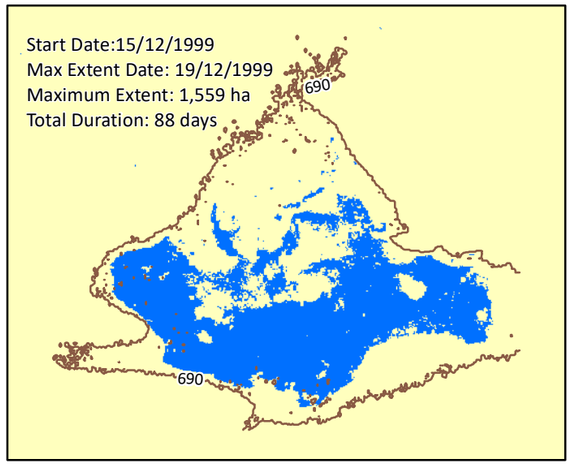
There was considerable variation in the duration and maximum extent of flood events (Figure 12). The duration ranged from 20 days (February 2011) to 112 days (April 1995) (Figure 12). Maximum flood extent ranged from 14 ha (February 2004) to 2,290 ha (February 2012). The expected positive association between duration and extent was not as close as expected. This is likely due in part to measurement error arising from gaps in the image archive, particularly where cloud was persistent prior to and during flood events.

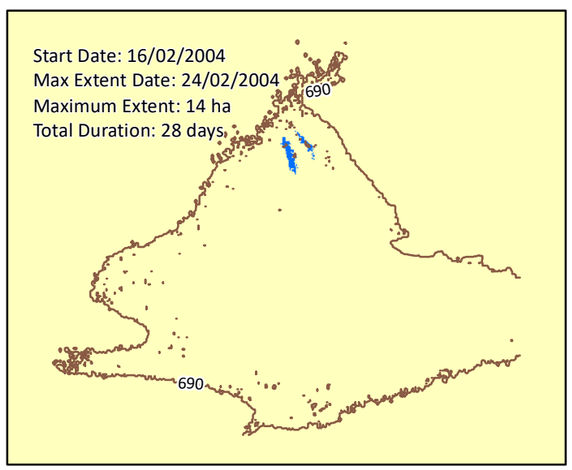
The location of maximum flood extents differs between small flood events and large flood events, with minimal overlap between extremes (Figure 12). While minor flood events show inundation on Lake Robinson, major floods show minimal if any flooding at this location at the time of maximum flood extent. South of Lake Robinson, elevation on the flats is as low as on the lake (Figure 5). It is highly likely that inundation occurs on Lake Robinson during extensive floods, but that inundation water then either flows overland, infiltrates the soil, or is evaporated in many instances by the time of maximum flood extent. The absence of surface water on Lake Robinson while water is present elsewhere on the flats can also be seen in the photograph of flooding in January 2012 (Figure 13). Further confirmation that the pattern of inundation detected in February 2012 was accurate was obtained by comparison of the results from other water detection indices and visual examination of the February 2012 Landsat imagery displayed in natural colour and false colour infrared (FCIR) (Figure 14). The FCIR image of February 2012 clearly shows the extent of water in blue.

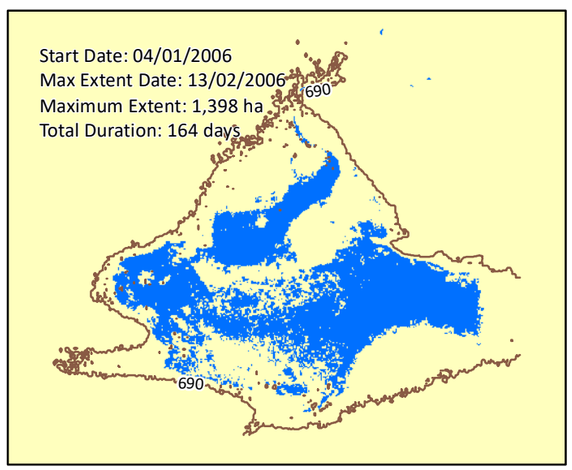


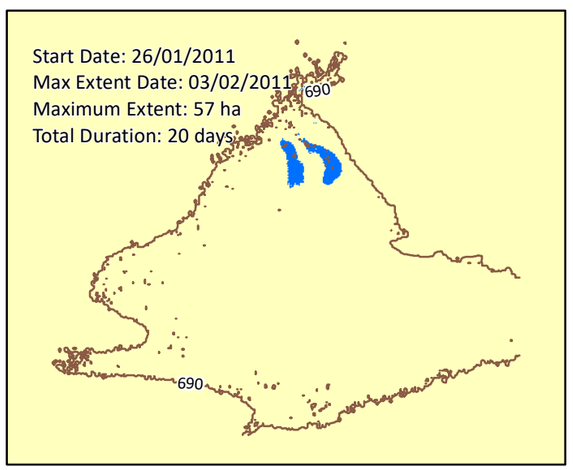


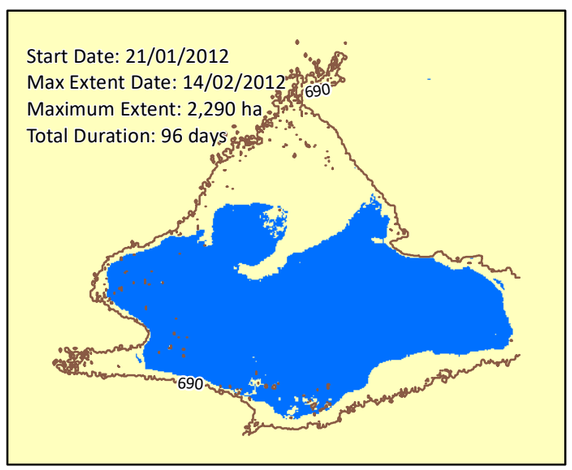


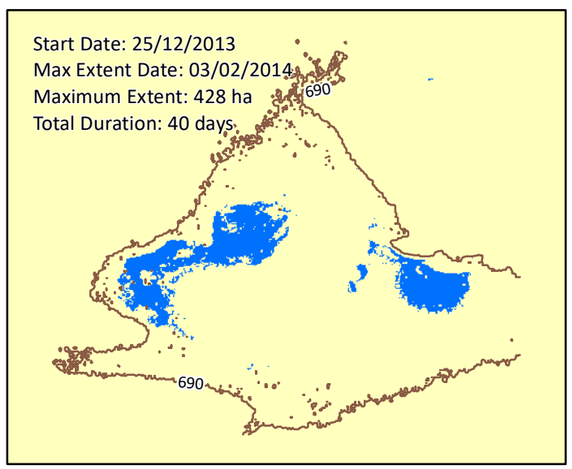


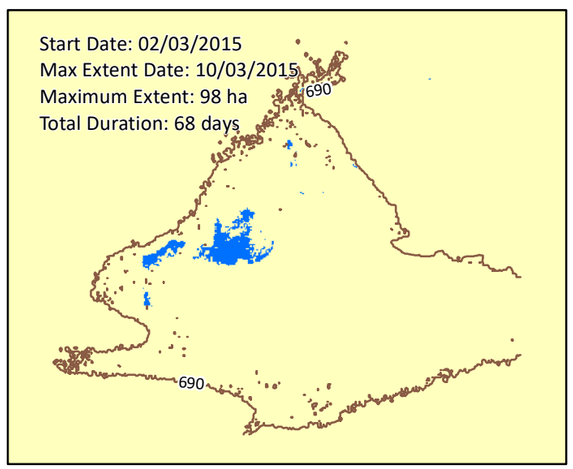


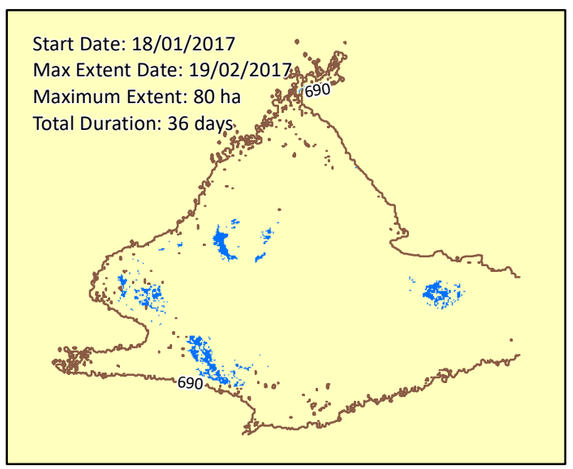


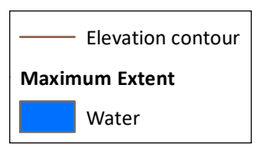












BHP

Coondewanna Flats Remote Sensing History

Figure 12: Maximum flood extent by year where floods greater than 10 ha were detected.

Scale: 1:130,000 at A3

Coordinate System: GDA 1994 MGA Zone 50

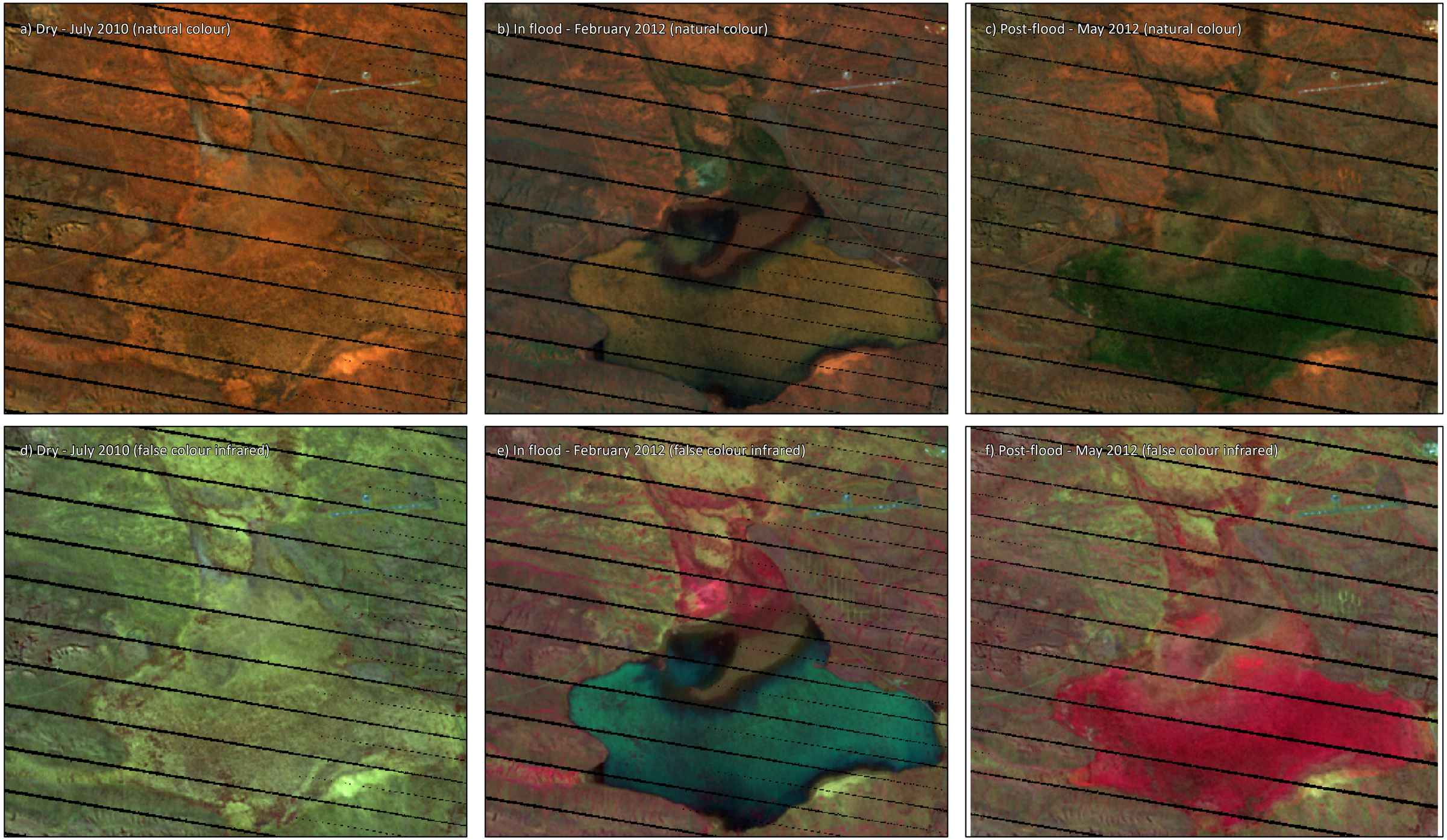
0 2 4 6 Kilometres

N

astron delivering environmental intelligence™

Author: R. Archibald | Drawn: L. Robinson | Date: 22-03-2023 | Figure Ref: 2481-22-BIDR-1RevB_230322_Fig12_Flood_Extent

Figure 13: Photograph of flooding on Coondewanna Flats in January 2012 reproduced from AQ2 (2015). Note that inundation on Lake Robinson is not evident at this time.



BHP

Coondewanna Flats Remote Sensing History

Figure 14: Landsat image time series showing a dry-wet sequence at Coondewanna Flats

Scale: 1:85,000 at A3

Coordinate System: GDA 1994 MGA Zone 50

0 1 2 3 4 Kilometres

N

astron delivering environmental intelligence™

Author: R. Archibald | Drawn: L. Robinson | Date: 22-03-2023 | Figure Ref: 2481-22-BIDR-1RevB_230322_Fig14_Image_Panels_Combined

3.5 Predictive Model for Flood Events

RF classification models were developed to predict flood events. The 1-month rainfall variable was of highest importance in all models (Table 6). Time since last flood, 1-month KBDI and 12-month rainfall were of least importance. To assess the accuracy of each model, 10% of the records consisting of both flood and dry records were removed from the training data to be used as testing data. Due to the number of dry records compared to flood records as well as the sampling method used, this generally meant that only one flood event was used for testing. All models were similarly accurate, ranging from 0.97 to 0.98 suggesting that the majority of dry events were correctly classified. The flood event used for testing was correctly predicted in Model 3 where the F1 score is highest. The flood event was not correctly identified in models 4 to 6. While the cross-validation approach to calculation attempts to address the imbalance within the data of few positives (start of flood events) and many negatives (dry periods), high accuracy scores are still very likely when the imbalance is substantial. The F1 score provides an indication of the degree of over-classification (i.e. for flood events, how many dry periods were classified as floods, and for dry periods, how may flood events were classified as dry periods). F1 scores were similar between models ranging from 0.27 to 0.33. The models are therefore prone to overclassifying. Considering accuracy, F1 score and the number of input variables, Model 6 involving simply 1-month rainfall is the most suitable model. The use of 1-month rainfall as a predictor aligns with the finding of AQ2 (2015). However, the fact that model was unable to correctly classify the actual commencement of flooding for the flood within the testing data set also highlights it inadequacies.

Appendix A contains the RF classification predictions of the commencement of flooding events using the entire range of data. The predictions presented were derived from the iteration of the model with the largest F1 score for each table. All models correctly classify all 10 flood records used to train the RF classification model. Model 1, the model run on all predictor variables, misclassified the least number of dry records as flood events despite the average F1 score not being the highest (four records incorrectly classified as floods). Model 2 and Model 6 misclassified six dry records as flood events and Model 3, Model 4 and Model 5 misclassified eight, 10 and 12 dry records as flood events, respectively.

Table 6: Results of Random Forest classification models to predict flood events (i.e. the start of flood events). The number of input variables was progressively reduced from Model 1 to Model 6 by eliminating the variable of least importance. The flood date used as validation data is presented in the Actual Flood Date column and predicted floods are shown in the Predicted Flood Date column. Predicted dates in red indicate that the prediction is actually a dry record.

Model no.	Importance values (0 to 1)						Accuracy: cross-validation (0 to 1)~	F1 Score (0 to 1)~	Actual Flood Date	Predicted Flood Date*
	Time since last flood	1-month rainfall	3-month rainfall	12 -month rainfall	1-month evap.^	1-month KBDI^				
1	0.067	0.444	0.281	0.037	0.123	0.048	0.980	0.300	02/01/2014	02/01/2014 28/02/2020
2	0.050	0.520	0.279		0.118	0.033	0.975	0.300	02/01/2014	02/01/2014 28/02/2020
3	0.044	0.547	0.298		0.110		0.977	0.333	02/01/2014	02/01/2014 28/02/2020
4		0.466	0.308		0.226		0.972	0.283	02/01/2014	28/02/2020
5		0.585	0.415				0.969	0.267	02/01/2014	28/02/2020
6		1.000					0.971	0.307	02/01/2014	28/03/2007

^ mean value for 1-month period

~ mean value calculated from 300 iterations of the model

* obtained from the model iteration returning the highest number of predicted flood events

The summary statistics for weather variables at the initiation of a flood event are presented in Table 7. Values for most variables show considerable dispersion. The mean value of 131 mm for 1-month rainfall is below 180 mm: the estimated threshold for flooding indicated by AQ2 (2015). This difference likely stems from the greater number of small flood events detected using Landsat image analysis. The low minimum 1-month rainfall value associated with a flood event, 42 mm, combined with the low F1 scores for the prediction models, indicates that accurate prediction of flood events is unlikely under the best available model.

The highest annual rainfall totals since 1986-87 have generally coincided with flood events (Figure 6). As with annual rainfall, low values for KBDI frequently coincided with flood events, but not consistently (Figure 9, Figure 10 and Table 7).

Model development is currently limited by the fact that only 10 flood events are available. The first flood event could be not used because the Time since last Flood variable required an initial flood to be used to set this variable to 0. Model development is also limited by gaps in Landsat imagery leading to difficulties in accurately identifying the beginning of a flood event. The identification of the start of a flood event to within an accuracy of a few days is needed if 1-month rainfall or any other variables across shorter time frames (e.g. 5-day rainfall) are to be used within a model.

Table 7: Summary statistics for records of variables at the start of flood events.

Variable	Mean	Standard deviation	Minimum	Maximum	Number of records
Time since last flood (number of days)*	668.0	417.5	220.0	1684.0	10
1-month rainfall (mm)	130.9	58.1	41.5	210.6	11
3-month rainfall (mm)	203.0	68.6	103.4	317.8	11
12-month rainfall (mm)	441.5	130.8	252.2	691.8	11
1-month evaporation (mm)^	9.1	1.8	4.7	10.6	11
1-month KBDI (mm)^*	58.8	38.3	6.1	106.9	11

^ mean value for 1-month period

* Keetch-Byram Drought Index

3.6 Vegetation Trends and Response of Vegetation to Flood Events

Mean MSAVI on the flats cycles higher and lower with wet and dry seasons, respectively (Figure 15 and Figure 16). Peaks in MSAVI occur after significant monthly rainfall totals and after flood events. Most of the highest MSAVI values across the 1987-88 to 2020-21 period occur in the months after a flood event. This response can be seen in the pronounced increase in vegetation cover and vigour across a dry-wet sequence of imagery that includes the 2012 flood (Figure 14). Some notable sharp reductions (short-lived) followed by sharp increases in MSAVI immediately following flood events can also be seen in some instances, particularly during 1998-99 and 2005-06. Surface water has a low MSAVI reflectance value and extensive inundation is likely temporarily depressing MSAVI in these instances.

The range of fluctuations in MSAVI for PEC P1, which occurs on Lake Robinson, are much more pronounced than those for PEC P3, which occurs across the remainder of the flats (Figure 15 and Figure 16). The explanation for this is not clear. However, as surface water mostly flows from the north over Lake Robinson, it may be that the frequency and volume of surface water inflow to Lake Robinson is greater than further south at PEC P3. Other factors relating to soil type or stratigraphy may also account for the differences between vegetation response of the two PECs.

Distinct patterns in the MSAVI times series exist according to species, with these patterns differing between PEC P1 and PEC P3. In comparison to all vegetation or *E. victrix*, MSAVI values for *A. aptanuera* tended to increase more quickly following significant rainfall events and floods (Figure 15 and Figure 16). This aligns with sap flow studies on the flats which show that water uptake in *A. aptanuera* responds to increased moisture availability sooner than *E. victrix* (Astron Environmental Services 2014). While physiological attributes may in part account for this difference, the contrasting distribution of rooting depth between *A. aptanuera* (shallow) and *E. victrix* (deeper) (Astron Environmental Services 2014) is undoubtedly an important factor. It is notable that the increase in MSAVI is more pronounced within PEC P1 than within PEC P3. The cause of this difference is not clear but most likely relates to differences in soil properties and/or infiltration regimes between PEC P1 and PEC P3. The finer textured soils in the Lake Robinson area (PEC P1) compared to PEC P3 (AQ2 2016) most likely lead to greater stores of soil moisture within the root zone at PEC P1 following flooding or significant rainfall.

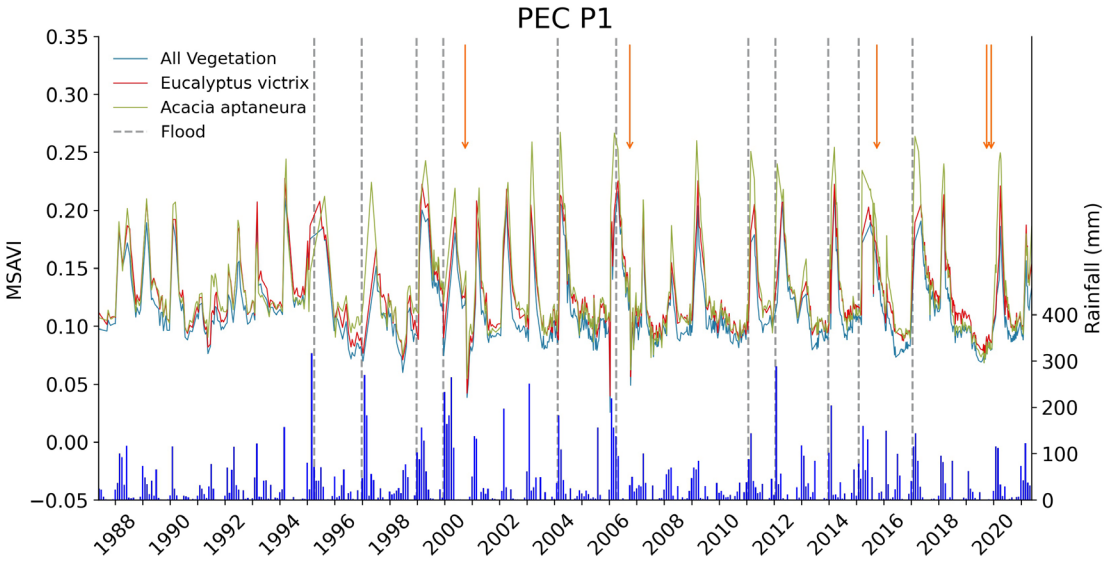


Figure 15: Mean Modified Soil Adjusted Vegetation Index (MSAVI) for all available cloud-free Landsat images from 1986-87 to 2020-21 for the Priority Ecological Community subtype Priority 1: *Coolibah woodlands over lignum (Duma florulenta) over swamp wandiree (Lake Robinson is the only known occurrence)* on Coondewanna Flats. MSAVI values are shown for all vegetation as well as the two main tree species. The start of flood events and monthly rainfall totals are also shown. The timing of fires since 2000 is indicated with orange arrows using data from Northern Australia Fire Information (2022).

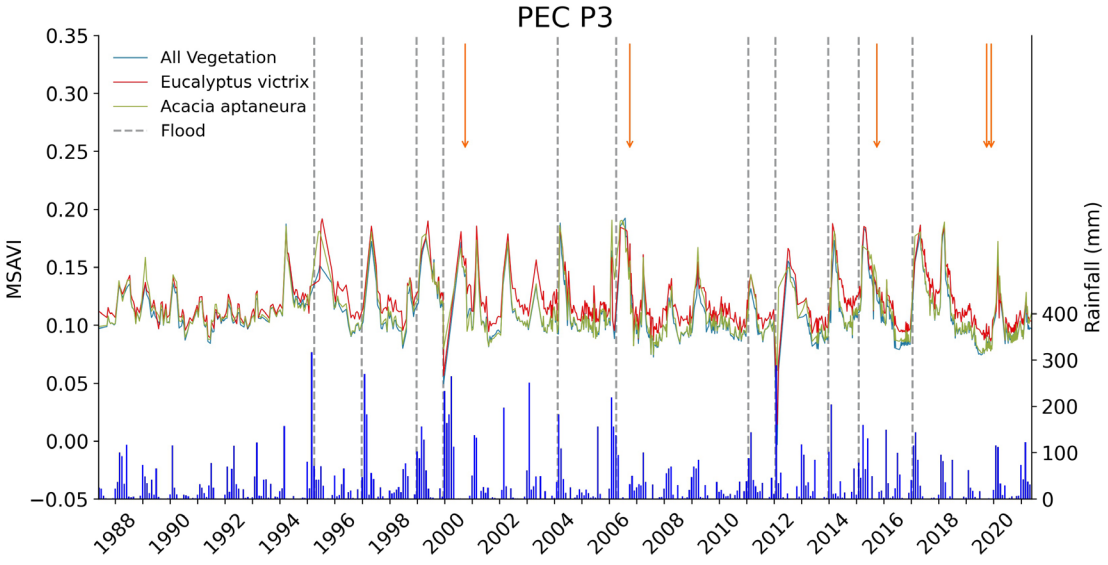


Figure 16: Mean Modified Soil Adjusted Vegetation Index (MSAVI) for all available cloud-free Landsat images from 1986-87 to 2020-21 for the Priority Ecological Community subtype Priority 3: *Coolibah and mulga (Acacia aneura) woodland over lignum and tussock grasses on clay plains (Coondewanna Flats and Wanna Munna Flats)* on Coondewanna Flats. MSAVI values are shown for all vegetation as well as the two main tree species. The start of flood events and monthly rainfall totals are also shown. The timing of fires since 2000 is indicated with orange arrows using data from Northern Australia Fire Information (2022).

When data are summarised to annual values (Figure 17 and Figure 18), some of the points described above are revealed more clearly. Peaks in MSAVI are most coincident with flood events in both PECs and vegetation within PEC P1, situated on Lake Robinson, is clearly more responsive to flood events and rainfall than the vegetation within PEC P3. However, the annual data illustrate a clear difference in *E. victrix* between the PECs. With PEC 1, MSAVI of *E. victrix* increases and declines in a similar way to *A. aptanuera* and all vegetation. However, within PEC 3, MSAVI values of *E. victrix* become distinctly higher and stay higher for longer, following flood events, in comparison to *A. aptanuera* and all vegetation. In fact, it is apparent that floods increase *E. victrix* vigour for more than one year following a flood, with a sequence of floods having an influence that can manifest for several years. This is evident in the relative similarity in MSAVI between *E. victrix*, *A. aptanuera* and all vegetation for the 1986-87 to 1994-95 period (relatively dry period with no floods) and the higher MSAVI values for *E. victrix* after sequences of floods that are sustained across the relatively dry flood-free periods of 2000-01 to 2003-04 and 2006-07 to 2009-10. This pattern accords with the characteristic deep rooting habit of *E. victrix* on the flats and the capacity of many trees of this species to maintain growth by accessing stores of soil water, and potentially occasionally groundwater, at depth (Pfautsch et al. 2011, Astron Environmental Services 2014). However, this pattern of water use was not found in all trees in the study by Astron (2014), with one sample tree of the species on Lake Robinson displaying more of a pulse-response to increases in water availability. This finding supports the differences in *E. victrix* function between PEC P1 and PEC P3 highlighted in the present report.

There were no distinct long-term trends in MSAVI for the vegetation on the flats (Figure 17 and Figure 18). The effect of the large fire in 1999-2000 (Figure 11) appears evident insofar as the post-flood peak in MSAVI subdued relative to other post-flood peaks. The effect of the fire appears relatively short-lived within the MSAVI time series.

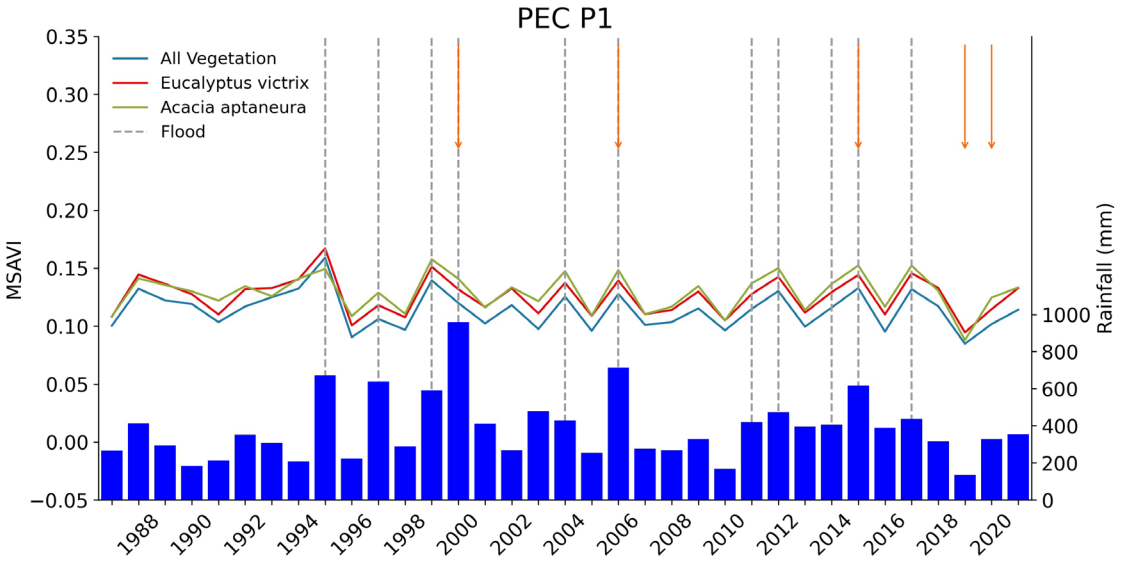


Figure 17: Mean Modified Soil Adjusted Vegetation Index (MSAVI) for each year from all available cloud-free Landsat images from 1986-87 to 2020-21 for the Priority Ecological Community subtype Priority 1: *Coolibah woodlands over lignum (Duma florulenta) over swamp wandiree (Lake Robinson is the only known occurrence)* on Coondewanna Flats. MSAVI values are shown for all vegetation as well as the two main tree species. The start of flood events and annual rainfall totals are also shown. Years are represented from 1 November to 31 October. The timing of fires since 2000 is indicated with orange arrows using data from Northern Australia Fire Information (2022).

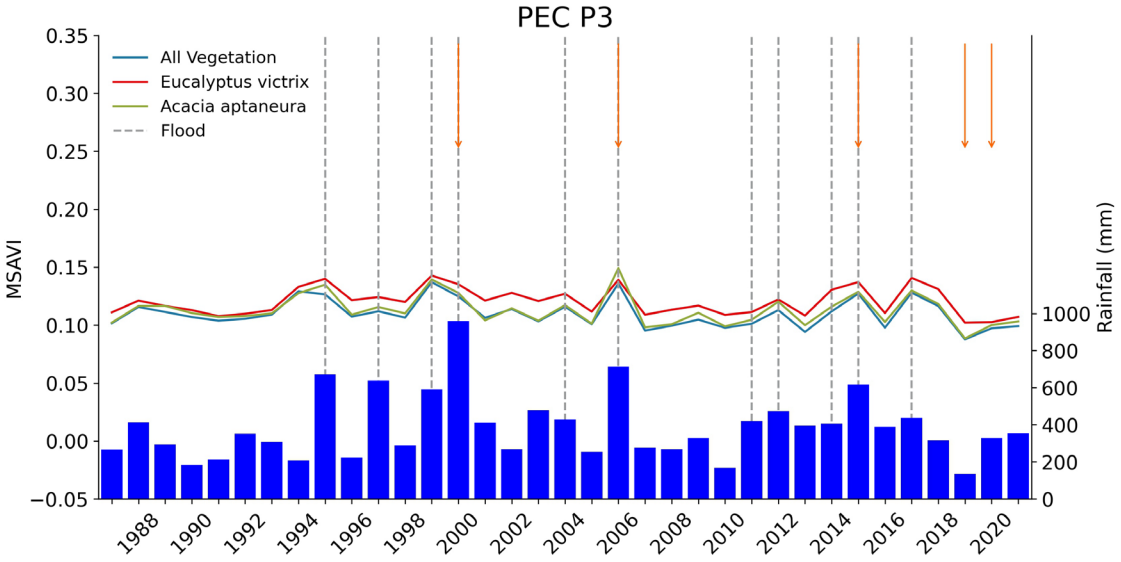


Figure 18: Mean Modified Soil Adjusted Vegetation Index (MSAVI) for each year from all available cloud-free Landsat images from 1986-87 to 2020-21 for the Priority Ecological Community subtype subtype Priority 3: *Coolibah and mulga (Acacia aneura) woodland over lignum and tussock grasses on clay plains (Coondewanna Flats and Wanna Munna Flats)* on Coondewanna Flats. MSAVI values are shown for all vegetation as well as the two main tree species. The start of flood events and annual rainfall totals are also shown. Years are represented from 1 November to 31 October. The timing of fires since 2000 is indicated with orange arrows using data from Northern Australia Fire Information (2022).

4 Conclusions

Based on Landsat image analysis, 11 flood events have occurred on Coondewanna Flats at an average frequency of one in every three years between 1986-87 and 2020-21. This frequency is higher than previous estimates of one in four years. However, one in four years may be a better estimate across a longer time frame that includes the drier decades prior to 1986-87. Three small flood events, which to our knowledge remain unreported, were detected as occurring after the well-known flood event of 2012.

The use of Landsat imagery to detect inundation enabled the distribution of water across the flats to be mapped for each flood event. These data showed that during the maximum extent of large flood events, inundation persists across broader areas of the flats and not on Lake Robinson. For smaller flood events, inundation is more prevalent on Lake Robinson.

Modelling to predict flood events was attempted using variables relating to rainfall, evaporation, drought, and time since previous flood events. The strongest predictor was 1-month rainfall total. This accords with findings presented in AQ2 (2015, 2016). However, these previous reports nominate 180 mm as the threshold 1-month rainfall total that initiates a flood event. Modelling based on floods detected from Landsat imagery suggests a lower value is more appropriate. The average 1-monthly total for the floods detected in this study was 131 mm. However, the accuracy of the 1-monthly total is uncertain at present given difficulties in identifying the start of a flood event with Landsat imagery.

With only 11 flood events since 1986-87 and considerable variation in the timing and characteristics of each flood, it will be difficult to build a model using climatic data which predicts flood events and their characteristics. As such, it is not likely to be possible to determine whether the presence/absence of floods, or the characteristics of floods when they do occur, deviate from natural variation or not. Instrumented monitoring of surface flow within drainage lines that flow onto the flats is likely to be more revealing, particularly if potential impact and control drainages can be instrumented.

While the use of Landsat imagery represents the best available technique for detecting historical flood events in remote areas where on-ground records are deficient, the technique has shortcomings that can lead to the start and end dates of flood events being difficult to estimate accurately or some flood events being completely undetected. These shortcomings arise from gaps in the available cloud-free images in the Landsat archive which are accentuated due to the common co-occurrence of cloudy periods with flood events. Data from an on-ground sensor network of surface flow gauges and cameras in suitable locations on the flats would enable the commencement of floods to be more accurately identified in future years. The use of cloud-penetrating radar data from the Sentinel satellite could also be utilised to detect patterns of flooding, noting that this satellite only became available from 2014. It is recommended that further studies examine the utility of these data for determining the timing and distribution of inundation on the flats.

The time series data for vegetation condition on the flats using the MSAVI index from Landsat data showed a strongly cyclical pattern where condition increases after heavy rainfall and flood events. The highest peaks in MSAVI tend to occur soon after flood events. Contrasts in the patterns of MSAVI response were observed between PEC P1, on Lake Robinson, and PEC P3, present across the remainder of the flats. The response of vegetation on Lake Robinson was more pulse-like, with similar patterns apparent for *E. victrix*, *A. aptanuera* and the vegetation as a whole. However, within PEC P3, the effects of flooding tend to manifest for longer, particularly for *E. victrix* which is capable of drawing on stores of water deep within the soil or from groundwater during extended periods following floods. The results further highlighted the importance of periodic floods in the ecology of this species on the flats.

5 References

Acharya, T., A. Subedi, H. Huang, and D. Lee. 2019. Application of Water Indices in Surface Water Change Detection Using Landsat Imagery in Nepal. Sensors and Materials 31:1429.

AQ2. 2015. Coondewanna Flats Eco-hydrology Review and Conceptual Model. Report prepared for BHP Billiton Iron Ore Pty Ltd.

AQ2. 2016. Coondewanna Flats Ecohydrological Conceptualisation. Report prepared for BHP Billiton Iron Ore Pty Ltd.

Astron Environmental Services. 2011. BHPBIO Coondewanna Flats Eco-hydrogeological Survey - Vegetation Component - Phase 1, November 2022. unpublished report to BHP Billiton Iron Ore.

Astron Environmental Services. 2014. Coondewanna Flats Ecohydrological Study – Ecological Water Requirements of Vegetation Report. Unpublished report prepared for BHP Billiton Iron Ore Pty Ltd.

Astron Environmental Services. 2021. Coondewanna Flats Vegetation Monitoring Program Annual Report 2020-21. Unpublished report prepared for BHP Western Australian Iron Ore.

Astron Environmental Services. 2022. Coondewanna Flats Vegetation Monitoring Program - Report Card - December 2021. Unpublished prepared for BHP Western Australia Iron Ore.

Baig, M. H. A., L. Zhang, T. Shuai, and Q. Tong. 2014. Derivation of a tasselled cap transformation based on Landsat 8 at-satellite reflectance. Remote Sensing Letters 5:423–431.

BHP. 2021. Central Pilbara Water Resource Management Plan.

Breiman, L. 2001. Random forests. Machine Learning 45:5–32.

Bureau of Meteorology. 2023. Tropical cyclone reports.

Cordeiro, M., J. M. Martinez, and S. Peña-Luque. 2021. Automatic water detection from multidimensional hierarchical clustering for Sentinel-2 images and a comparison with Level 2A processors. Remote Sensing of Environment 253.

Department of Fire and Emergency Services. 2014. What is the Keetch-Byram Drought Index?

Dzinotizei, Z., A. Murwira, F. M. Zengeya, and L. Guerrini. 2018. Mapping waterholes and testing for aridity using a remote sensing water index in a southern African semi-arid wildlife area. Geocarto International 33:1268–1280.

Fisher, A., N. Flood, and T. Danaher. 2016. Comparing Landsat water index methods for automated water classification in eastern Australia. Remote Sensing of Environment 175:167–182.

Herndon, K., R. E. Muench, E. A. Cherrington, and R. E. Griffin. 2020. An Assessment of Surface Water Detection Methods for Water Resource Management in the Nigerien Sahel. Sensors (Basel, Switzerland) 20.

Keetch, J. J., and G. M. Byram. 1968. A drought index for forest fire control. US Department of Agriculture, Forest Service, Southeastern Forest Experiment Station 38:1–32.

McFeeters, S. K. 1996. The Use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features. International Journal of Remote Sensing 17:1425–1432.

NAFI. 2022. Northern Australian Fire Information. https://www.firenorth.org.au/nafi3/.

Ogilvie, A., G. Belaud, S. Massuel, M. Mulligan, P. Le Goulven, and R. Calvez. 2018. Surface water monitoring in small water bodies: potential and limits of multi-sensor Landsat time series. Hydrology and Earth System Sciences 22:4349–4380.

Pfautsch, S., C. Keitel, T. L. Turnbull, M. J. Braimbridge, T. E. Wright, R. R. Simpson, J. A. O'Brien, and M. A. Adams. 2011. Diurnal patterns in water use of Eucalyptus victrix indicate pronounced desiccation-rehydration cycles despite unlimited water supply. Tree Physiology 31:1041–1051.

Pringle, M., M. Schmidt, and J. Muir. 2009. Geostatistical interpolation of SLC-off Landsat ETM+ images. ISPRS Journal of Photogrammetry and Remote Sensing:654–664.

Qi, J., A. Chehbouni, A. Huete, Y. Kerr, and S. Sorooshian. 1994. A Modified Soil Adjusted Vegetation Index. Remote Sensing of Environment 42:119–126.

Queensland Government. 2022. SILO climate database: https://www.longpaddock.qld.gov.au/silo/.

R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rouse et al. 1973. Monitoring vegetation systems in the great plains with ERTS. Third ERTS Symposium, NASA 1:309–3017.

RPS Aquaterra. 2011. Coondewanna Village Development: Surface Water Assessment.

Sarp, G., and M. Ozcelik. 2017. Water body extraction and change detection using time series: A case study of Lake Burdur, Turkey. Journal of Taibah University for Science 11:381–391.

Subramaniam, S., and M. Saxena. 2011. Automated Algorithm for Extraction of Wetlands from IRS Resourcesat LISSIII Data. Page ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.

Williamson, G. 2016. ecbtools: Environmental Change Biology Tools.

Xu, H. 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. International Journal of Remote Sensing:3025–3033.

Zhai, K., X. Wu, Y. Qin, and P. Du. 2015. Comparison of surface water extraction performances of different classic water indices using OLI and TM imageries in different situation. Geo-spatial Information Science 18:32–42.

This page has been left blank intentionally.

Appendix A: Random Forest Model Outputs

This page has been left blank intentionally.

Table A.1: Model 1 Random Forest classification of flooding events run on all available data: records where a flood first occurs (flood start) and all dry records. Predictions are presented in the model iteration where the F1 score was highest.

Date	Actual Classification
11/01/1997	Flood start
17/01/1999	Flood start
19/12/1999	Flood start
18/02/2002	Dry
21/02/2003	Dry
24/02/2004	Flood start
12/01/2006	Flood start
03/02/2011	Flood start
14/02/2012	Flood start
02/01/2014	Flood start
10/03/2015	Flood start
17/02/2016	Dry
25/02/2016	Dry
19/02/2017	Flood start

Table A.2: Model 2 Random Forest classification of flooding events run on all available data: records where a flood first occurs (flood start) and all dry records. Predictions are presented in the model iteration where the F1 score was highest.

Date	Actual Classification
11/01/1997	Flood start
17/01/1999	Flood start
19/12/1999	Flood start
18/02/2002	Dry
24/02/2004	Flood start
12/01/2006	Flood start
03/02/2011	Flood start
14/02/2012	Flood start
31/01/2013	Dry
02/01/2014	Flood start
10/03/2015	Flood start
17/02/2016	Dry
25/02/2016	Dry
19/02/2017	Flood start
12/02/2020	Dry
28/02/2020	Dry

Table A.3: Model 3 Random Forest classification of flooding events run on all available data: records where a flood first occurs (flood start) and all dry records. Predictions are presented in the model iteration where the F1 score was highest.

Date	Actual Classification
11/01/1997	Flood start
17/01/1999	Flood start
19/12/1999	Flood start
18/02/2002	Dry
24/02/2004	Flood start
12/01/2006	Flood start
11/02/2008	Dry
03/02/2011	Flood start
14/02/2012	Flood start
31/01/2013	Dry
02/01/2014	Flood start
10/03/2015	Flood start
17/02/2016	Dry
25/02/2016	Dry
19/02/2017	Flood start
27/01/2020	Dry
04/02/2020	Dry
28/02/2020	Dry

Table A.4: Model 4 Random Forest classification of flooding events run on all available data: records where a flood first occurs (flood start) and all dry records. Predictions are presented in the model iteration where the F1 score was highest.

Date	Actual Classification
11/01/1997	Flood start
17/01/1999	Flood start
19/12/1999	Flood start
18/02/2002	Dry
24/02/2004	Flood start
12/01/2006	Flood start
11/02/2008	Dry
03/02/2011	Flood start
14/02/2012	Flood start
31/01/2013	Dry
02/01/2014	Flood start
10/03/2015	Flood start
01/02/2016	Dry
17/02/2016	Dry
25/02/2016	Dry
19/02/2017	Flood start
27/01/2020	Dry
04/02/2020	Dry
12/02/2020	Dry
28/02/2020	Dry

Table A.5: Model 5 Random Forest classification of flooding events run on all available data: records where a flood first occurs (flood start) and all dry records. Predictions are presented in the model iteration where the F1 score was highest.

Date	Actual Classification
11/01/1997	Flood start
17/01/1999	Flood start
19/12/1999	Flood start
24/02/2004	Flood start
12/01/2006	Flood start
05/04/2007	Dry
11/02/2008	Dry
03/02/2011	Flood start
14/02/2012	Flood start
31/01/2013	Dry
02/01/2014	Flood start
06/02/2015	Dry
10/03/2015	Flood
14/06/2015	Dry
25/02/2016	Dry
19/02/2017	Flood start
07/03/2017	Dry
06/02/2018	Dry
10/03/2018	Dry
27/01/2020	Dry
04/02/2020	Dry
28/02/2020	Dry

Table A.6: Model 6 Random Forest classification of flooding events run on all available data: records where a flood first occurs (flood start) and all dry records. Predictions are presented in the model iteration where the F1 score was highest.

Date	Actual Classification
11/01/1997	Flood start
17/01/1999	Flood start
19/12/1999	Flood start
24/02/2004	Flood start
12/01/2006	Flood start
05/04/2007	Dry
11/02/2008	Dry
03/02/2011	Flood start
14/02/2012	Flood start
31/01/2013	Dry
02/01/2014	Flood start
19/02/2014	Dry
06/02/2015	Dry
10/03/2015	Flood start
14/06/2015	Dry
19/02/2017	Flood start