Understanding debris sources and transport from the coastal margin to the ocean

Final Report to the Australian Packaging Covenant Organisation Ltd.

Britta Denise Hardesty, Qamar Schuyler, TJ Lawson, Kimberley Opie and Chris Wilcox

30 August 2016

Australian Packaging Covenant
Funded Project

CSIRO Oceans and Atmosphere

Citation

Hardesty, et al. (2016) Understanding debris sources and transport from the coastal margin to the ocean. CSIRO: EP165651

Copyright

Important disclaimer

Acknowledgments

This work has been supported by the Australian Packaging Covenant and CSIRO Oceans and Atmosphere. We acknowledge the critical participation of, and stimulating conversations with stakeholders who engaged in the Australian Litter Workshop in November 2015. We also thank the many volunteers, citizen scientists, state, federal and government employees and interested members of the public who are committed to reduce litter, marine debris, and illegal dumping in the environment. We thank Kathy Willis and Jack Buckingham for carrying out additional surveys in the Brisbane region. Finally, we thank the New South Wales Environmental Protection Agency for contributing financially and to the organization of the litter workshop and we thank Clean Up Australia, Keep Australia Beautiful and Keep South Australia Beautiful for providing data that has enabled the analyses presented in this report.

Table of Contents

Tables

Figures

1 Executive Summary

Marine debris is defined as any persistent solid material that is manufactured or processed and, directly or indirectly, intentionally or unintentionally, disposed of or abandoned into the marine environment. While this project is focused on what we call 'marine debris', it might be better simply called 'debris' to encompass all anthropogenic litter. The data collected and the results presented here focus on land-based surveys and, accordingly, the land-based sources of the debris or litter recorded.

The CSIRO and Australia Packaging Covenant project aimed to understand where littering and other losses of rubbish into the environment are occurring. The project specifically focused on the following main questions: 1) what is the relationship between debris in the marine environment and debris from nearby sites; and 2) are there identifiable pathways through which debris reaches and moves into the marine environment.

The project achieved its aim of improving our understanding of litter across Australia. We identified hotspots of coastal debris, which enables us to cost-effectively target regions for waste reduction. We evaluated whether quantities of debris varied with different site types (beaches, highways, recreational parks and so on). In this analysis, we found smaller quantities of debris at beaches and in residential areas than along highways and in industrial areas, retail strips and shopping centres. We identified a number of variables that affect debris loads, including socioeconomic factors, population density and accessibility. Using this information, we have developed a model that allows the prediction of the debris load at unsurveyed sites. We also mapped litter flows to the ocean using meteorological datasets and other variables (wind force, water flows, human transport), using the watersheds in the greater Brisbane area (approximately 110,000 square kilometres) to test whether such features are significantly correlated with the transport of debris. We found that human deposition was by far the most important in determining the load at a site; transport by water was second, with a discernible but smaller effect from wind transport.

As a result of the analyses, we are now poised to present opportunities or suggestions to reduce debris inputs to the environment. We suggest that addressing sites with high littering rates should be the first priority. The strongest effect on loads is due to direct deposition (that is, littering behaviour), which contributed more than transport in determining the load (or debris count) at a site. Second, it is critical to target particular population segments. The three strongest predictors of debris at a site had to do economic wealth and social disadvantage in the population near the site. Similarly, the political context (e.g. state or territory) matters—the Northern Territory had particularly high loads, both in comparison to other jurisdictions and in the light of other possible explanations for load. Fourth, land use and people's activities in the land-use context tended to correlate with high loads. Thus creeks and natural areas had high debris loads, as did sites that had parking, residential or shopping-related activities. Since the data available for analyses did not sample all of those contexts equally, results need to be interpreted with some caution, but they do give a basis for inferring contexts where loads are particularly high and interventions might make the maximum difference.

While deposition is a key driver of the amount of debris at a site, transport by wind and water are also important. Areas downwind and downstream act as sinks, accumulating debris from other sites. Water transport was more important than wind, suggesting that debris traps in both surface and stormwater systems are a useful intervention. Moreover, using the maps of loads and flows developed in this project, it is possible to anticipate the load that would be expected given the location of the trap. This information can support the allocation of funds and staff time in installing and maintaining infrastructure to maximum efficiency. Critically, loads appear to be high in watercourses, likely due in part to direct deposition; and given the propensity for transport to the marine environment, reducing loads in watercourses is a vital action.

There are a number of opportunities for improvement in both data and analysis concerning littering, transport and loss to the marine environment. We suggest three major possibilities.

First, the highest quality information is currently concentrated along the coast and in the major cities. Increasing the spatial coverage and ensuring that important site types are sampled on urban margins and in rural areas will significantly improve the information available. This is particularly important, as many critical variables are under-sampled, such as areas of socioeconomic disadvantage, residential areas, and natural areas that are estimated to have high loads. It may be possible to extend the data collection to address this issue using a single designed survey, avoiding the need for the long-term expansion of any of the existing data collection programs.

Second, the transport modelling completed for this project is a useful first step in understanding the movement of debris and identifying critical intervention points. Increasing the detail in this analysis would improve the capacity to identify and evaluate intervention points. For instance, it might be possible to estimate whether funds would be better spent installing and maintaining a litter trap in a storm drain, or clearing the existing debris load in a creek receiving the storm water. It is likely that limited funds and staff time force decisions on local governments responsible for waste management, so there is a significant opportunity in providing information on the likely success or impact of such allocation decisions.

Third, the load analysis in this project could be used in conjunction with interventions such as litter signage, bin installations or community outreach programs to evaluate the impact of those programs on load. The current statistical model can be used to predict the expected load in these contexts. Thus, sites that have had or will have interventions can be compared to assess the effect of the intervention on the debris load. This analysis can be done on an ad hoc basis, using load sampling in the area at some point after a known intervention.

We encountered a number of challenges in this project, some of which are now resolved and others which are likely to remain issues into the future. First, existing large-scale debris datasets have been collected using different methodologies and provide data of varying quality. Combining those datasets proved challenging, but we were able to integrate at least two of them in some analyses. Critical in addressing this are robust sampling designs, which quantify both sampling effort and observer error. We were able to make use of three of the four major datasets we identified during the project scoping phase; however, the information we could obtain from the data was to some extent limited by the sampling design. Particularly for volunteer-collected data, addressing some of these issues would significantly increase the utility of the information.

Second, obtaining data from existing data holders was a significant task. In some cases, the data required extensive preparation by the project team to bring it to a useable standard. For instance, several datasets did not include geographic coordinates, making it impossible to find information such as land use or the socioeconomic context in order to analyse the data. We were able to address this issue; however, future work with these datasets will be much more efficient if data holders incorporate a small amount of additional information in future monitoring. A larger problem, which remains partially unresolved, is the willingness of data holders to share their data. We experienced the full range of responses from data holders. Some were openly willing to share data with no restrictions. Others required confidentiality agreements, which took time to negotiate. At the extreme, one of the data holders we identified in the scoping phase declined to share data in any context, despite offers of confidentiality or terms of use agreements. This aspect of openness about data, which ultimately is turned into useful information only through evaluation and analysis, is a critical consideration in future projects, be they extensions of this work or otherwise. The APC, in its role as a funder and coordinating body, and through its relationship with the Department of the Environment, may be able to improve data sharing at the national level, which would increase the value of investments for all stakeholders.

2 Introduction

2.1 Rationale

Anthropogenic debris has myriad impacts on the environment, on people, on biodiversity, and on industry. We need to significantly reduce debris and increase recycling at local, regional and national levels. To do this requires robust analyses to identify key loss points as well as feasible and efficient responses to reduce loss rates into the environment. Critically, we need solid data and information to ensure evidence-based decisions underpin actions. This will make certain that on-ground activities are successful in meeting targets for debris reduction.

This project set out to estimate the load of debris at the landscape scale, identify the important variables driving debris hotspots, and analyze the transport mechanisms and pathways debris takes on its journey into the marine environment.

We compiled and analyzed existing land-borne and marine debris data from a variety of sources. Our previous research had suggested that isolated areas may be important sources of plastic pollution through illegal dumping. Hence, in addition to focusing on major metropolitan areas, considering remote and regional sites is key to understanding loss rates and flows. This can help to target infrastructure and outreach campaigns, and improve success of incentives and enforcement actions to reduce littering and improve packaging materials recovery.

By understanding where littering takes place as well as issues and current trends in littering, we can know how and where best to put resources so that we improve recovery, particularly away from individual consumer households. The work we propose directly addresses Goals 2 (Recycling) and 3 (Product Stewardship) of the Australian Packaging Covenant's Key Performance Indicators (as per the website; www.packagingcovenant.org.au/data/Resources/Performance_Goals_and_KPIs_29.08.13.pdf).

CSIRO's analyses into policy and practices to reduce littering and identify source point losses into the environment relied upon data from multiple partners. The analyses undertaken in the project provide a benchmark for the investigation of the amounts, types, sources of litter and policies and practices that may be associated with litter and its reduction.

The team's approach to understanding where littering and other losses of debris into the environment are occurring focused on two main questions: 1) what are the important factors driving debris hotspots on land and in coastal environments; and 2) are there identifiable transport mechanisms and pathways through which debris moves into the marine environment.

2.2 Aim

The aim of this project was to better understand the pathways of land based litter into the marine environment. Future work will focus more extensively on the influence of the policy context, infrastructure and other action programs. The ultimate goal is to use this work to inform future projects and policy in litter mitigation around Australia.

2.3 Project Description

This project sets out to investigate the relationship between marine environments and litter from land-based sources. Using reference to a range of existing datasets, we used a spatial statistical model across the Australian coastline to evaluate likely routes for debris to move in the marine environment, which looked at linkages between systems such as rivers, watersheds, wind and roads. We also look at the type of site (industrial, retail etc.) development and proximity to the marine environment.

We focused on two main questions: 1) what is the relationship between debris in the marine environment and debris from nearby sites; and 2) are there identifiable pathways through which debris reaches and moves into the marine environment.

To answer these questions we undertook 3 bodies of analyses:

1. We analysed the available data on debris loads at coastal and inland sites to understand the factors leading to high and low debris loads at survey sites;
2. We used these models to make predictions of the load at the landscape scale across a major urban region in Australia. We then evaluated three different possible transport mechanisms for that load to move into marine systems.
3. We identified hotspots of high debris loads, both directly from the data and through development of a spatial statistical model of loads to capture trends outside surveyed areas.

3 Description of the works undertaken

This was a very ambitious project that aimed to achieve results reliant upon a number of outside factors, including the acquisition of data from multiple data custodians. As a result of the additional time required to procure data used in the analyses herein, subsequent steps of the project were somewhat delayed. Furthermore, quality assurance and data curation took significant time beyond that which was anticipated. Overall, however, the project achieved its substantive aims, and we gained significant insight into the debris issue in Australia. Working closely and having good communication with the staff at the Australian Packaging Covenant (APC) meant that the key participants were aware of changes in the project plan along the way.

The fulfilment of the project required a number of steps to be undertaken. These steps included:

1) Data compilation (see further discussion in Section 4)
    a. Identification of data holders/data partners
    b. Outreach to data holders
    c. Data acquisition from data holders
        i. Developing and procuring data sharing agreements between data holder(s) and CSIRO
        ii. Quality assurance of data provided
        iii. Curation of data, including geo-referencing in many cases
        iv. Input of data provided by data sharers into electronic format
    d. Compilation of data into a consistent format for analysis
2) Stakeholder engagement (see further discussion in Section 3.1)
    a. Identification of stakeholders
    b. Survey of stakeholders responses
    c. Stakeholder workshop
3) Data analyses (see further discussion in Section 5)
    a. Comparison of data collection methods
    b. Compilation of data for site characterisation including roadways, land cover, railways, census data, socio-economic index, watershed information, wind data
    c. Statistical exploration of data
    d. Statistical analyses of data
    e. Visualization of outcomes
4) Project reporting

a. Risk-management and risk register report (provided October 2015, with update provided November 2015)
b. Stakeholder engagement plan and project methodology (provided October 2015)
c. Stakeholder engagement summary report (provided November 2015)
d. Progress report (provided November 2015)
e. Final report (provided August 2016)

3.1 Stakeholder outreach and engagement

From the myriad coastal debris surveys and coastal clean ups that have taken place around Australia, there is increasing data from many places around the country. This data has been kept and collated from a number of local, regional, state and nationally associated clean up activities and/or surveys such as those run by Clean up Australia, Keep Australia Beautiful, Tangaroa Blue and others. While this information is of variable quality and collects different types of information, it has tremendous potential to help inform the amounts, types, and locations of debris losses into the environment, as previously outlined in the project proposal. With the data included in this work, we aim to quantify factors associated with higher and lower loss rates into the environment, with a goal of reducing debris inputs (and impacts) in the environment.

To best address the knowledge gaps identified in the proposal to the APC, we set out to connect and communicate with stakeholders from around the country. This outreach included:

a) Informal conversations with members of the public regarding debris and loss rates in their communities;
b) Direct conversations with those involved in littering awareness campaigns, clean up activities, waste management and educational/outreach as part of their jobs and/or community volunteer activities;
c) Discussions with the key staff members at the federal department of the environment regarding the relationship between coastal debris and impacts on marine wildlife;
d) Providing input to the federal government (at their request) to inform decisions regarding the revision of the marine debris Threat Abatement Plan (TAP);
e) Participating in the department organized TAP working group meeting regarding revision of the TAP;
f) Participating in monthly meetings with multiple state debris teams to identify and coordinate debris efforts and to better understand key issues, questions, and activities underway in various jurisdictions (particularly New South Wales, Queensland and Victoria);
g) Organization of the 2015 National Litter Workshop which included invited participants from all states and territories, includes special interest groups, volunteers, industry partners, federal government personnel, and major data holders;
h) Development of a web-based survey tool targeted to stakeholder groups from around the country.

The process by which stakeholder engagement has been undertaken included identification of key stakeholders with whom to engage. This involved reaching out to broader networks and web-based searching to identify any debris related programs or activities. Colleagues and collaborated involved in CSIRO's debris related projects over the past 7+ years were also contacted.

The two main stakeholder activities included:

1) A web-based survey which targets a broader audience engaged in litter or debris management, reduction and removal (in final stages of development and trial); and
2) A national workshop called the 2015 Australia Litter Workshop with approximately 30 stakeholders and participants, held in November 2015.

The Australian 2015 litter workshop was comprised of representatives from the state agencies charged with managing litter, debris and illegal dumping (as well as participants from non-governmental organizations, industry and the federal government). The stakeholder meeting took place from Wednesday 25-27 November 2015. Participants included a range of stakeholders a summary from the workshop has been provided under separate cover.

* Discussion of the problem in each of the states and territories

* Unique aspects

* Existing programs

* Lessons regarding success of programs

* Outstanding issues with respect to understanding or addressing the problem

* Development of a shared strategy and set of resources for addressing the problem

* An accessible catalogue of existing policies and outcomes

* Joint projects

The workshop focused on topic areas which were identified as high priority for by participants. These included a summary of existing legislation for various states and territories; investigating stakeholder interest, feasibility and opportunities for a collaborative approach to debris management; information on hand and required to increases understanding and quantify debris and illegal dumping problems across the country, including common issues, unique aspects, existing programs and lessons learned from successful programs (as well as unsuccessful programs to identify pitfalls to avoid); potential reduction strategies, gaps in programs and data, data sharing and coordination of data collection efforts, specific and generic solutions, targets for change, and methods for achieving behaviour change; and funding opportunities and impediments to addressing littering, debris and illegal dumping holistically and at large (and small) spatiotemporal scales.

The workshop provided an open, inclusive and structured forum for exploring key identified issues. It facilitated developing knowledge around the structure of the problem, and identification and sharing of effective solutions. In order to make the workshop as effective as possible the number of key stakeholders attending was limited to just under thirty participants.

3.2 Data acquisition

The parties that have been approached to contribute data to the project included:

1) CSIRO for their national marine debris dataset
2) Keep Australia Beautiful (KAB) for their national survey dataset, excluding South Australia
3) Keep South Australia Beautiful (KSAB) for their South Australian survey dataset
4) Clean Up Australia (CUA) for their national clean up data
5) Tangaroa Blue for their national coastal clean up data, which is derived from multiple groups who add their data to the Australian Marine Debris Initiative (AMDI) database
6) Other state/territory/volunteer-collected data

Each party was approached prior to the commencement of the project, between late 2014 and early 2015, during the scoping phase for the project to assess the data they held and their interest in participating in the project. All of the data holders expressed interest in the project, and the potential for sharing their data in some form. After the commencement of the project in July 2015, the project leaders made contact with the aforementioned potential data providers with the intent of negotiating secure, uncompromised, and mutually acceptable access to their data. Given the project's goal of understanding the status of debris nationally, it was hoped that data holders would subscribe to our objectives and share their data for the joint purpose. Further details and information about the data provided to CSIRO is discussed in Section 5.1.

Ultimately, when the data holders were approached they expressed the full range of responses in terms of willingness to share data. CSIRO, Keep Australia Beautiful South Australia, and Clean Up Australia were openly willing to share data with no restrictions. Keep Australia Beautiful (national) required confidentiality agreements, which took some time to negotiate and Tangaroa Blue was unresponsive with respect to data sharing during the project lifetime, despite offers of cost support, confidentiality, or terms of use agreements.

3.3 Issues and challenges

There are multiple stakeholder groups operating in the litter, waste management, illegal dumping and marine debris space. This provides some excellent opportunities for engagement and collaboration, and it also highlights some potential challenges with regards to stakeholder sensitivities, overlap and complementarity of projects and overall inclusiveness on the topic.

The project tracked well against the deliverables, in spite of a few setbacks that arose during the course of the project. In brief, there were significant changes in staffing at APC, data acquisition did not transpire as had been anticipated and hoped and the stakeholder meeting took place later than planned. Furthermore, data required significant quality assurance and was much more analytically cumbersome than had been anticipated. Due to scheduling and other constraints, the 2015 National Litter Workshop was held in November rather than in August, as had previously been planned. The advantage of this, however, is that the workshop was able to build on a previously held marine debris workshop focused on updating the Threat Abatement Plan, under the EPBC Act organized by the Department of the Environment earlier in the year, in which CSIRO project staff participated.

There were multiple unforeseen risks that arose with the project, all of which were around data sharing and data provision and the time required to procure data. Given the tight timeframe for the proposed work, 14 months from project inception to final reporting, securing data for the project in a timely fashion was of highest import. Data was secured within one to two months from some groups, while negotiating confidentiality agreements for some groups required significantly more time.

In a more extreme case, Tangaroa declined to share the data their volunteers and stakeholders had provided to the AMDI. Despite CSIRO staff contacting Tangaroa Blue Foundation founder, and the head of the Tangaroa Blue Foundation board of directors, agreeing to sign a non-disclosure agreement and offering to provide financial compensation for the database person’s time that would be required to provide the data in a secure manner, permission was not granted. As of September 23, 2016, we have been unable to obtain access to the data Tangaroa Blue holds despite trying to negotiate access for more 18 months.

Tangaroa’s networks are vast, and we had hoped to include the data collected by the multitude of partners who contributed to the AMDI database. This was an excellent opportunity to realize the substantial contribution of volunteer collected data and for this data to be included in a national analysis. If included, it would have increased the sample sizes (e.g. sample sites) and geographic extent of surveys included in our analyses. Furthermore, this was a unique and unparalleled opportunity to provide the numerous volunteers and data contributors from around the country statistically robust feedback on the data they had collected. Unfortunately, this opportunity remains unrealized.

The third unforeseen risk or project issue that has arisen was with respect to data provided by KAB. In verbal conversations and written correspondence, KAB staff stated that they had explicit spatial data (GPS locations) for survey sites. Upon receipt of the data, we were advised that such data were not held by the organization. To overcome the issue so the data can be included in analyses, CSIRO staff required significant time to identify and input the spatial locations for the data provided. Subsequently, CSIRO have provided the data back to KAB. As with the other risks mentioned above, this was an issue that arose, was identified, and resolved. To date, there have been no additional costs to the APC, although these issues have required a substantial additional time and staffing resource investment by CSIRO.

To increase the value of the report, CSIRO staff engaged in additional fieldwork, collecting data from 31 additional coastal sites in the greater Brisbane area. While this was not a task identified early in the project, the project leaders felt it could significantly improve the quality of the final outcomes from the research and would be of overall benefit to the project. Hence, this additional work was carried out and data from these survey sites were included in the overall project analyses.

4 Data used in Analyses

4.1 Survey Data

Site debris survey data has been collected by a number of different organisations, including CSIRO. For this study, five sources of consistent site surveys were assessed (Table 1). These data are quite distinct in nature. Some surveys observe debris, some remove debris and some are a combination of both over time. The sourced survey data covers different periods of time, uses different categories to audit the mix of debris types and practice different survey techniques. This section of the report will cover details on the survey characteristics.

Table 1. Survey grouping for analysis. Note: Keep Australia Beautiful counts are for combined national and South Australian sites.

Data source	Description
CSIRO staff transect	National 100km transect method coastal surveys
CSIRO Fixed area surveys (CSIRO Schools Program)	Schools program using CSIRO fixed area coastal surveys
CSIRO Public transect	Public surveys using CSIRO transect method coastal surveys
Clean Up Australia (CUA)	National annual public clean ups
Keep Australia Beautiful (KAB) including KESAB	National representative debris counts and public clean ups

Data was constrained to include sites that were surveyed from 2007 onwards. The decision to limit the time period of records was made to enable the most accurate covariate data collection (Section 4.2). The entire survey data consists of 18,394 records across all five data sources (Figure 1). This includes 3503 site locations with some sites being surveyed at multiple times. Table 2 details the number of dates over which data was collected for each type of dataset for the time period of this study. KAB surveys make up the majority of data, with an average of 16 surveys at different times per site.

Table 2. Distribution of collected data by date

Description	Number of sites	Number of dates	First date (m/d/y)	Last date (m/d/y)
CSIRO staff transect	202 (668 transects)	90	6/9/2011	18/6/2016
CSIRO Schools fixed area survey	39	22	17/2/2012	10/8/2015
CSIRO Public Transect	150	46	28/3/2012	10/8/2015
Clean Up Australia	1663	9	4/3/2007	1/3/2015
Keep Australia Beautiful, including KESAB	983	100	12/11/2007	25/5/2015

Surveys were each given a unique identifier which was carried throughout the analysis and used to identify unique combinations of site location and the date of the survey. The data was projected into World Mercator and all distances and areas were calculated in kilometres unless otherwise noted. Various covariates were collected to answer questions about what important factors correlate with or influence the volume and mix of debris identified in the surveys (see Section 4.2).

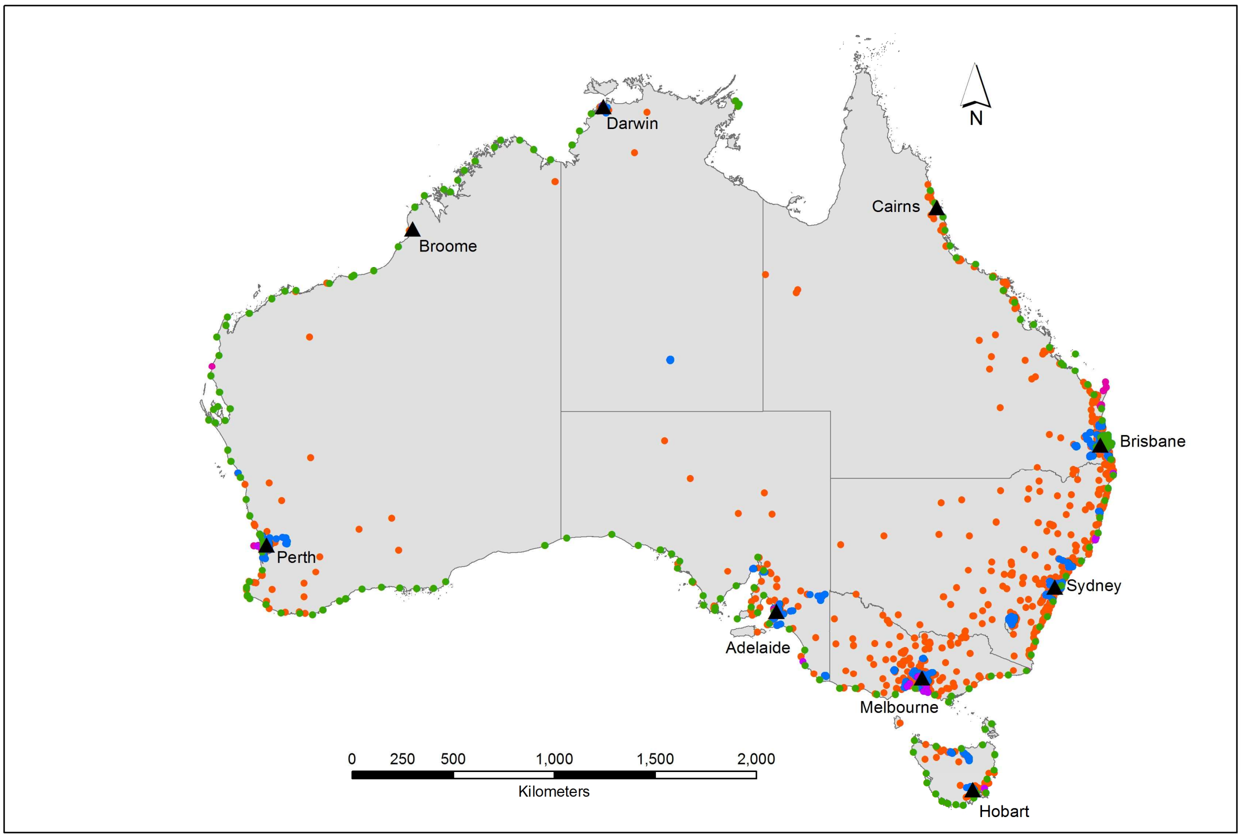

Figure 1. Map of CSIRO (green circles), KAB (blue circles) and CUA (orange circles) survey sites between 2007 and 2016.

4.1.1 CSIRO surveys

CSIRO staff have conducted marine debris surveys at over 202 sites around the coast of Australia, including Tasmania (Figure 2). This was part of a national marine debris project (see Hardesty et al. 2014) set out to look at gaps identified in the Commonwealth Department of Environment Threat Abatement Plan (TAP). This project aimed to answer 4 questions:

1) What are the sources, distribution and fate of marine debris?
2) What is the exposure of marine wildlife to debris?
3) When wildlife are exposed to debris, what factors determine whether animals ingest or are entangled by debris?

4) What is the effect of ingestion or entanglement on marine wildlife populations?

The coastal surveys were conducted to address the first question above. At each coastal site, staff carried out 3 to 6 transect surveys, resulting in 668 records used in analyses. The coastal surveys started at Cape Tribulation in far north Queensland and were conducted clockwise around the coast approximately every 100km to Darwin. In instances where staff did not have access at 100km from the last site, staff used the nearest access point to the coast. Conducting the surveys in this manner resulted in both heavily used and rarely used coastal areas being sampled. Site information was collected at the point we accessed the beach. This included information related to weather, time of day, number of people on the beach, and the location of the access point.

To avoid bias such as higher traffic (and increased debris) at access points, we conducted the first transect survey at least 50m from the access point. Each following transect was a minimum of 50 meters from the last. Where there was different substrate types (sand, rock slab, boulders etc.), transects were distributed evenly among them. Transects ran inland (perpendicular) from the water's edge to two metres into the backshore vegetation or until a cliff/sea wall was encountered. Each transect was 2 meters wide. Two observers, one each side of a measure tape looking at a one meter swath, walked the length of the transect recording any piece of debris found. Debris was recorded into categories based on material type and colour. Size class was also recorded at equal intervals along transects. For a full description of the methods, material types, colours and size categories used for the beach/transect survey please see Hardesty et al. (2014) and Hardesty et al. (*in press*).

As part of this project, CSIRO engaged with over 7,000 students and teachers in an education program. This activity involved supervised beach surveys. Refer to sections 4.1.2 (Fixed area surveys) and 4.1.3 (CSIRO public transects) for further detail on survey methodology.

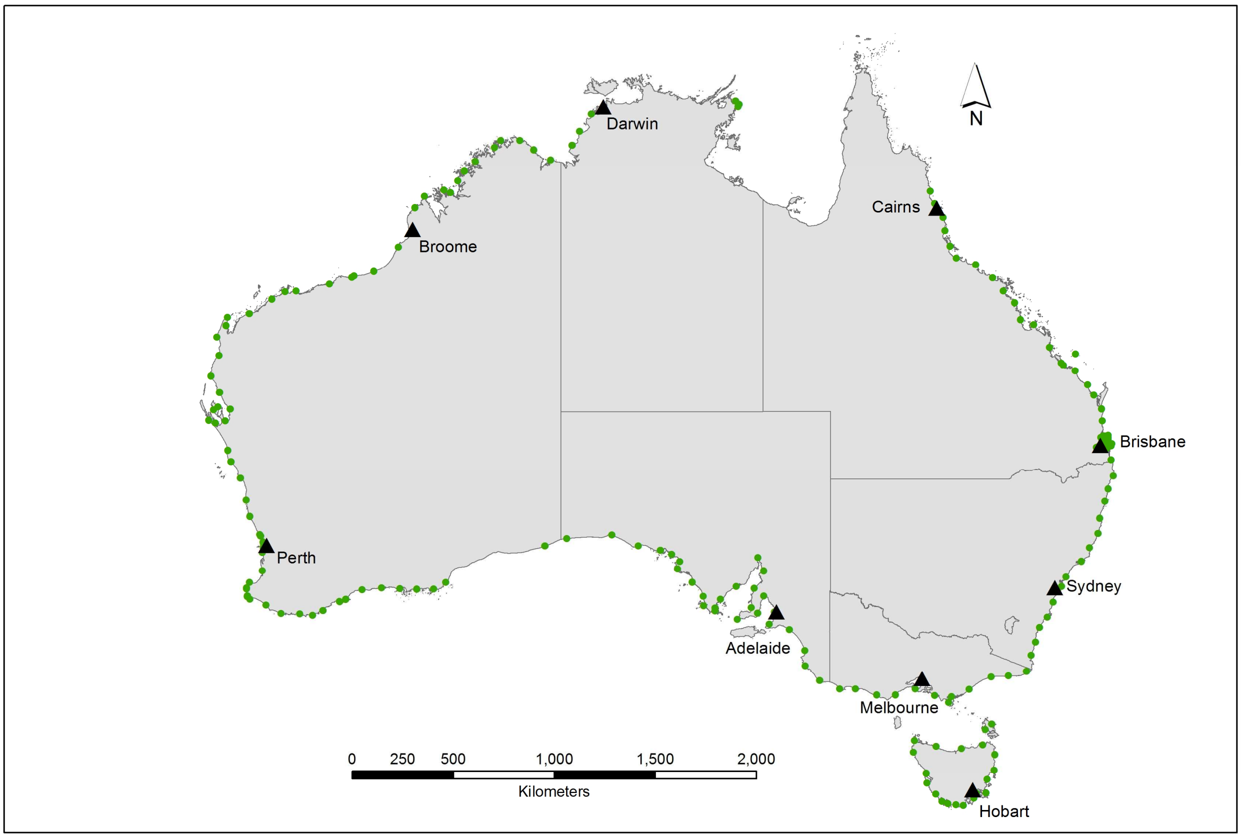

Figure 2. Location of transect surveys carried out by CSIRO staff

4.1.2 CSIRO Fixed area surveys

As part of the CSIRO national marine debris project, a three year partnership was entered into by Shell Australia (funding body), Earthwatch Australia and CSIRO. The goals of this partnership were to:

1) Increase science learning and uptake for individuals, schools, communities and industry across the country and
2) Contribute to decreased marine debris deposition from a change in behaviour based on science learning at local scales.

To achieve these goals, CSIRO and Earthwatch staff developed and led engagement with students (primary and secondary), community groups and Shell employees through the Teachwild "Scientist for a Day" program. The participants of the program learned about marine debris, its impacts on wildlife and how to collect data in a scientific manner. Due to the complexity of conducting scientifically rigorous transect surveys it was determined that fixed area surveys would be best suited for primary school age children. Classes were divided into groups of 10 or less with each group assigned a specific section of the beach. Fixed area surveys were typically 30 meters wide and ran inland from the water's edge to two metres into the backshore vegetation or until a cliff/sea wall was encountered. To be consistent with the transect method used by CSIRO staff, fixed area surveys were conducted at least 50 meters from the beach access point. Where multiple surveys occurred at one site, they were spaced a minimum of 25 meters apart. Survey areas were walked in a regular grid patterns, with observers spaced approximately one meter apart to maintain a constant 1 meter swath per observer. All debris found on the fixed area surveys was collected, identified, categorised and recorded after returning to the classroom. The information collected on the fixed area survey is the same as that collected by the transect method (Section 4.1.3). Data was then entered into the National Marine Debris online database. A total of 38 fixed area surveys were included in analyses (Figure 3).

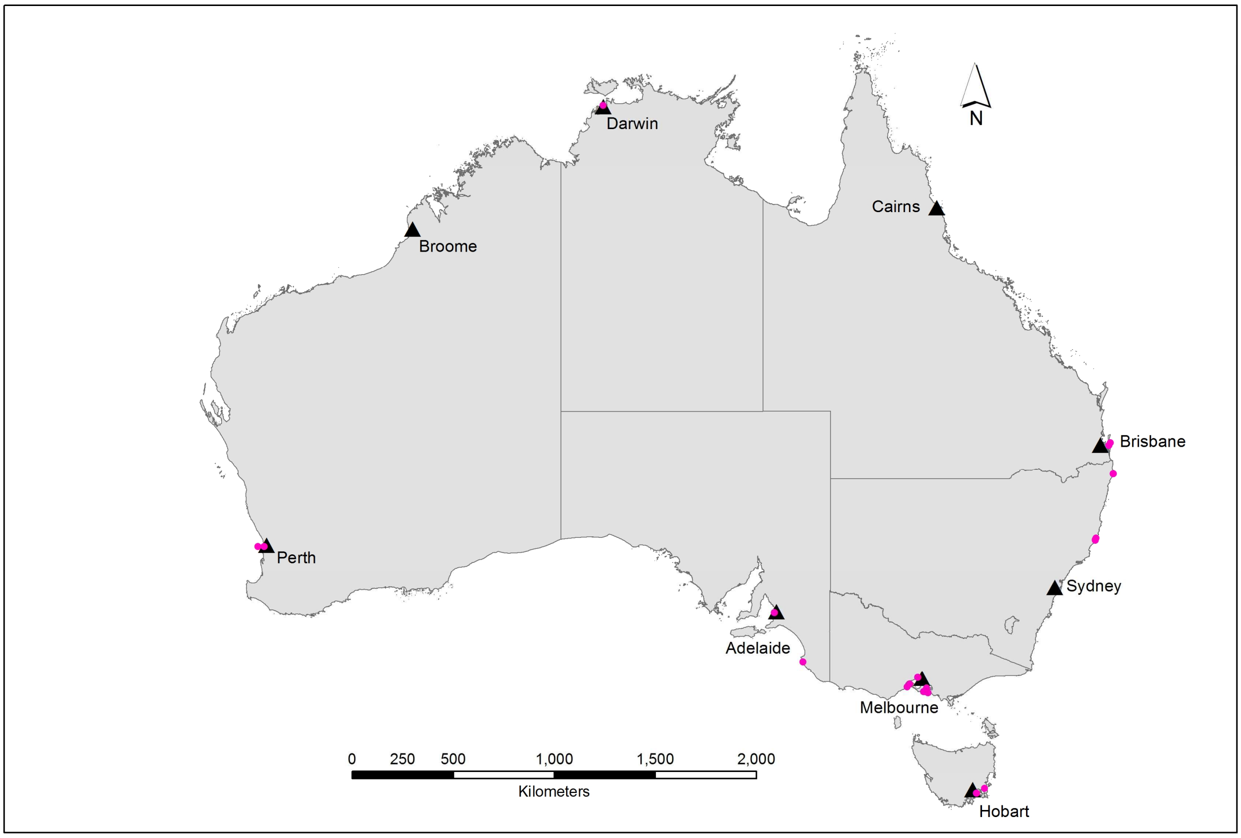

Figure 3. Location of CSIRO fixed area survey sites

4.1.3 CSIRO Public Transects

Transect survey data was collected using the same methodology as the CSIRO coastal debris survey (Section 4.1.1) by secondary students, Shell employees and community groups. Once again the data collected was entered into the National Marine Debris online database. A total of 149 transect surveys were included in analyses (Figure 4).

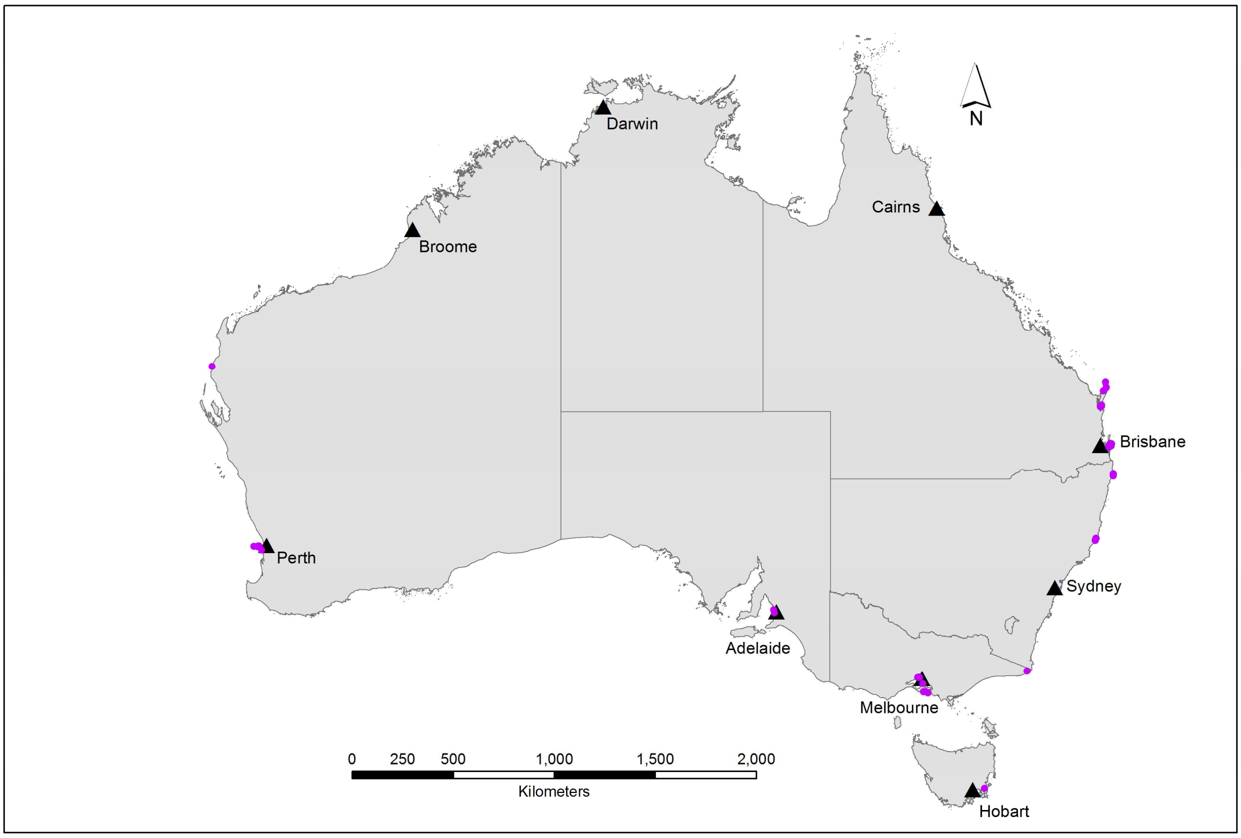

Figure 4. Location of transect survey sites carried out by members of the public following the CSIRO transect method.

4.1.4 Clean up Australia Data

Clean up Australia (CUA) is an annual event where the public are encouraged to complete debris clean ups. Only data from 2007 to 2016 are used in analyses (Table 3, Figure 5). Locational organisers are provided with equipment (gloves, rubbish bags, etc.) and survey forms. Surveys collect details such as the location of the clean up, the area targeted for the clean up, how many people are in attendance and detailed debris categorisation. The information is used by CUA to generate annual reports detailing debris concentrations. Typically, a community group will organise one or more clean ups in their area. Every site is surveyed on the same day each year. Location identification information in the surveys was used to generate a Geospatial Information System (GIS) dataset. Every effort was made to accurately locate the site of the clean up. Sites within 20 meters of one another in subsequent years were assumed to be the same location. Those surveys that did not have sufficient information regarding its location where removed from the data used in the analysis. Also removed were surveys with incomplete or inconsistent records. For example, no assumptions were made about the intended value (number) of items when the answer supplied for a count of a certain type of debris was entered as ‘Many’ or ‘Lots’. Appreciation must be given that the completion and accuracy of surveys lies with the general public volunteers for this event.

Table 3. Count of site surveys provided by CUA and the final number of sites assessed in the analysis over the years of interest.

Year	Sites provided	Sites used
2007	1203	249
2008	1058	226
2009	2870	316
2010	691 (pre cleaned by CUA)	103
2011	2660	244
2012	453 (pre cleaned by CUA)	148
2013	1238	245
2014	540	241
2015	668	172

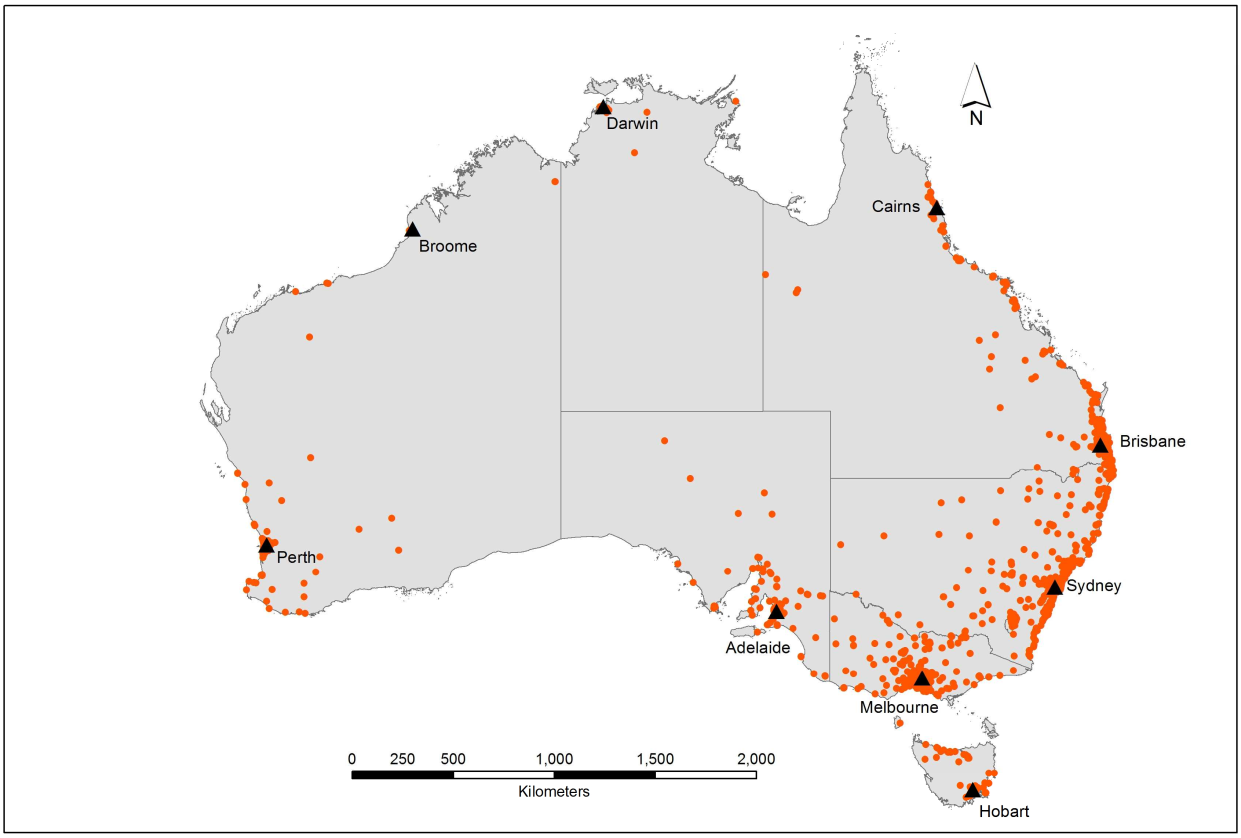

Figure 5. Location of Clean up Australia survey sites used in analyses

CSIRO staff cleaned and verified data collected in CUA surveys. This reduced the number of surveys used in analyses, as per Table 3. The number of sites was reduced in part because not all CUA data included location information (e.g. latitude/longitudes or descriptions that would enable staff to pinpoint locations in geographic space). Furthermore, CSIRO staff then amalgamated data from the 94 categories used by CUA into 21 categories for comparison with CSIRO collected data surveys and methodology.

4.1.5 Keep Australia Beautiful Data

Keep Australia Beautiful and Keep South Australia Beautiful (hereafter referred to as KAB collectively) provided their survey data from 2005 to 2015 (Figure 6). Only data onwards from 2007 was considered for this analysis. This decision was made because of the large changes in the method of data collection and recording prior to and after 2007. Textual and photographic information was provided for each site location. This was used to build a GIS dataset for analysis purposes. Every effort was made to correctly identify the location of surveys. KAB surveys occur at representative anonymous sites nationally several times a year. Between 2007 and 2015 sites were surveyed on average 16 times. The anonymity of the site location allows for an unbiased sample of debris type and volume. From this data, KAB produces an annual report. CSIRO staff cleaned and verified data collected in KAB surveys. Furthermore, CSIRO staff then amalgamated data from the 84 categories used by KAB into 21 categories for comparison with CSIRO collected data surveys and methodology.

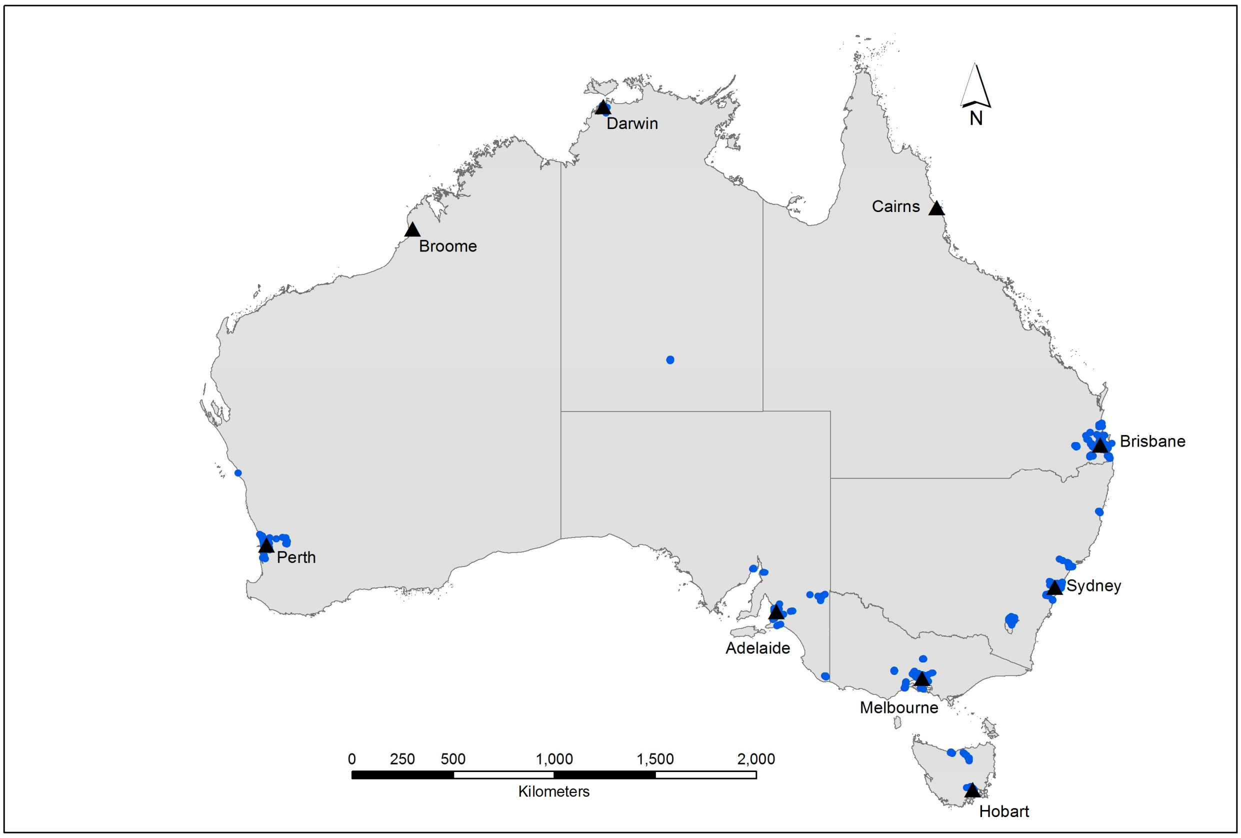

Figure 6. Map of KAB survey locations (2007-2015).

4.2 Geospatial information in relation to survey sites

Data were collected to characterise each sample site and to determine if site characteristics have a discernible influence on debris load and type. A single distance to features was calculated or the density of attributes was collected using the covariate sample design. Covariate segments are concentric circles with a radius of a given distance (1km, 5km, 10km, 25km and 50km) from the survey sites (Figure 7). This allows for an analysis of the distance at which site characteristics have an influence on debris type and volume. Where data can be safely be assumed to be constant over the time that the surveys occurred, the most reliable data source was used. For data that changes through time, data was derived from a linear model of multiple samples across time for the day each survey occurred.

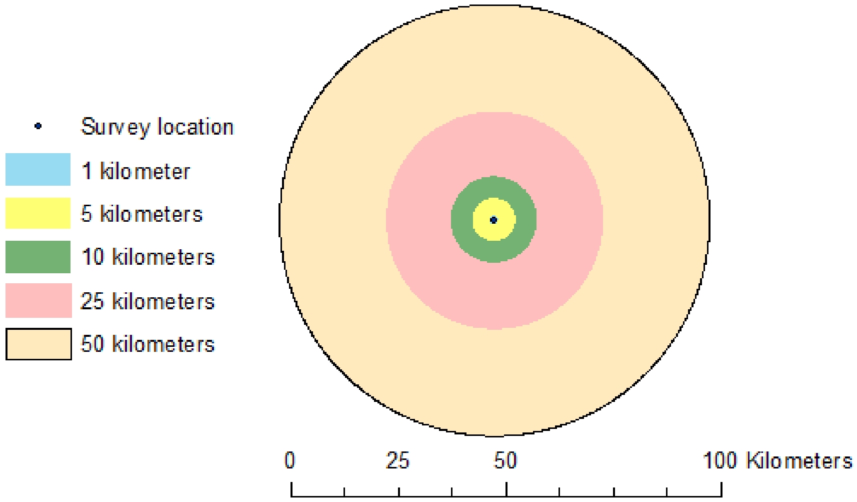

Figure 7. Covariate sampling for each survey location summarised values within a concentric circle with a radius of 1km, 5km, 10km, 25km and 50km.

4.2.1 Roads

The data used in this analysis is the GEODATA TOPO 250K Series 3 Topographic Data (Geoscience Australia, 2006) dataset. The distance to the nearest road was used as a proxy for the potential number of people accessing sites (accessibility). To do this, we determined the distance in kilometres to the nearest road. Sites were further characterised by the total length of different road types (see road classifications, (Table 4) within the covariate sample segments (within 1km, 5km, 10km, 25km and 50km).

Table 4. GEODATA TOPO 250K Series 3 Topographic Data (Geoscience Australia, 2006) dataset road classifications

Road class
Dual Carriageway
Principal Road
Secondary Road
Minor Road
Track

4.2.2 Land cover

Land cover was classified by the Catchment Scale Land Use of Australia (CLUM) dataset developed by the Department of Agriculture, Fisheries and Forestry (DAFF, 2015). Land use is classified according to the Australian Land Use and Management (ALUM) Classification version 7. CLUM is compiled from vector land use datasets collected as part of the state and territory mapping programs through the Australian Collaborative Land Use and Management Program (ACLUMP). Catchment scale land use data was produced by combining land tenure and other types of land use information, fine-scale satellite data and information collected in the field. CLUM is the most recent national land cover product. The ALUM classification defines land cover in three tiers (primary, secondary and tertiary) of classifications (Appendix A) consistently across Australia at a spatial resolution of 50 meters (Figure 8).

The land use identifier that correlated to all three tiers of categories was captured at each site location. The changes in version of CLUM do not necessarily reflect land use change. CLUM is constantly updated using better techniques and more accurate data. For this reason only the latest (most accurate) land use is captured for each site regardless of the time that the survey occurred.

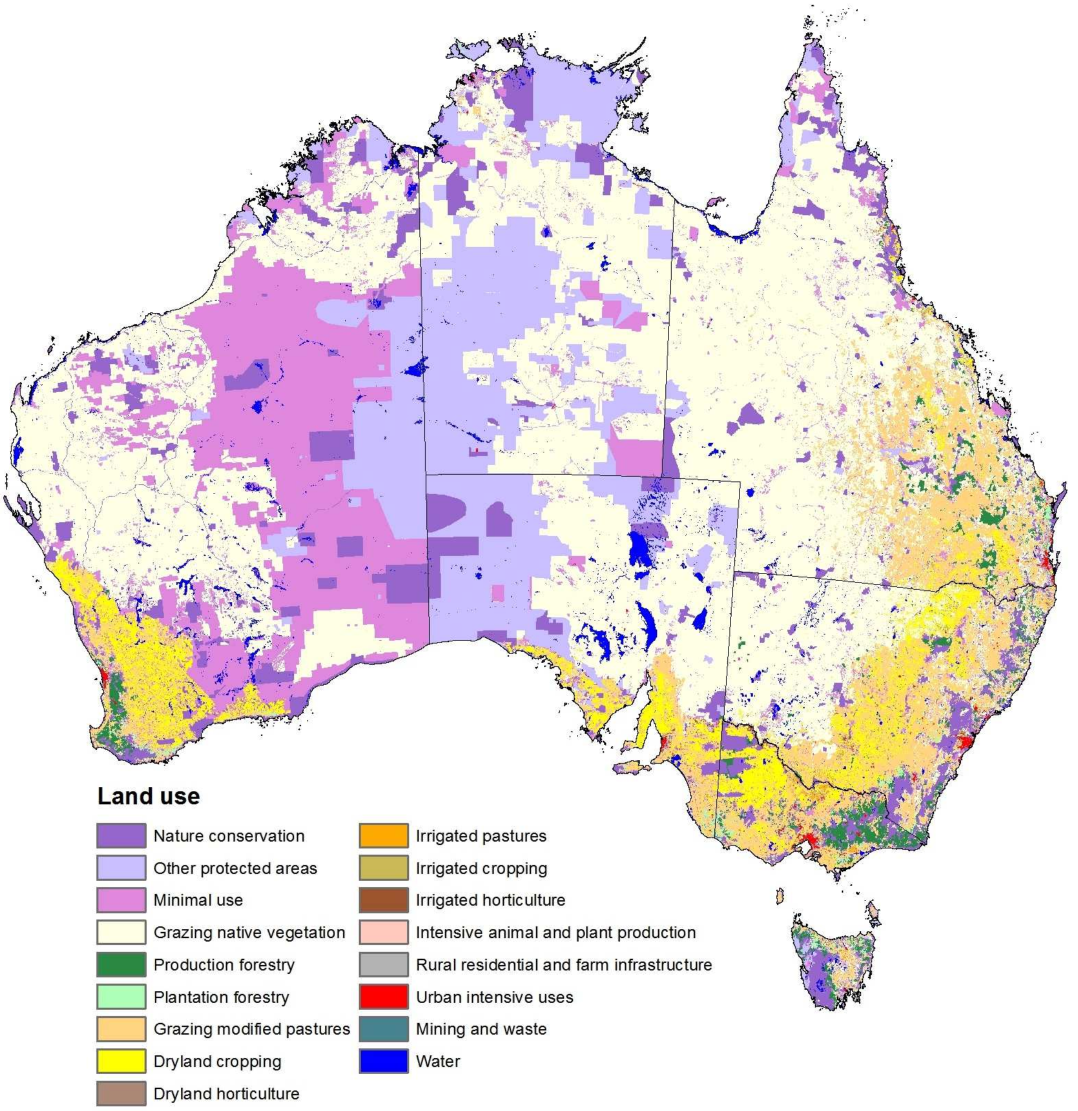

Figure 8. National Catchment Scale Land Use of Australia (CLUM) map showing primary (tier one) categories

4.2.3 Railways

The rail data used for this analysis was the GEODATA TOPO 250K Series 3 Topographic Data (Geoscience Australia, 2006) dataset. As a proxy for determining the accessibility and presence of people on each site we looked at the proximity to railway stations. For this we collected the distance in kilometres to the closest railway station from each survey site.

4.2.4 Population density data

Population census data was evaluated for each of the covariate segments for the date of each survey at each individual site. The highest spatial and temporal resolution data available for this task is the Australian Bureau of Statistics (ABS) Statistical Area 2 (SA2) Estimated Residential

Population (ERP) (ABS, 2016). The geography of this data is according to the 2011 edition of the Australian Statistical Geography Standard (ASGS2011). ERP is the official estimate of the Australian population, which links people to a place of usual residence within Australia. Usual residence within Australia refers to that address at which the person has lived or intends to live for six months or more in a given reference year. To enable the comparison of regional populations over time, historical population estimates based on consistent updated geographic boundaries are prepared for this dataset. The dataset contains ERP from 1991 to 2014. There are 2214 SA2 regions covering the whole of Australia without gaps or overlaps. The ASGS2011's SA2 geographic sampling areas are not the highest possible resolution data available from the ABS (Figure 9), however the ERP at this spatial resolution is high enough to estimate the surrounding population at the date of a survey accurately.

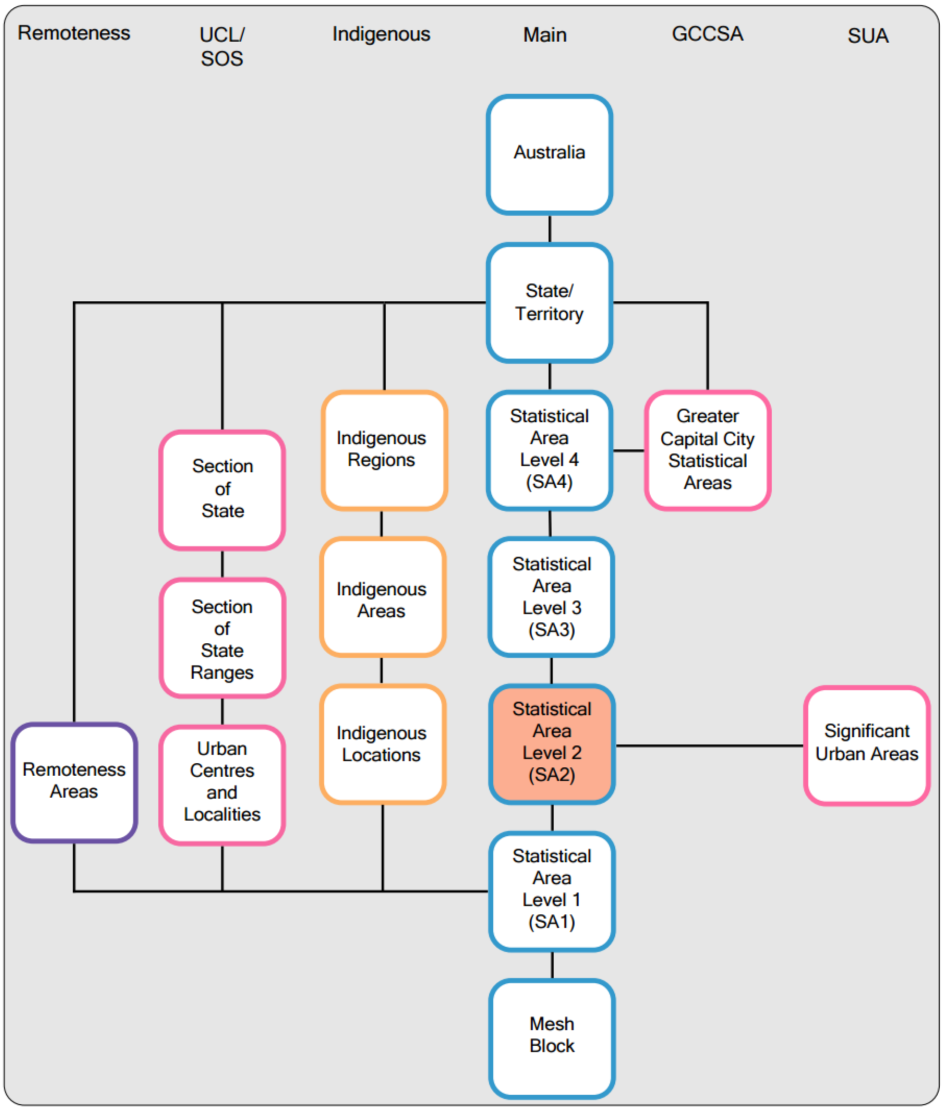

Figure 9. Geographical sampling architecture of the Australian Bureau of Statistics.

Total values were estimated for each site at each sampling distance by summing the percentage area of the SA2 covered by the segment, multiplied by the ERP value for that SA2. From this, a linear model over time was determined and the total values at the date of the survey calculated. Data from 2006 to 2014 was used in this analysis.

For example, the CSIRO site at -27.53° latitude, 152.95° longitude has data for a survey completed 12 July 2016. If we look at the 5km covariate sample segment overlayed on the ABS ERP data (Figure 10) we see that a number of SA2 are intersected. The total population for this segment is calculated for 2006 to 2014.

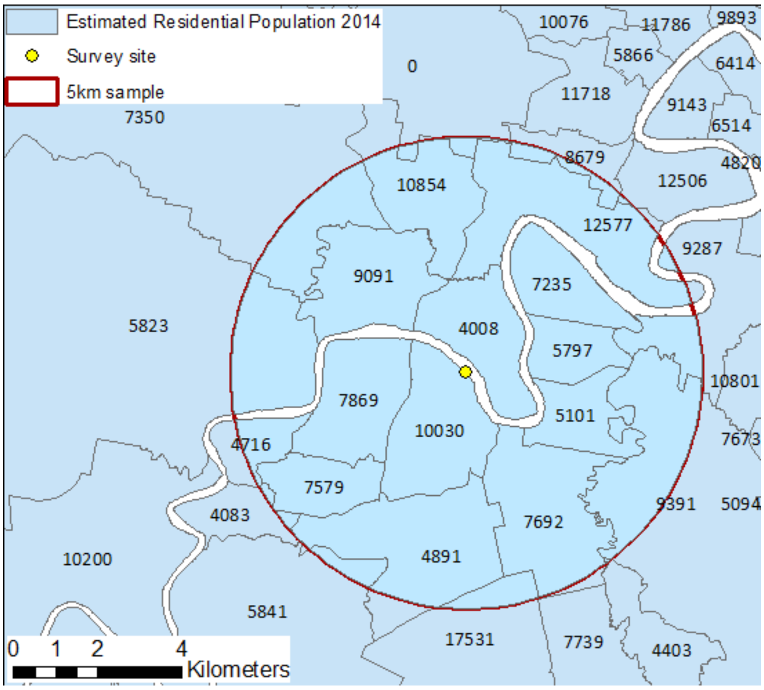

Figure 10. Australian Bureau of Statistics (ABS) Statistical Area 2 (SA2) Estimated Residential Population (ERP) overlayed with the 5km covariate sample segment of a CSIRO survey that occurred near Sinnamon Park, Queensland, 12 July 2016. Numbers are the ERP for the SA2 they fall within.

To calculate the best approximation of the population that existed in this 5km sample a linear model is used. The values used to make up the linear model are the values from the ABS ERP SA2 data. For the date that the site was surveyed, the census value can be calculated by fitting it on this linear model.

We also calculated the population changes across years and the estimated population value that is within a 5km radius of the survey site (Figure 11). Values used and calculated for the 5km sample and model are presented in Table 5.

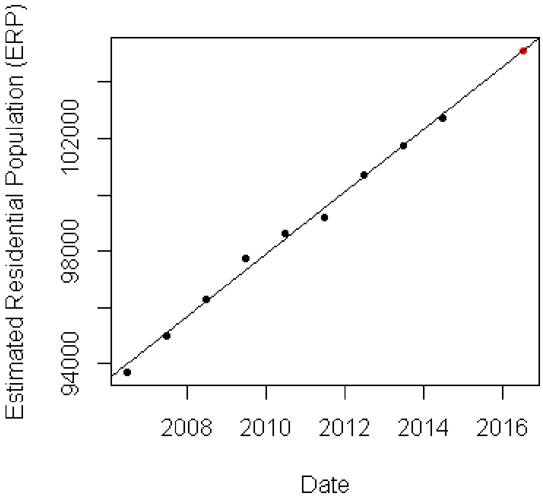

Figure 11. Linear model (black line) and fitted values using the Australian Bureau of Statistics (ABS) Statistical Area 2 (SA2) Estimated Residential Population (ERP) showing are total population within 5km of the survey site (black dots) and estimated population (red dot) for the date that the survey occurred (12 July 2016).

Table 5. ASC Population values within a 5km radius that were used to generate a linear model and derived population for the dates of CUA clean ups and KAB surveys.

Type	Population	Date (day/month/year)
ABS SA2 ERP	93715	30/6/2006
ABS SA2 ERP	95000	30/6/2007
ABS SA2 ERP	96316	30/6/2008
ABS SA2 ERP	97765	30/6/2009
ABS SA2 ERP	98625	30/6/2010
ABS SA2 ERP	99183	30/6/2011
ABS SA2 ERP	100714	30/6/2012
ABS SA2 ERP	101761	30/6/2013
ABS SA2 ERP	102707	30/6/2014
Fitted value	105105	12/7/2016

4.2.5 Socio-Economic Indexes for Areas (SEIF)

The Socio-Economic Indexes for Areas (SEIFA) contains four summary measures from Australian Census data. The summary measures are represented as relative indices for every statistical area in Australia. Each index summarises a different aspect of the socio-economic conditions of people living in an area. They each summarise a different set of social and economic information. The indexes take into account a range of factors in determining socio-economic conditions. SEIF data from 2006 (ABS, 2006) and 2011 (ABS, 2011b) were used in this analysis.

The four indices are:

i. The Index of Relative Socio-economic Disadvantage

The Index of Relative Socio-economic Disadvantage summarises variables that indicate relative disadvantage. The index is designed to focus on disadvantage only. A low score on this index indicates a high proportion of relatively disadvantaged people in an area. We cannot conclude that an area with a very high score has a large proportion of relatively advantaged ('well off') people, as there are no variables in the index to indicate this. We can only conclude that such an area has a relatively low incidence of disadvantage.

ii. The Index of Relative Socio-economic Advantage and Disadvantage

The Index of Relative Socio-economic Advantage and Disadvantage summarises variables that indicate either relative advantage or disadvantage. This index can be used to measure socio-economic wellbeing in a continuum, from the most disadvantaged areas to the most advantaged areas.

An area with a high score on this index has a relatively high incidence of advantage and a relatively low incidence of disadvantage. Due to the differences in scope between this index and the Index of Relative Disadvantage, the scores of some areas can vary significantly between the two indexes. For example, consider a large area that has parts containing relatively disadvantaged people, and other parts containing relatively advantaged people. This area may have a low Index of Relative Disadvantage, due to its pockets of disadvantage. However, the Index of Relative Advantage and Disadvantage may be moderate, or even above average, because the pockets of advantage may offset the pockets of disadvantage.

iii. The Index of Economic Resources

The Index of Economic Resources summarises variables relating to the financial aspects of relative socio-economic advantage and disadvantage. These include indicators of high and low income, as well as variables that correlate with high or low wealth. Areas with higher scores have relatively greater access to economic resources than areas with lower scores.

iv. The Index of Education and Occupation

The Index of Education and Occupation summarises variables relating exclusively to education, employment and occupation. This index focuses on the skills of the people in an area, both formal qualifications and the skills required to perform different occupations.

A low score indicates that an area has a high proportion of people without qualifications, without jobs, and/or with low skilled jobs. A high score indicates many people with high qualifications and/or highly skilled jobs.

Census districts with very low populations, or high levels of non-response to certain Census questions, were excluded from the analysis. Mean SEIF indices were calculated for each sample segment and these were used to derive the SEIF index for the date the survey occurred based on a linear model.

For example, the CSIRO site at -27.53^o latitude, 152.95^o longitude has data for a survey completed 12 July 2016. If we look at the 5km covariate sample segment overlayed on the ABS SEIF (Figure 12) we see that a number of census districts are intersected.

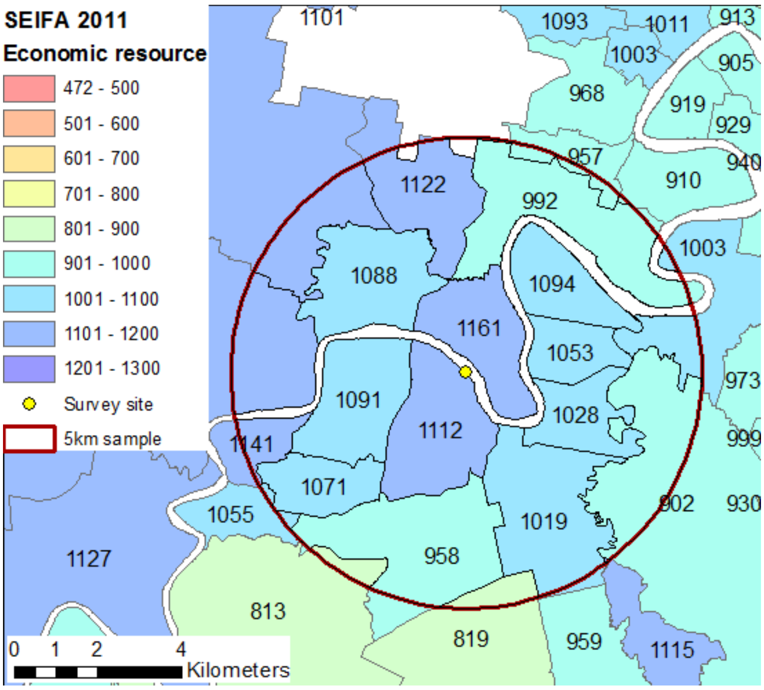

Figure 12. Australian Bureau of Statistics (ABS) Census District (CD) 2011 Socio-Economic Indexes for Areas (SEIFA) Economic Resources overlayed with the 5km covariate sample segment of a CSIRO survey that occurred near Sinnamon Park, Queensland, 12 July 2016. Numbers are Economic resources indices for the CD they fall within.

Mean SEIF indices are calculated for the from 2006 and 2011 within each of the five distances from the surveys (1km, 5km, 10km, 25km and 50km). Means are generated by multiplying the index of a census district by the proportion of total SEIF census district areas intercepted. From this, the mean SEIF indices were predicted using a linear model for the date of the survey (Figure 13).

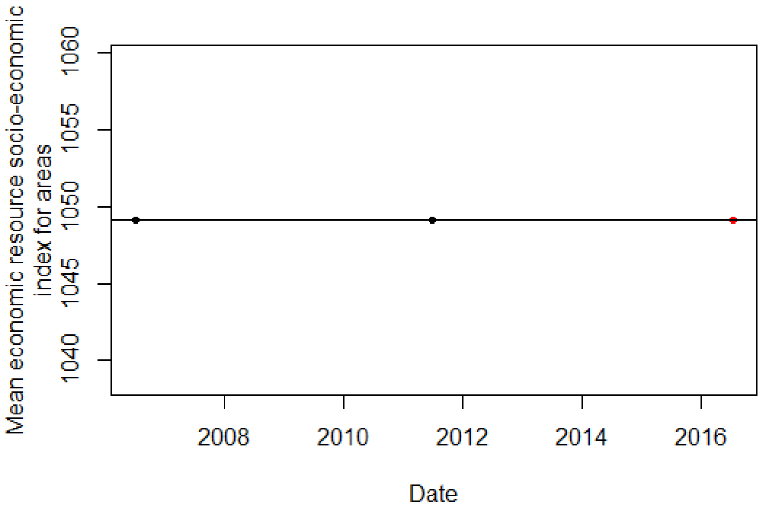

Figure 13. Linear model (black line) and fitted value using the Australian Bureau of Statistics (ABS) Census District (CD) Socio-Economic Indexes for Areas (SEIFA) Economic Resources Index with mean index within 5km of the survey site (black dots) and estimated index (red dot) for the date that the survey occurred (12 July 2016).

4.3 Developing Parameters for the Debris Transport

Our analysis of the transport of debris by wind and water required several types of data that needed to be either collected or developed from other data sources. This analysis required information equivalent to that used in modelling the load of debris in survey data, as outlined in section 4.2. In addition, it required information on wind movement for each location in the study area, outlined in 4.3.2 below. It also required information on surface water movement at the watershed scale in the study area, covered in 4.3.3.

4.3.1 Definition of the study grid and extraction of covariates

The study area for the transport modelling was defined around a single major urban area in Australia, extending the study area boundaries to include the watersheds of all major watercourses passing through that urban area. We chose to focus on the greater Brisbane area, encompassing 3 major rivers (Brisbane, Logan, and Pine Rivers) and a number of minor drainages (Figure 14). The boundaries of the study area encompass approximately 110,000 square kilometres, extending to the outer edges of the three watersheds relevant for the area. We selected the Brisbane watershed because there were numerous surveys conducted both on the coast and along various river and road networks. To ensure adequate representation, we further sampled additional sites following the CSIRO transect methodology (22 additional transect surveys were carried out in July 2016).

To determine the covariate values and forcing functions for transport across the landscape we created a 300m x 300m grid over the study area. We calculated the centre-point of each grid cell and used this point to extract the data. The same study points were used as locations in the calculation of surface water flows and wind velocities.

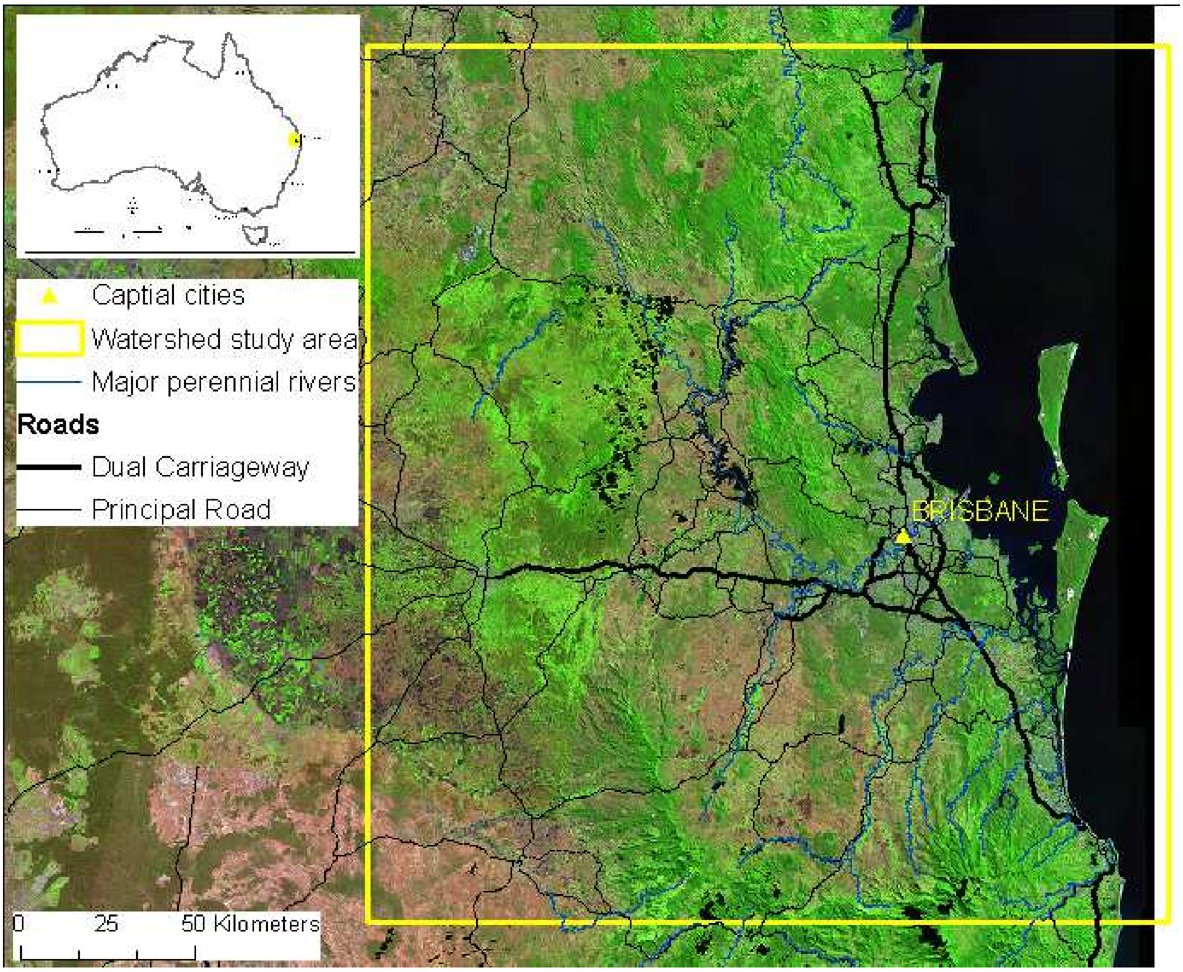

Figure 14. Area included in the Brisbane case study analysis with major perennial rivers and principal road networks.

For each of these grid locations the same covariates as above (5.2.1 – 5.2.5) were extracted from the geospatial database to be used in predicted debris loads at each grid site. In addition to these covariates, we also calculated:

- Distance to each river mouth (Brisbane river, Logan River and Pine River, in km)
- Nearest distance to any point along the three rivers (km) and
- Nearest distance to the coast (km)

4.3.2 Wind Transport model

We modelled transport by wind using the downwind velocity between sites at each pair of points on the 300 m grid. We obtained daily wind speed for each of the 51 weather stations across the Brisbane watershed area (Figure 15). For each of the 51 weather stations we received the daily average wind speed and direction from The Bureau of Meteorology (BOM).

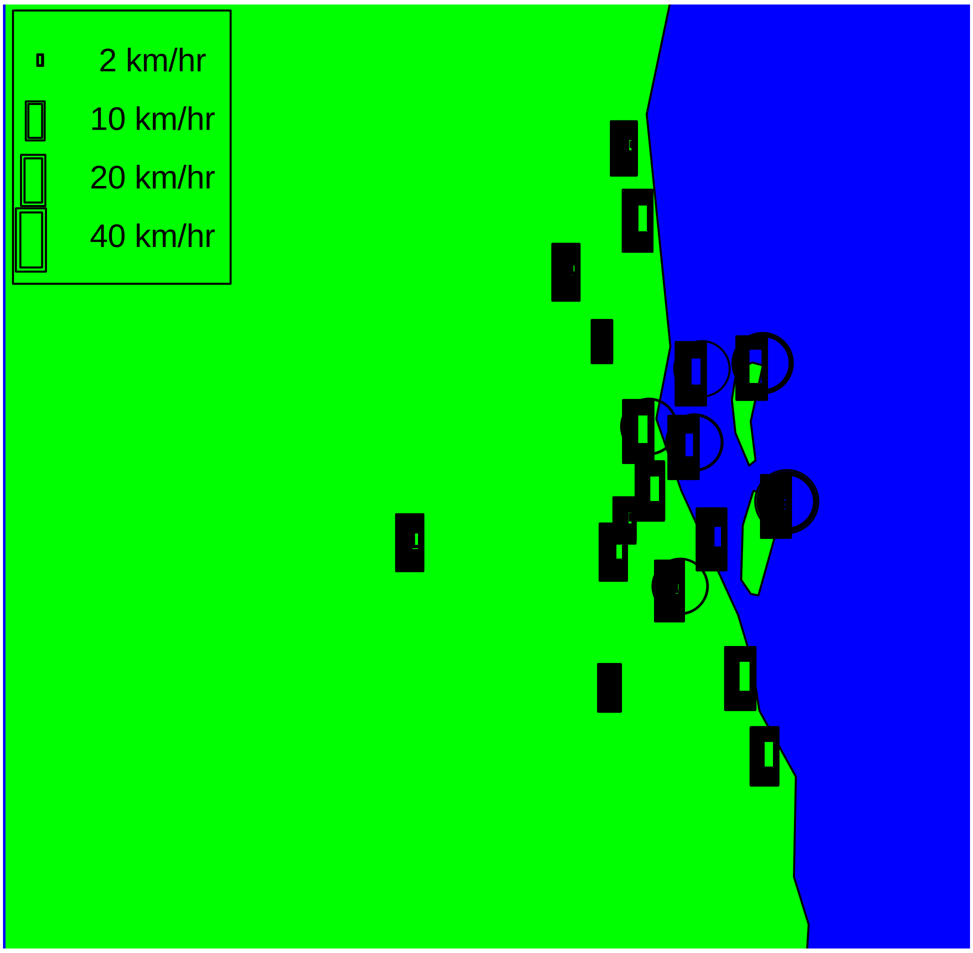

Figure 15. Wind velocities at weather stations in the Brisbane area. Daily wind velocities over the data set are plotted as concentric circles, scaled in proportion to the velocity.

We used these data to evaluate the correlation in speed across the sites, by comparing the velocities to the nearest neighbour among the stations. Sites were locally correlated, but there was significant variation across sites (Figure 16). In addition, the correlation among sites was not evenly distributed with respect to distance. Some sites appear to be more similar to each other than others. Further complicating the pattern is the fact that the wind stations are not evenly distributed throughout the study area covered by the 300 m grid, but tend to be concentrated along the coast.

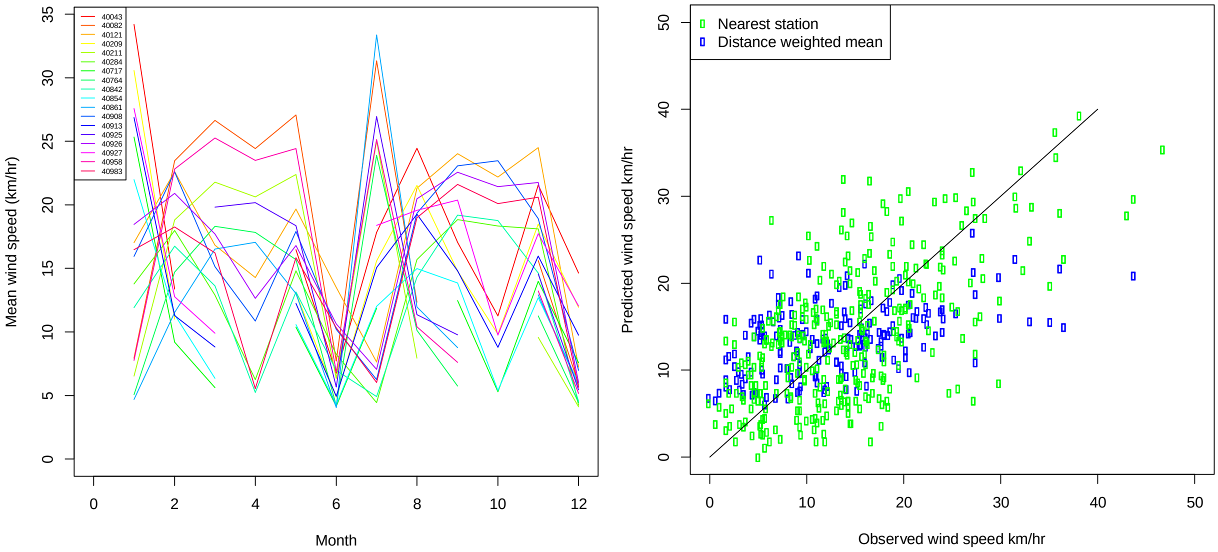

Figure 16. Relationship among wind speeds at stations in the Brisbane region. The left panel shows the monthly mean wind speeds at stations across the region. The right panel shows a plot of the relationship between observations at a station and those at either the nearest station (green) or across all stations, inversely weighted by their distance from the observation location (blue).

We attempted to create an interpolated wind field for the Brisbane region, however, due to the distribution of the stations locations this appeared to give unsatisfactory estimates for many of the 300 m grid locations. We also attempted to use a spatial statistical model to infer velocities at the 300 m grid locations, but again, the concentration of stations at the coast left the statistical model with highly uncertain estimates for the inland portion of the grid. In the end we opted to use the mean daily velocity and direction across all of the observation sites in the region to produce a mean wind speed.

We calculated the downwind speed from the vector component of wind aligned along an axis from a proposed source location to a proposed sink location on the grid. We also calculated the great circle (i.e. following the curvature of the earth) distance between each pair of locations. Due to the very large volume of data, we chose a single year during which we had relatively complete observations to estimate the distribution of velocity and direction. We then took the mean of the daily downwind speed across the year as a measure of transport.

We used this process to estimate two data sets, one providing the downwind velocity from every point on the 300 m grid to each point where we had data on debris load (see Table 2, section 4.1), and a second providing the inverse, the downwind velocities from each location with data to all of the grid points in the Brisbane region. The first data set gives a measure of wind transport to the survey points from all possible sources, the second a measure of the locations that might receive debris from the survey sites where it was measured.

4.3.3 Water Transport model

Topographic slope is an important variable controlling water flow direction and speed. For the purpose of modelling the influence of water transport on debris, the water transport potential was determined from everywhere on the landscape to the survey points was calculated. A flow accumulation map was created from radar interferometric measurements of global land surface heights recorded during the February 2000 Shuttle Radar Topographic Mission (SRTM) (Gallant, 2011). We used the DEM-S product derived from the SRTM. The DEM-S product is a bare earth digital elevation model which is smoothed to remove instrument artefacts. The SRTM data provides national coverage at 1 second (~30m) resolution (Figure 17). The

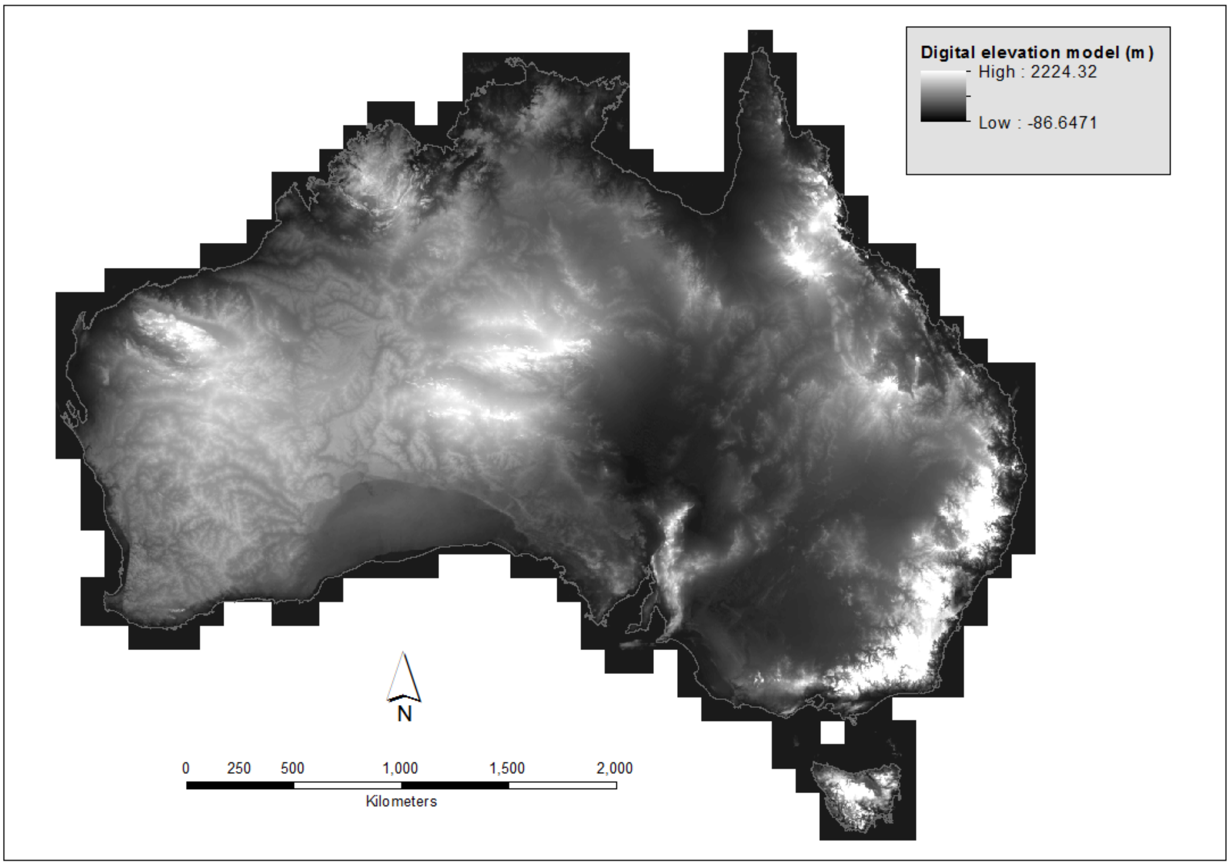

Figure 17. Shuttle Radar Topographic Mission (SRTM) DEM-S digital elevation model (DEM) showing topographical elevation above sea level in meters.

Figure 18accumulated flow (Figure 18) is based on the number of cells flowing into each cell in a downstream trajectory. We calculated the transport of water from any location on the landscape to the survey locations within the Brisbane watershed.

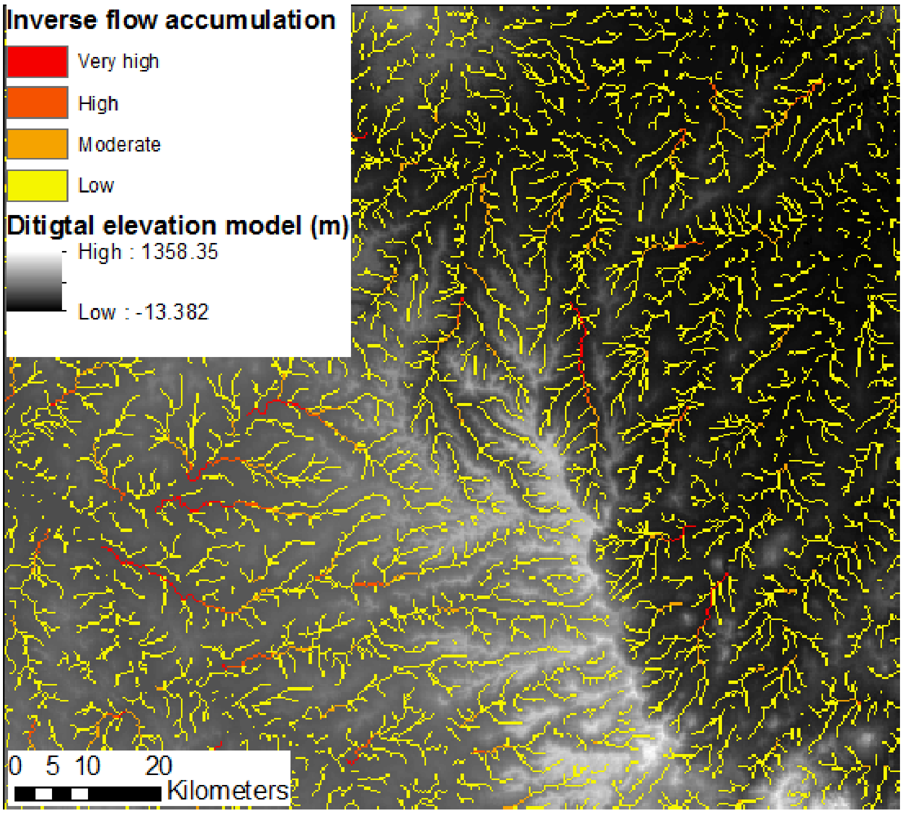

Figure 18. Flow accumulation derived from the SRTM DEM-S overlayed on the digital elevation model.

5 Analyses and interpretation

We conducted three separate analyses aimed at achieving the objectives outlined in Section 2.3. These analyses in some cases share data sets and/or analysis methods. Section 5.1 presents an analysis evaluating the evidence for competing hypotheses about transport of debris at the landscape scale. Section 5.2 discusses the analysis of the KAB data at the national scale, evaluating the geographical, infrastructure, socio-economic and land use drivers of high and low debris loads at the national scale. Section 5.3 summarizes the information available, from both observations and statistical models, to identify hotspots of debris load to support targeting interventions at the national scale.

5.1 Modelling debris with respect to wind and water transport

There were three basic steps in our analysis of debris transport. First, we fit a statistical model to observations of debris densities at survey sites using covariates described in section 4.2. This model was then used to make predictions at locations where there were no surveys using covariates described in Section 4.3.1, allowing us to estimate the debris load across the entire study region. Second, we used the transport models developed in 4.3.3 and 4.3.4 to predict the relative amount of transport between pairs of sites. In this case we used all sites within the study region as sources, but only sites where we had observations as sinks. Third, we compared these predicted loads at the survey sites to the debris loads reported in the survey data. We used the correlation between the predicted and the observed loads at the survey sites to test the support for the three possible transport models, human deposition, wind transport, and water transport.

Testing the relative influence of both wind and water transport on debris at a watershed scale is a multi-step process that involves a tremendous amount of data mining and analysis. Because of the intensive analytical component we limited our analysis to a single major urban area. Even with this limitation, our analysis covered more than 350,000 source locations, with transport modelling delivering debris to a total of 2716 survey locations. The 2716 surveys that fell within the Brisbane watershed area included 2303 KAB surveys, 292 CUA surveys, 109 CSIRO transects, 3 fixed area surveys, and 39 public transect surveys.

We assessed each of the datasets to determine which would be most appropriate to use in predicting debris levels at the source locations on the study area grid. Because each survey type has slightly different techniques, we could not simply amalgamate all of them into one large prediction. KAB surveys report on a standardised area of 1000 square meters, and include a range of debris amounts from 0-3345. The CUA data is not inherently standardised, but the area of the clean up is reported. However, when we standardised the total debris to 1,000 square meters, we had a range of values from 1 – 134,000. CSIRO data ranged from 0-9500 after standardisation, with potentially larger numbers due to the inclusion of small fragments in the sampling scheme. We determined that KAB and CSIRO data had fewer unknown issues that could affect the data, so we restricted our predictions analyses for the transport modelling to these two datasets.

We used both datasets to individually model the relationship between the total amount of debris and a suite of covariates. We used conditional inference trees to develop a predictive model for the total amount of debris at a site, separately for each dataset. Conditional inference trees are an excellent tool for exploration of large complex datasets, and are well suited to making predictions based on sampling data. Essentially this approach uses covariates to split the observations on the variable being modelled into sub-groups which are very similar. The idea is to choose a value of one of the covariates, for instance distance to the nearest road, which if used to split the observations gives two groups with similar values. So, in this case distance to road might be split at 100 meters, giving a set of load observations that are high (and near roads) and low (and far from roads). The process then repeats for each subgroup, with the algorithm examining all of the possible covariates to find the one and its value that best split each of the subgroups from the previous step. This process continues to split the data further down the “branches” until it cannot find a significant difference within a data subset. These models allow for a very complex structure, taking into account multiple covariates and their interactions to explain the variability in the data. The ctree function in the R package partykit was used to implement conditional inference trees for our analysis. We used the entire KAB dataset of surveys at 983 different sites around Australia to create our prediction model. We removed 16 surveys from the analysis, due to the lack of complete set of covariates from these sites. We therefore analysed 15,678 surveys from 982 different locations. The ctree function does not require data to be normally distributed, so we used the total number of debris items per 1000m2 area, as reported by the KAB surveyors. We then repeated the analysis with the CSIRO data, with a total of 665 surveys at 196 sites.

The tree models incorporated the following covariates: Distance to nearest rail, sum of all roads within 50 km of the site, primary land use category, population within 50 km of the site, state, and the suite of socio-economic index factors within 50 km of the site. Additionally, we incorporated measures of the differences in urbanisation and population density close to the site (within 5 km) from those within 50 km around the site. We used the residuals of the relationship between the length of road within 5km of the survey location and 50km of the survey location, and similarly, the residuals of the relationship between the population within 5km of the survey location and 50 km of the survey location. By using the residuals we are able to remove any issues with the correlation between the 5km and 50km values, allowing us to distinguish between areas where there is urbanization both locally and regionally, locally only, regionally only, and neither locally nor regionally.

The KAB tree model was incredibly complex, with 195 branches in the tree. By contrast, the CSIRO tree model was very simple, and split only on population within a 50km buffer. The difference between the two models may in part be due to the variation in the data and the covariates we used to explore them. The CSIRO data has a much wider range, fewer observations, and is all fairly similar in that it samples the coast along with a few estuaries at the national scale. In contrast, the KAB data includes a wider range of sites and characteristics, and resulting counts are somewhat less variable. Given that our primary focus is on using these models to make predictions, as opposed to exploring the variables that explain high and low debris loads, we have not included the tree structures in the main portion of the report.

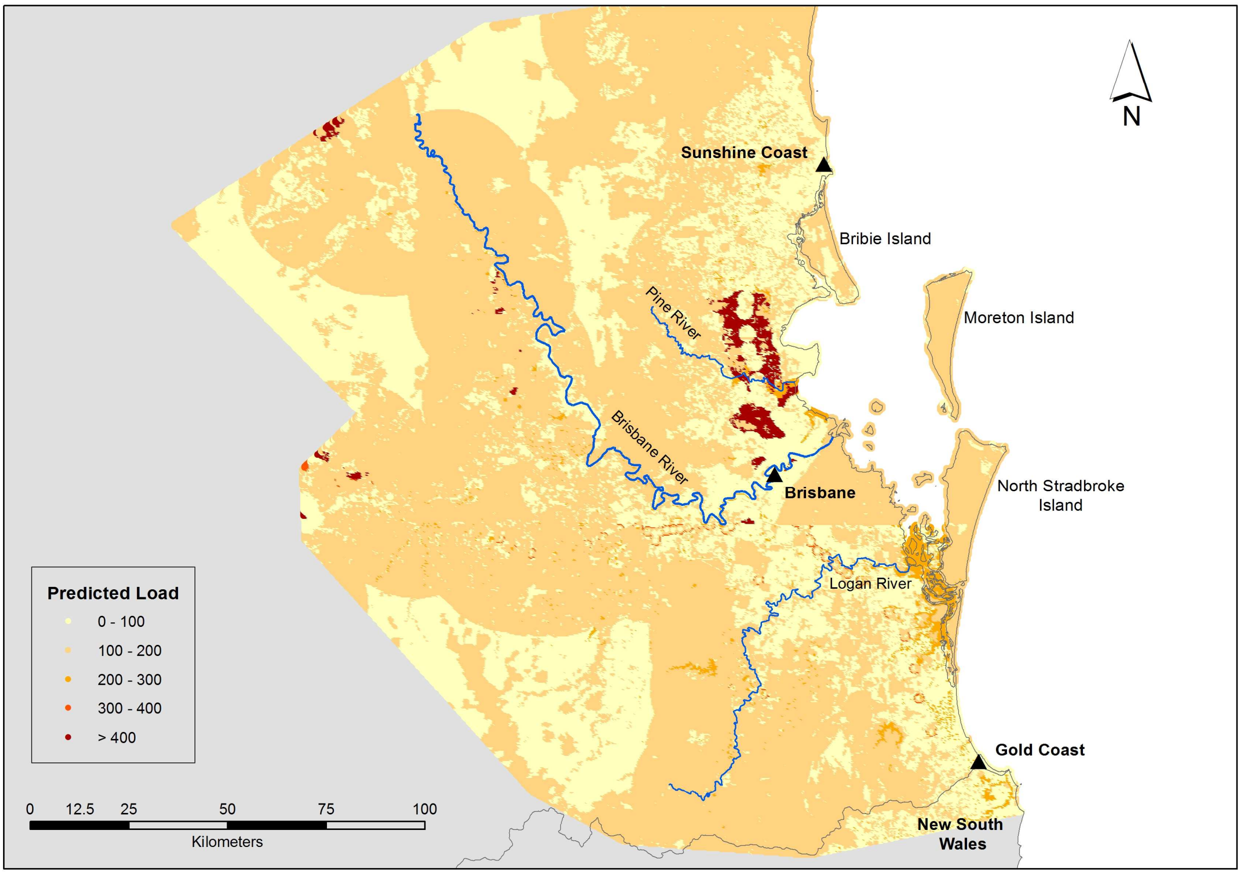

Figure 19. Predicted amount of debris within each 300m grid cell within the Brisbane watershed. The model used to create this prediction was trained on the KAB dataset, and applied to the covariates on each individual grid cell.

Next, we used the KAB tree model to predict the amount of debris, or debris load that would be found in every cell of the 300 meter grid of the greater Brisbane area, based on the covariates associated with that grid cell (Figure 19, Section 5.3).

We used wind and water transport functions to model the influence of upwind and uphill sites on the amount of debris load expected at each survey site in the greater Brisbane region. We posed a number of hypotheses about how these transport functions would be related to the debris load at the 416 survey sites in the study region where we had observed data. The labels in parentheses after each hypothesis below correspond to the test names in Table 6.

H0: Debris is deposited in place and does not move, therefore there will be no correlation between the predicted transport to a site (by wind or water) and the amount of debris at that site.

H1: The total number of grid cells that are upwind (or uphill from a given grid cell will be positively correlated with the amount of debris at that site. In other words, the more potential inputs, the greater the total amount of debris at a given site (Upwind, Uphill).

H2: Debris will increase relative to the number of upwind grid cells, but the total amount of debris will be inversely proportional to the distance from each upwind (or uphill) site. So if a cell is upwind but far away, it will contribute less debris than a cell that is upwind but close (Wind Sink dist).

H3: Debris inputs will be proportional to the sum of the magnitude of the wind flow from upwind grid cells. (Wind Sink)

H4: Debris inputs will be scaled by the total amount of debris predicted to be found in each source grid cell in either an absolute sense (Upwind Debris) or a proportional sense (Wind Sink Prop Debris, Water Sink Prop Debris), and may also include an effect of distance between sites (Wind Sink Dist Debris).

H5: Debris inputs will be proportional to the sum of the magnitude of the water transport between all source cells and the sink cell (Water Sink). *Note that the water transport model is calculated differently from the wind transport model, and already incorporates both elevation difference and distance.

H6: Debris inputs will be proportional to the sum of the magnitude of the water transport between all source cells and the sink cell, weighted by the predicted load in the source cell (Water Sink debris). *Note that the water transport model is calculated differently from the wind transport model, and already incorporates both elevation difference and distance.

We used the wind and water transport models to create predictions for each of the survey sites based on each of our hypotheses. For example, for H3 (Wind Sink) we would use the average wind speed from each possible source on the grid for a survey site (i.e. all sizes with a nonzero downwind velocity for a given survey site), and sum those velocities to get an estimate of the amount of transport to that site. That would be done for each of the 416 sites where we had observational data, to get the estimates for those 416 sites for hypothesis H3. To estimate the values for H4 (Upwind Debris) we would use a similar process, with the addition that each site would also be multiplied by the debris load predicted at that site by the KAB tree model (Figure 19).

Next we tested the correlation between the predicted debris loads at survey sites based on each of the hypotheses above with the total amount of debris observed at each survey site. We also tested the predicted loads under each hypothesis against the residual variation in the data, left over after fitting the tree and GAMs (Generalized Additive Models; see section below) to the survey data. In regression models, such as tree and GAM models, residuals are a measure of the variation in the data that is unexplained by the model. Thus, if a residual is negative, it suggests that there is less debris at a site than would be expected, given the site characteristics. Similarly, if a residual is positive that indicates that there is more debris than would be expected at a site given its characteristics. These residuals could be related to transport, as negative residuals might be expected to occur in locations where debris is transported away from a site, leaving less than would otherwise be expected given the site characteristics alone. Similarly, positive residuals might occur at sites where debris is transported to the site, meaning more appears there that would be expected from site characteristics. So negative residuals mean that more debris is found at the site than would be predicted by the covariates at that location. In other words, the site is a debris sink. Positive residuals indicate that there is less debris than would be predicted, and the site is a debris source.

Our transport functions and the hypotheses around how they are related to the data are abstractions of the full mechanistic process. For instance, we do not consider buoyancy or drag of the debris items, nor do we consider friction or surface roughness which all likely affect transport. Given the level of abstraction, we expect our semi-mechanistic models to produce predictions that are correlated with the observations from the survey data if they are correct. However, we do not expect them to produce predictions that are accurate on an *absolute* scale. Because our predictions are on a relative scale, we used a correlation test between the predictions and the observations to evaluate the support for each of the hypothesized transport mechanisms. The Spearman rank correlation test compares the rank of two different vectors to determine whether or not they are correlated. It does not distinguish between values, simply between their rank order. This test is useful in this instance, because although we can predict where the debris might travel with wind and water, it is much more difficult to predict how much debris will move, or how far it will move. The rank test, therefore, is independent of the actual magnitude of the debris, and compares the relative order of debris at different survey sites.

We used the same methodology for KAB and CSIRO data, and tested the predictions against both the actual amount of debris as well as against the two residuals measurements. The results are summarised in Table 6 below.

Table 6. Spearman rank correlation between observed debris load at each survey site and predicted load based on transport function hypotheses. KAB and CSIRO Ctree and GAMs were created individually from the full Australia-wide data sets. Rho is the correlation between the rank of a site out of the 416 sites (thus 1 has the lowest debris, site 416 the highest). The statistical significance is based on the p value, which is the statistical likelihood of the result. The low the p value, the more likely the correlation is due to the hypothesized effect. By convention, p value of less than 0.05 is taken to mean there is strong evidence for support of a hypothesis. For convenience, "*", indicates a p value less than 0.05.

Hypothesis	**a. Observed Debris**				**b. Tree residuals**				**c. Gam residuals**		
KAB Data	rho	p			rho	p			rho	p	
Predictions only	0.642	0.000	*		-0.073	0.000	*		0.121	0.000	*
Upwind	0.041	0.047	*		0.016	0.450			0.030	0.149	
Wind Sink	0.041	0.048	*		0.016	0.441			0.029	0.169	
Wind Sink with distance	0.065	0.002	*		0.013	0.522			0.015	0.487	
Upwind with debris	0.042	0.046	*		0.016	0.447			0.030	0.149	
Water Sink	0.073	0.000	*		0.055	0.008	*		0.056	0.007	*
Wind Sink prop. debris	0.007	0.729	*		0.005	0.817			0.011	0.591	
Wind distance prop. debris	0.013	0.528	*		-0.041	0.049	*		-0.003	0.868	
Uphill	-0.175	0.000	*		-0.017	0.412			-0.050	0.017	*
Water Sink debris	0.168	0.000	*		0.017	0.418			0.064	0.002	*
CSIRO Data	rho	p			rho	p			rho	p	
Predictions only	0.489	0.000	*		-0.281	0.003	*		0.120	0.212	
Upwind	-0.297	0.002	*		0.079	0.415			0.061	0.527	
Wind Sink	-0.243	0.011	*		0.118	0.221			0.085	0.380	
Wind Sink with dist	0.345	0.000	*		0.054	0.574			0.071	0.465	
Upwind with debris	-0.297	0.002	*		0.079	0.415			0.061	0.527	
Water Sink	-0.387	0.000	*		-0.083	0.391			0.036	0.708	
Wind Sink prop deb	-0.398	0.000	*		0.169	0.079			-0.133	0.169	
Wind dist prop deb	-0.192	0.045	*		0.142	0.142			-0.116	0.228	
Uphill	0.333	0.000	*		-0.010	0.915			-0.050	0.609	
Water Sink debris	-0.400	0.000	*		-0.123	0.204			-0.057	0.555	

Overall, the correlations between the predictions under the various hypotheses, and the values at each of the sites were strongest when we considered the observations (Table 6a), as opposed to considering the residual variation, either from the tree models (Table 6b) or the GAM models (Table 6c). Examining these relationships with the observation data first (Table 6a), the strongest relationship was with the predicted values from the tree model. This suggests that the site characteristics are more important than transport processes for determining the load at a site. This is likely due to direct deposition by people at the site. For the KAB data, the next strongest relationship was with the number of uphill sites in the watershed of the survey site (H1). Interestingly, the number of uphill sites is negatively correlated with the load observed at a site. Thus, as the area above a site in a watershed increases, the site tends to have less debris. That could be due to an increase in runoff passing through the site, transporting debris away from it. The next strongest relationship for the KAB observations was with the water transport between sites, incorporating the load at the source site (H6). This transport function is modelled based on the elevation and distance between the source and the sink site, as detailed in Section 5.3.3. It also incorporates the load of debris at the source site. The correlation is almost equivalent to that for the uphill sites, but positive as one would expect. Thus the debris load at a survey site goes up as more sites with large predicted loads and greater water transport export debris to the site. The next strongest relationship for the KAB surveys was with the wind transport, taking account of the mean velocity and distance between the source and the survey site (H3). However, in contrast to water transport, incorporating the load at the source sites (H4) did not improve the predictions.

Comparing the predicted values from each of the transport function hypotheses with the observations from the CSIRO survey data yields somewhat different patterns. Again, all of the transport models make predictions that are significantly correlated with the values observed during the surveys. For the CSIRO data, the strongest relationship was again with the water transport, incorporating the load at the source sites (H6), as was the case for the KAB data. However, for the CSIRO data this relationship appears to be negative. Thus more load at source sites that have large flows to a survey site, the less debris was observed at a survey site. The number of sites uphill from a survey site also has the opposite sign of the KAB comparison. For the CSIRO data, the sign is positive, indicating that as there are more uphill sites from a given survey location, there is more debris at the survey sites.

One possible explanation for this difference is related to the survey methodology, and the survey locations. The KAB surveys are focused on relatively large intact or partially intact items. By contrast, the CSIRO surveys include these items but also extend down to smaller fragments, on the order of 1.0 mm in diameter. Importantly, these small fragments are typically much more common in waterways and on the coast that the large items. Second, the KAB surveys are distributed throughout the Brisbane region, with a small number along watercourses. By contrast, the CSIRO surveys in the area are along the Brisbane River, on the shores of Moreton Bay, and in the surrounding open coastline. Thus the CSIRO surveys are primary at the very bottom of the watershed, and encompass much smaller items.

Considering the differences in the size of the items and the location of the surveys, the differences in the important water transport processes does make some sense. The KAB surveys are primarily up higher in the watershed, closer to the source locations, and cover larger items. Thus one might expect to see large volumes of water flushing debris items out of the location, but proximity to sites with heavy loads could lead to additional loads at the survey sites. By contrast, the CSIRO surveys are measuring debris much further down the watershed, and likely further along in its transport process. Given this, a CSIRO site close to lots of sites with high loads is likely to be up in the Brisbane River, as opposed to out in Moreton Bay or along the open coast. Given that the surveys are at the water line, water flow will likely be higher, potentially scouring the debris from the site. Lower down in the river, and out in Moreton Bay, there is likely to be much less scour due to directional movement of water, and onshore transport due to wave action and wind, resulting in higher concentrations. These Moreton Bay sites will also have higher counts of sites uphill from them, given they are at the bottom of the watershed. So, in summary, one could reasonably expect strong effects of load and water transport on both sets of survey locations. In the KAB case we are likely seeing the direct effects of debris transport to survey sites. In the CSIRO survey site case, we are likely seeing the opposite result due to the effects of stronger retention of debris at sites at the bottom of the watershed.

Turning to wind transport for the CSIRO data, we see a strong positive correlation with the number of sources that have downwind transport to a survey site. Incorporating the debris load causes the sign of the correlation to change, resulting in a negative relationship between the prediction and the values observed at the CSIRO survey sites. Our explanation for this pattern is that wind transport is likely important for the CSIRO survey sites, with additional transport from sources toward a survey site resulting in more debris at the site. The change of sign, when load at the source is incorporated, is likely due to loads being related to the position in the watershed. Loads are highest away from the coast, in the mid and upper portion of the watersheds. Thus survey sites close to high load sites are likely to be in the Brisbane River, and thus affected by scouring due to directional flow toward Moreton Bay.

The patterns in the correlations between the transport model predictions and the residuals of the tree and GAM model predictions of debris loads at the KAB survey sites were not nearly as strong, or as clear, as for the observations themselves. This is similarly true for the comparison of the predictions and the residuals for the CSIRO survey sites. For the CSIRO sites, none of the predictions was significantly correlated with the residuals for either the GAM or the tree models. For the KAB data, the relationship with the GAM residuals mirrored that with the observations. Thus, water transport appeared to explain additional variation among the sites that was not captured by the GAM model for direct deposition. For all three data sets, the water transport from source sites was correlated with the values at the survey sites. While this result was not the strongest correlation for the observation data, it was significantly correlated with the residuals from both the tree model and the GAM. This suggests that the number of source sites above a survey site does explain additional variation in the data, on top of that explained by direct deposition at a site based on the site characteristics. Wind transport did not show any clear pattern with the residuals for either the GAM or the tree model. There was a weak negative relationship with the wind model including velocity, distance and load at a site. However, it is unclear how to interpret this result.

The difference in the strength of the relationships between the predicted loads and the value of the observations versus those of the residuals is reasonable. The tree and GAM models will both try to account for any pattern in the data, including that due to (unobserved) transport. For instance, one important variable in the both models is the site time. Moreover, watercourses are the highest of all the site types in terms of debris loads in these models. Thus, the model is probably already capturing water transport to some extent, via estimating that creeks and rivers are expected to have high loads. It is important to note that this is not to say the model includes transport by water explicitly, it just recognizes that if a site is in a creek it is expected to have higher loads. Thus, if a site is in reality affected by transport of debris from other sites to the site, it still may not have a very strong pattern in its residuals as the statistical models have already captured that pattern. In this context the residual for the hypothetical creek site would be expected to be zero, and thus even though it is likely to be strongly affected by water transport processes, that would only show up in the observation data, but not in the residuals from either the GAM or the tree models.

5.2 Exploring drivers for debris loads at the national scale

Conditional inference tree models can be quite complex. While they are useful for prediction, their complexity does not lend itself well to exploring patterns and understanding drivers in data. This is particularly true when they have a large number of branches, as in the case of the tree model for the KAB data. In order to explore the relationships between the covariates and the total amount of debris at a survey site, we fit a generalised additive model (GAM) to the full Australia-wide KAB and CSIRO data sets. Instead of partitioning the data into homogeneous subsets, the focus of GAM models is on fitting functional relationships between a variable of interest, in this case the debris load at a site, and potential explanatory variables, also called covariates. This allows us to understand how debris load varies at different levels of the covariates as they interact with each other.

The total debris data from KAB surveys varies widely, with loads distributed between 0 and 3345 items/1000m^2. The median value is 40 items/1000m^2, while the mean is 72 items/1000m^2. Because of the large number of small observations, the data are not normally distributed, so we transformed them with a log transform. The distribution of CSIRO data is similar, with loads distributed between 0 to 9500 items per 1000m2. The median value is 40.5, similar to the KAB data, while the mean is significantly higher, at 227. This indicates several outlying values in the CSIRO data. We also log transformed the CSIRO data.

We incorporated the same covariates as we had for the conditional inference tree models (see Section 5.2), and also included a spatial smooth for the KAB data. We did not include spatial factors for the CSIRO data, because the data are not distributed evenly across latitude and longitude; instead they form a ribbon along the coast of Australia (Figure 2) In order to control for site-level variability, we included site as a random factor. To answer the question of whether certain site types tend to have different debris levels, we incorporated site type into the KAB model. This information was unavailable for the CSIRO data, so we did not include it.

In contrast to conditional inference tree models, there is no automated selection process for GAM models to find the best model. For both CSIRO and KAB data we began by fitting a model with all smoothed terms. We dropped insignificant smooth terms from the model, then graphed each significant smooth term, and where appropriate, moved them into a parametric form. If a parametric term was significant, we left it as parametric; if not, we moved it back into a smooth.

For the KAB data we found that a model incorporating population within a 25 km radius had a smaller AIC value than one incorporating population within a 50 km radius, so the final model is based on the 25km value.

5.2.1 KAB Survey Analysis

The final KAB GAM explained 63.7 % of variability in the data. Table 7 provides the parameter values for each of the terms in the model, the standard error (a measure of uncertainty) for the parameter value, the t and p values which measure the statistical significance (by convention, a p value < 0.05 is considered to be significant statistically), the median value of the variable for that the coefficient is multiplied by, the product of the coefficient and the median (which measures how strong the effect is, and can be compared among variables), and a ranking of these products (or effect sizes) and a ranking of their absolute values.

The four socio-economic variables (Economic advantage, Economic resources, Education and occupation, and Economic disadvantage) had the strongest effect on the debris load at a site (Table 7, see Absolute effect rank and Median effect size columns). Interestingly, debris levels are higher in areas with a high proportion of economic disadvantage (within a 50km buffer), but also in areas of higher overall economic resources (Eco_resource). Conversely, debris levels are lower in areas with relatively high socioeconomic status and with relatively high levels of education and occupation. However, the effect of the education and occupation term is much weaker, ranked as 13th in the order of effect sizes among the model terms (Table 7).

The next most important factors in determining the debris load at a site are the site type and land use. Site types, as classified by KAB during their surveys with elevated levels of debris were Highways, Retail Strips, and Car Parks. Interestingly, all site types recorded by KAB were elevated with respect to Beach, which was included in the intercept as the reference level. Land use was also important in determining loads, with water in particular being associated with elevated debris loads. Production landscapes in a relatively natural state, such as grazing lands, were also associated with elevated debris levels. Thus in terms of debris loads, sites with one of these three site types and a creek or wetland, or open area nearby would be expected to have particularly heavy debris loads.

The political unit (state/territory) was an important predictor of the density of debris at a site, for the Northern Territory at least. This estimate should be taken with some caution, however, as the KAB survey data includes only a few sites in the Northern Territory. In general, all of the available datasets are relatively sparse in the Northern Territory in comparison with the other states. The political unit was not particularly important for the other states, with effect size ranks in the range 14 to 23 (Table 7).

A number of measures of urbanization are relevant for debris loads. Debris levels increase with the population within 25km of the site and proximity to rail stations. The levels decrease with increasing length of roads within a 50 km, and to an even greater extent when roads are particularly concentrated near the survey site. However, these factors have relatively low effect sizes, suggesting they are less important than socio-economic variables, land use, and site type.

Overall the results paint a picture of high debris loads in relatively economically disadvantaged areas, particularly where site types are related to transitory human use (highways, shopping, parking lots). This is particularly true when those site types occur in a context where water courses or other semi-natural areas occur nearby. Together these results paint a picture of two general types of debris sources, littering near transitory locations (such as parking lots) and potentially illegal dumping in natural areas. Both occur more frequently in socio-economically disadvantaged areas. Social context may be important in determining these behaviours, as site types where people are not transitory (i.e. residential) or where there are recognized aesthetic or public use values (beaches, parks) appear to have particularly low levels of debris (Figure 20). Basically, the site types fall nearly into two non-overlapping groups in this respect.

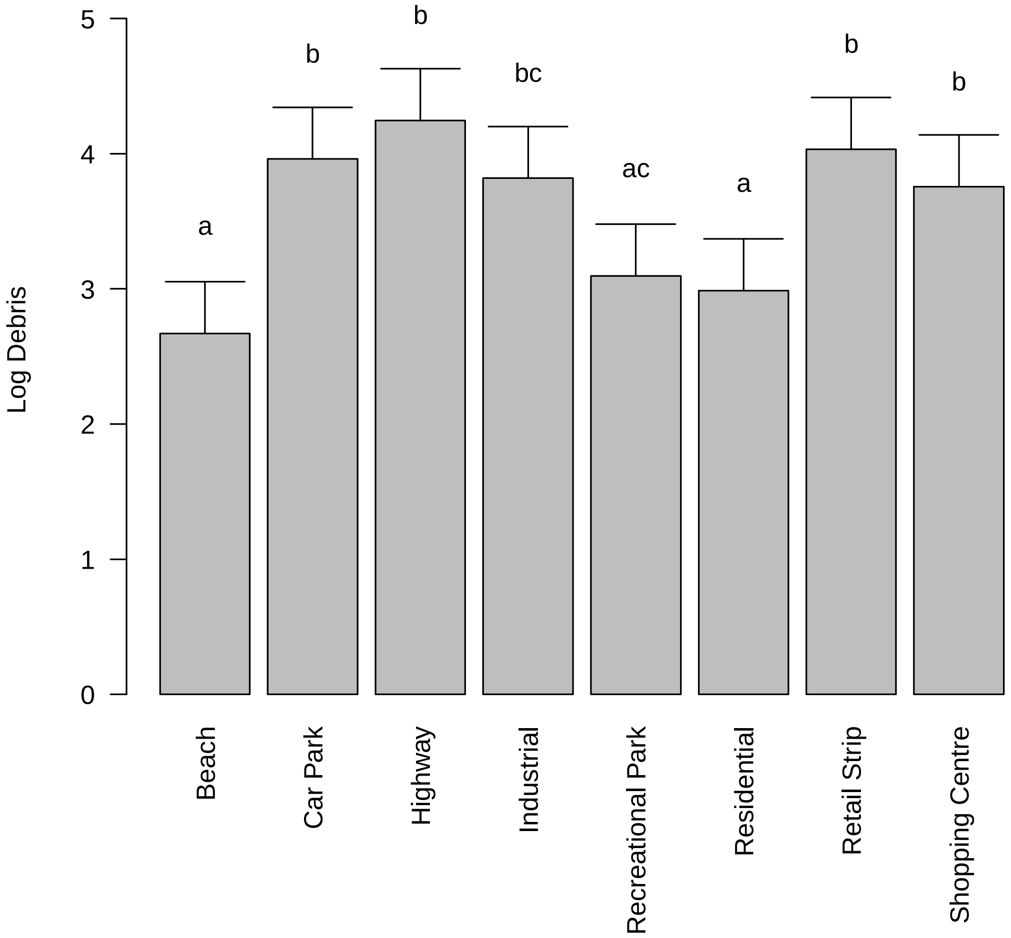

Figure 20. Comparison of the amount of debris present at each of the KAB site types. This model incorporates the full Australia-wide KAB dataset. Shared letters above the bars indicate non-significant differences, otherwise differences are significant.

In addition to these predictive variables, the GAM analysis also identified a very strong spatial pattern in the data and a strong individual site effect. The spatial pattern is in addition to the site characteristics discussed above, and suggests that there are regions where debris loads are higher or lower, even after the site characteristics have been accounted for. These could be due to a wide range of reasons, including variation in the behaviour of people in particular areas, availability or cost of waste disposal facilities at a local or regional scale, or other factors that could drive debris loads in a manner that is spatially correlated. Interestingly, individual sites tended to also be biased high or low. This could be due to idiosyncrasies of the sites themselves. For instance, if existing litter at a site attracts people to litter there themselves due to an assumption of social acceptability, that would be expected to lead to a strong (but potentially unpredictable prior to the even) site effect. It could also be that this site effect represents additional characteristics of a site that were not included in our analysis. For instance, if the speed limit changes near a site that could prompt drivers to litter from their vehicle before accelerating. However, as we did not consider speed limit, this effect would appear as a high debris effect at some sites but not others in the site effect.

Table 7. Parameter estimates for a statistical model of total debris for the KAB data (Australia-wide). Debris values are log(x+1) transformed. Note that the terms for State: ACT, SiteType: Beach, and Landcov: Conservation and Natural Environments are incorporated into the intercept in the model. The median is the median value of the relevant covariate, multiplying it times the coefficient gives a measure of the effect size of each term. Factors can be taken to have a value of 1 using treatment contrasts, as in this case. Smooth terms in the model are constrained to have mean values of zero, and thus are best interpreted as deviations around the parametric components. The effective and reference degrees of freedom for the smooth terms measure the deviation from linearity in these terms, the F and p value terms measure statistical significance.

Parametric terms							
	Coefficient Estimate	Standard Error of the Estimate	t value	Pr(>|t|)	Med. value	Med. effect	Absolute effect rank
(Intercept)	2.523	0.173	14.547	0.000	NA	2.523	4
RailDistKM	-0.024	0.001	-23.473	0.000	2.97	-0.07128	24
All_roads_50	-0.00004	0.00001	-4.758	0.000	5044.90	-0.201796	19
Use: Grazing	0.527	0.070	7.546	0.000	NA	0.527	12
Use: Dryland agriculture	0.324	0.041	7.911	0.000	NA	0.324	15
Use: Irrigated agriculture	0.020	0.098	0.206	0.837	NA	0.02	27
Use: Intensive uses	0.320	0.027	11.783	0.000	NA	0.32	17
Use: Water	0.720	0.056	12.823	0.000	NA	0.72	11
Pop_25km	1 E -07	1 E -08	5.528	0.000	550439	0.0550439	26
State: NSW	-0.196	0.042	-4.647	0.000	NA	-0.196	20
State: NT	1.316	0.072	18.258	0.000	NA	1.316	6
State: QLD	-0.141	0.066	-2.151	0.031	NA	-0.141	21
State: SA	-0.400	0.032	-12.445	0.000	NA	-0.4	14

State: TAS	-0.077	0.082	-0.932	0.351	NA	-0.077	23
State: VIC	-0.323	0.042	-7.687	0.000	NA	-0.323	16
State: WA	0.086	0.099	0.869	0.385	NA	0.086	22
Eco_resour_50km	0.012	0.002	5.577	0.000	1035	12.42	2
Eco_advan_50km	-0.016	0.001	-15.180	0.000	1008	-16.128	1
Edu_occupa_50km	-0.0005	0.0002	-2.653	0.008	987	-0.4935	13
Eco_disadv_50km	0.004	0.001	2.999	0.003	1014	4.056	3
SiteType: Car Park	1.027	0.027	38.231	0.000	NA	1.027	8
SiteType: Highway	1.399	0.026	53.238	0.000	NA	1.399	5
SiteType: Industrial	0.870	0.029	30.014	0.000	NA	0.87	9
SiteType: Recreational Park	0.253	0.029	8.572	0.000	NA	0.253	18
SiteType: Residential	0.065	0.027	2.463	0.014	NA	0.065	25
SiteType: Retail Strip	1.112	0.029	38.076	0.000	NA	1.112	7
SiteType: Shopping Centre	0.830	0.030	27.689	0.000	NA	0.83	10
Pop5to50km_resids	0.000001	0.0000002	8.863	0.000	-114.30	-0.0001143	29
roads_5to50km_resids	-0.001	0.0003	-3.683	0.000	-1.442	0.001442	28
Smooth terms							
	Effective degrees of freedom	Reference degrees of freedom	F	p-value			
te(Lat,Long)	1.674	1.900	5.015	0.025			
s(Sitecode)	903.009	979.000	15.763	0.000			

5.2.2 CSIRO Survey Analysis

The GAM model for the CSIRO data explained 41.5% of the variability in the data. The CSIRO GAM model was significantly simpler than the KAB GAM, corresponding with the tree model results. The most important parametric term was the distance from a rail station (aside from the intercept term, which just scales the response), which was negatively related to the debris density at a site. This variable is likely a proxy for a number of other correlated things, such as local population density, coastal access, and other factors. Population density in the region was the next most important variable, with debris loads increasing with population size.

The political unit was also important, with debris loads along the coast of various states differing significantly from New South Wales, which served as the reference level and was incorporated into the intercept. The patterns for the states are similar to what we found in previous analysis of the CSIRO coastal dataset, with relatively high loads in Western Australia due to prevailing onshore transport from wind and wave action, along with potential transport from currents in the Indian Ocean. On the east coast of the mainland, the debris loads increase from Queensland south to New South Wales, and further increase on the Victorian coastline. This is likely due to transport of materials along the coast southward in the East Australian Current. Thus debris from Brisbane appears to be exported southward, and transported onshore by wind and waves. This plume is steadily joined by additional debris from sources down the populated eastern coast, with deposition along the way, leading to the highest levels on the Victorian coastline (after correction for local population size). Tasmania, South Australia, and the Northern Territory have relatively low debris loads, compared with the other states. This could be due in part to the relative inaccessibility of coastal regions in these states.

Three socio-economic variables were included in the final model, although we were only able to include them as smooth terms (Figures 21 – 23). Overall, the relationships with these variables are parallel to those found in the analysis of the KAB data. Debris loads increase with economic disadvantage, decrease with economic advantage, and decrease less strongly with education and employment levels. Smooth functions in the model are centred on 0, and describe deviations from the mean level predicted by the model as the term in the smooth component varies. Thus, the importance of smooth components can be read directly from the graphs of its magnitude against the covariate value. For the CSIRO data, the level of economic advantage again has the largest effect, but the second most important is the level of disadvantage, followed by the term for education and employment opportunity. Although the advantage and disadvantage terms appear to represent the same drivers, they may in fact reflect differing ones. For instance, in areas of high advantage councils may be under significant pressure to maintain amenities, for instance via street sweeping. By contrast, in areas of disadvantage there may be differences in the availability or cost (maybe relative to income) of waste disposal facilities. Thus, both terms could be important simultaneously, and may represent different underlying drivers of the debris load in the coastal environment.

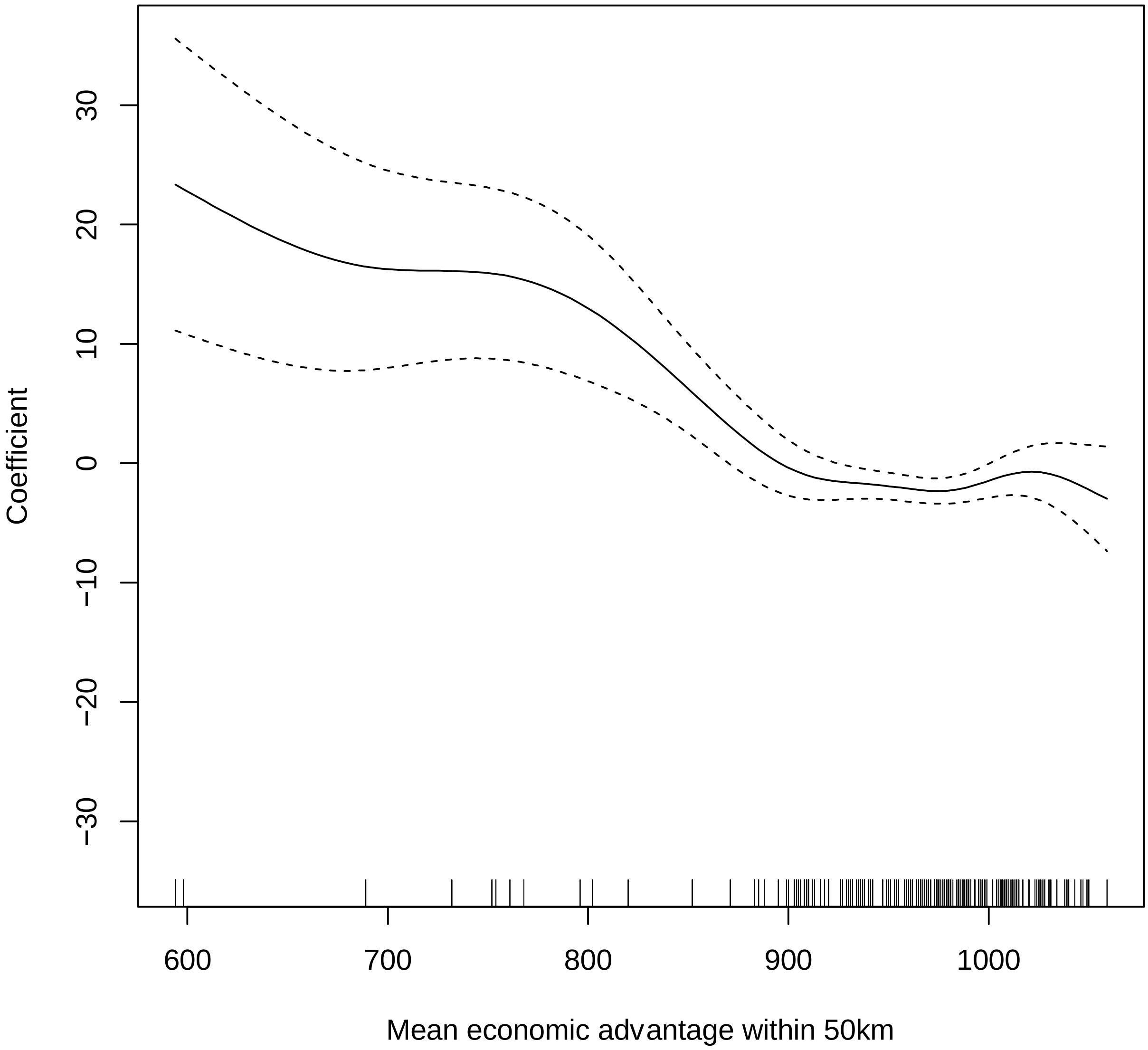

Figure 21. Smooth component for the index of relative economic advantage in the CSIRO Australia-wide GAM model. As socioeconomic status increases, total debris loads decrease. Lines along the x axis in the plot show where observations occur.

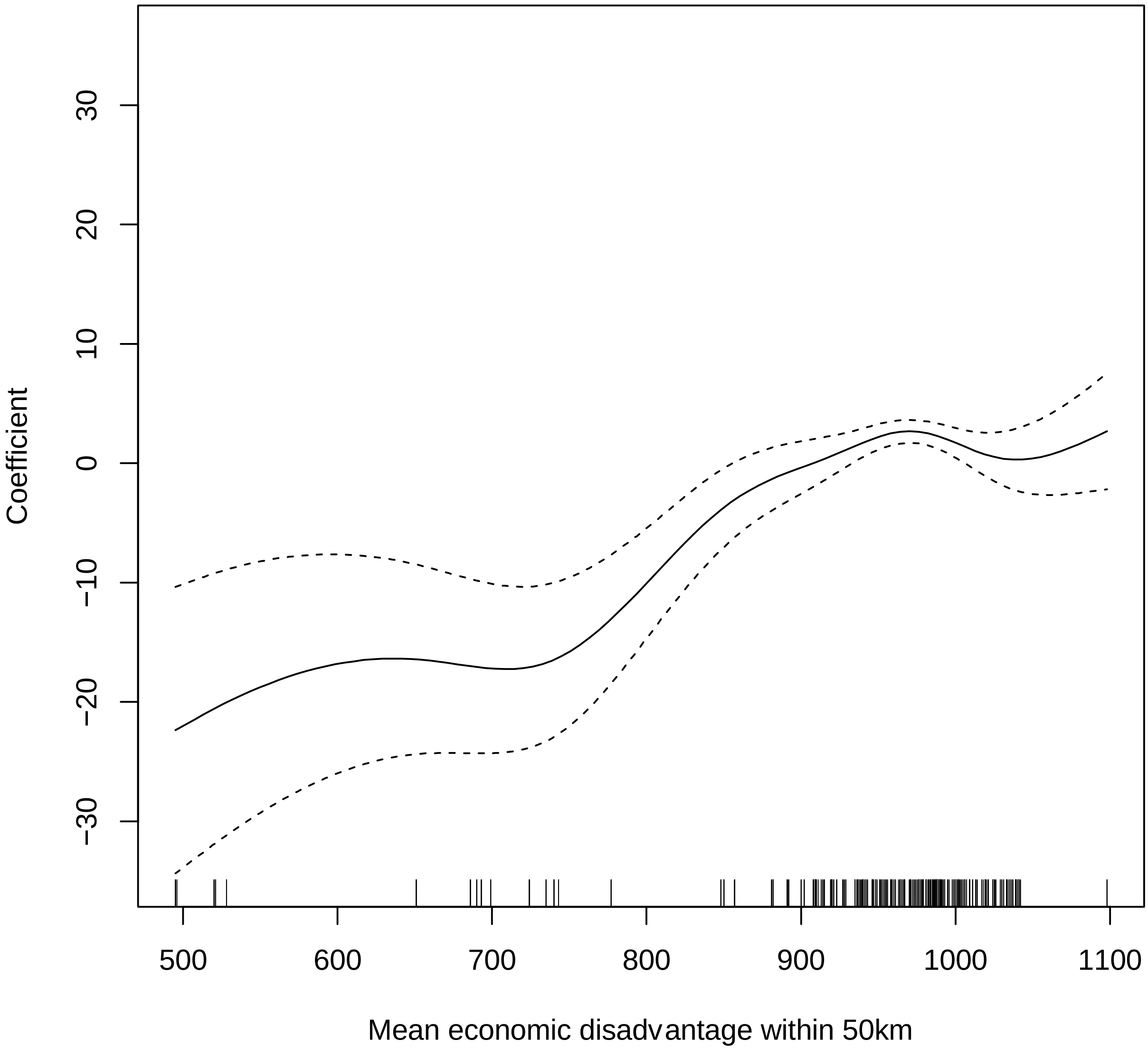

Figure 22. Smooth component for the index of relative economic advantage in the CSIRO Australia-wide GAM model. As the level of socioeconomic disadvantage increases, total debris loads increase correspondingly. Lines along the x axis in the plot show where observations occur.

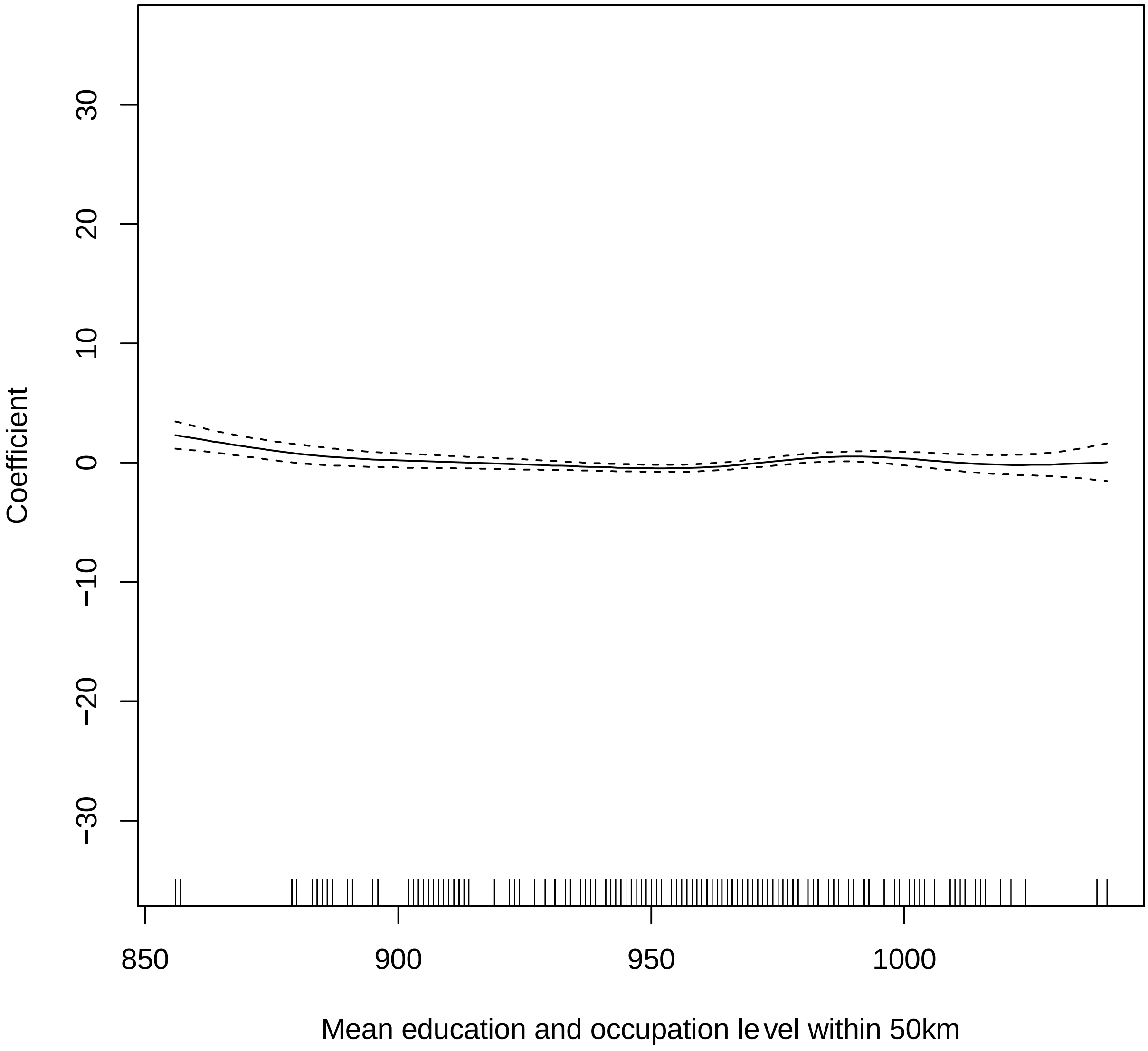

Figure 23. Smooth component for the index of relative education and occupation in the CSIRO Australia-wide GAM model. Debris levels do not vary much with education/occupation levels. Lines along the x axis in the plot show where observations occur.

Table 8. Parameter estimates for a statistical model of total debris for the CSIRO data (Australia-wide). Debris values are log(x+1) transformed. Note that the terms for State: NSW is incorporated into the intercept in the model. The median is the median value of the relevant covariate, multiplying it times the coefficient gives a measure of the effect size of each term. Factors can be taken to have a value of 1 using treatment contrasts, as in this case. Smooth terms in the model are constrained to have mean values of zero, and thus are best interpreted as deviations around the parametric components. The effective and reference degrees of freedom for the smooth terms measure the deviation from linearity in these terms, the F and p value terms measure statistical significance.

Parametric terms	Coefficient Estimate	Standard Error of the Estimate	t value	p-value	Med. value	Med. effect	Absolute effect rank
(Intercept)	3.823	0.329	11.635	0.000	NA	3.823	1
RailDistKM	-0.004	0.001	-4.203	0.000	40.952	-0.151	2
Pop_50km	0.000001	0.0000003	3.555	0.000	14825.000	0.015	3
StateNT	-1.890	0.956	-1.978	0.048	NA	-1.890	2
StateQLD	-0.903	0.371	-2.432	0.015	NA	-0.903	3
StateSA	-0.412	0.350	-1.177	0.240	NA	-0.412	4
StateTAS	-1.628	0.396	-4.112	0.000	NA	-1.628	3
StateVIC	0.279	0.407	0.686	0.493	NA	0.279	6
StateWA	0.367	0.395	0.930	0.353	NA	0.367	6
Smooth terms	Effective degrees of freedom	Reference degrees of freedom	F	p-value			
Eco_advan_50km	7.225	7.793	5.181	0.000			
Edu_occupa_50km	6.423	7.561	4.106	0.000			
Eco_disadv_50km	7.463	7.871	6.463	0.000			

5.3 Identification of debris hotspots (Catalogue of coastal debris hotspots for cost-effective targeting waste reduction)

There are a couple of ways to identify coastal debris hotspots. One method is simply to plot the raw data, the amount of debris (per unit area) for every study location (Figure 24). Because of the large variability in total debris, we represent this as the log of the total debris count. Figure 24 includes all of the study types; CUA, CSIRO, KAB, Emu, and Transect. One challenge is with standardisation of the data sets. Because all three have different methodologies, the values can vary dramatically. Additionally, because many of the surveys are conducted and reported by volunteers, quality control of the data can be complicated.

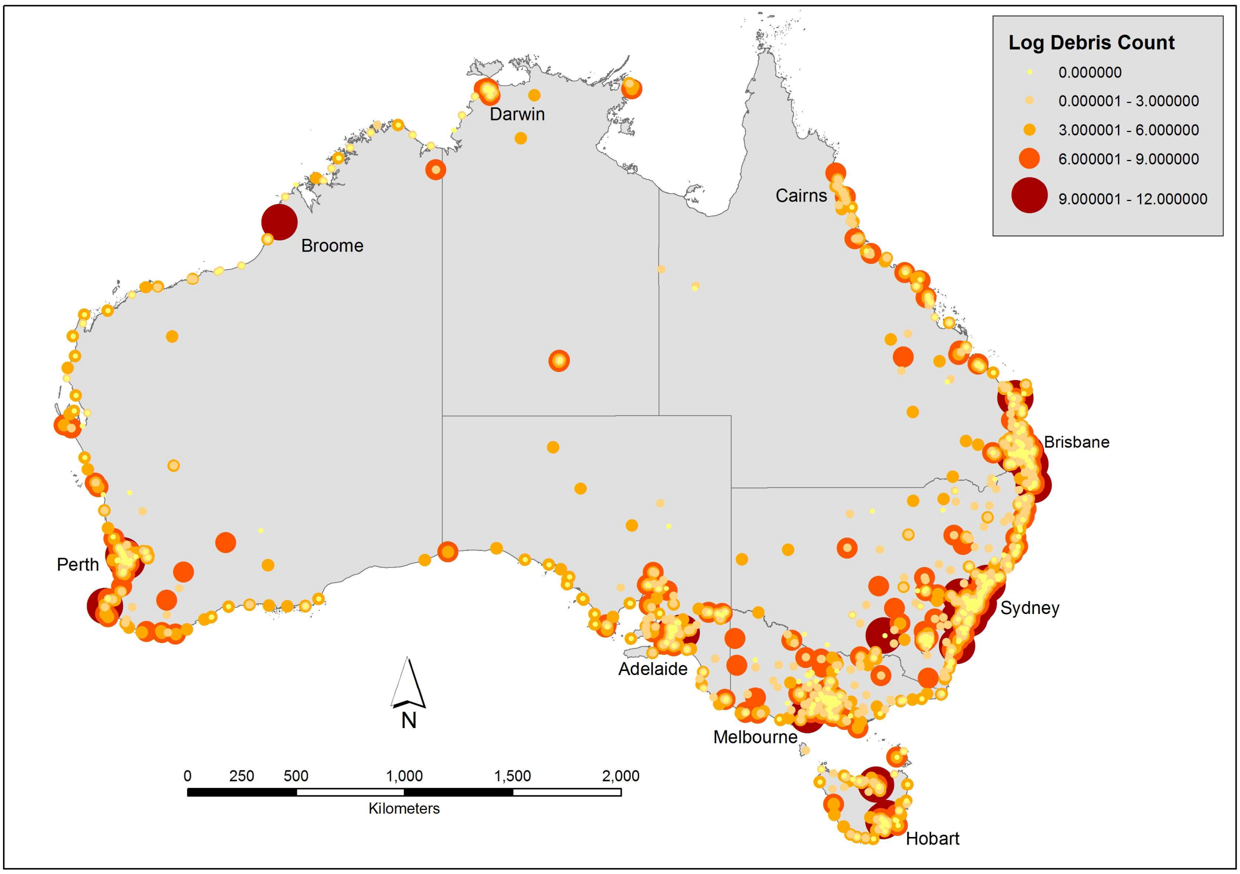

Figure 24. Map of debris hotspots based on all survey data from CSIRO, CUA and KAB. Note higher debris loads in urban cities around Australia's coastline. Data are reported as the log base 10 of the total amount of debris per 1000m^2.

The raw data are helpful because they can show where in the country absolute debris levels are highest. However, there are other ways of representing debris hotspots that can tell us more about geographic variability.

Because the KAB GAM model incorporates a smooth component jointly over latitude and longitude, we can determine the amount of variability in the data that is due to geography, after the variability related to the other covariates has been accounted for. We found significantly higher levels of debris in the northern and western survey sites (Figure 25).

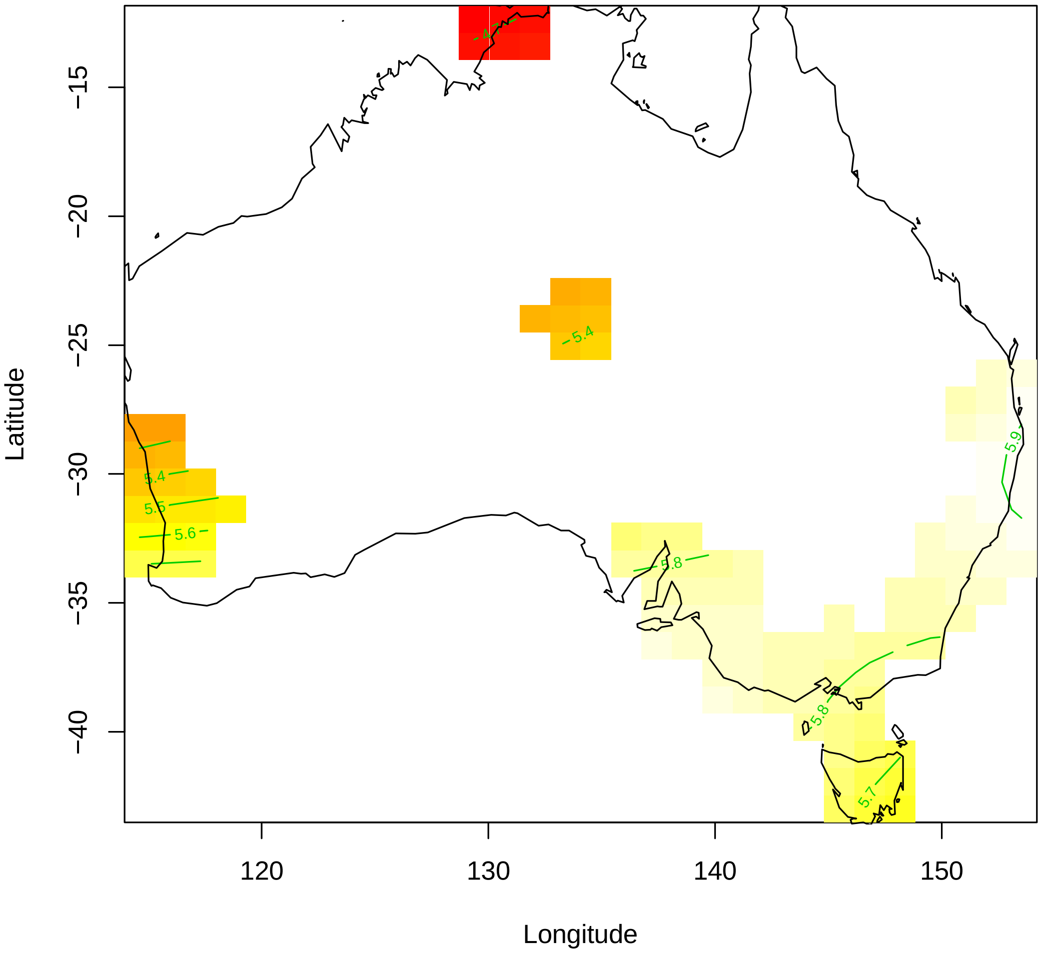

Figure 25. Coefficient values for the spatial surface included in the model of KAB data. The response variable is debris collected per 1000m^2. The full model is given in Table 2. The plot shows the values of the coefficient from the spatial surface. These are constrained to have a mean value of 0, thus they can be interpreted as spatial deviations from the expected values given the other model components

It is important to note however, that the variation in the spatial component of the GAM model is the left over variation that is not accounted for by any of the other model components. Thus, low socio-economic areas that have retail strips and highways, in densely populated areas would have particularly high debris levels and thus would be expected to be hotspots. However, even considering these relationships there is something about the regions coloured red or orange in Figure 25 that are associated with unexplained elevated debris levels.

One complementary approach that could be explored in future would be to extract the relevant covariates for the KAB GAM model on a grid at the national scale and use the fitted values in the model to predict where hotspots are likely to occur. We have not done this to date as it is a sizeable undertaking. However, if there is an emphasis to make a national estimate of the debris load in the Australian environment, this model could be used for that purpose. It could also be used to guide additional data collection in areas of suspected high and low load, or in areas where predictions are most uncertain.

6 How this project contributes to research into the understanding of marine litter

Marine debris is defined as any persistent manufactured or processed solid material discarded, disposed of, lost or abandoned into the marine and coastal environment (Coe and Rogers, 1997; Galgani et al., 2010). It includes those items made or lost by people, and those deliberately discarded into or unintentionally lost in the marine environment including, items made of plastic, wood, metal, glass, rubber, clothing and paper (Galgani et al., 2010; OSPAR, 2007). Recent work reports that approximately two-thirds of the debris found on Australia's coastline is comprised of plastic (Hardesty et al., in press), and globally plastic typically represents 60-80% of debris observed on beaches (Derraik 2002). Numerically speaking, plastic items are consistently among the most abundant types of marine debris in Europe and throughout the rest of the world (OSPAR, 2007; Thompson et al., 2009; UNEP, 2005, 2009).

The state of knowledge around marine debris and coastal litter has been growing dramatically in the last several years. Marine debris is derived from myriad sources including maritime shipping and transport, fishing related activities, stormwater rains, rain water runoff, littering, is known to pose a significant threat to biodiversity, due to its persistence, abundance and distribution in the marine environment. A recent review of marine debris impacts on wildlife reported that at least 17% of the nearly 700 species reported to interact with marine debris are on the IUCN Red List of threatened or near threatened species (Gall and Thompson 2015). Given the ubiquity of land-based litter or debris that enters the marine environment, debris can affect not only individuals, but also populations, assemblages, species and ecosystems.

The interest in marine debris has been growing within Australia (and worldwide). In August 2003, marine debris was identified as a key threatening process under the Australian *Environment Protection and Biodiversity Conservation Act 1999* (*EPBC Act*) (DEWHA, 2009). Furthermore, the federal government identified marine debris as a 'key threatening process' and has enacted a federal Threat Abatement Plan (TAP) due to the harm posed by marine debris to threatened vertebrate fauna in Australia and Australian waters. In 2015 this TAP came under review, and the threat abatement plan is currently under revision by the government.

In addition to biodiversity impacts, marine debris has economic consequences, particularly for coastal communities. Higher quantities of debris or litter are associated economic losses. For example, a 2013 study in Orange County, California interviewed residents asking how marine debris influences decisions about going to the beach and applied the travel cost model to estimate the value people derive from recreating at a particular site based on expectations compared to alternative sites. The authors reported that avoidance of littered beaches reportedly costs local residents roughly $32M annually, due to tourism losses in the summer months alone (Leggett et al. 2014). A study in South Korea considered the impact of a single extreme rainfall event which resulted in a particularly high level of litter to arrive on beaches on Goeje Island, estimating revenue losses from reduced tourism in that year were between 29-37M USD (Jang et al. 2014).

Furthermore, in the Asia-Pacific region, marine litter is estimated to cost around 1 billion euros per year to marine industries alone (McIlgorm et al. 2011).

This project focused not on the impacts of litter or debris on marine fauna, but rather on the quantification, distribution and density of litter from land-based sources. This work provides a significant step forward in our understanding of the movement, transport and drivers that result in litter hotspots in terrestrial and coastal environments. This is the first study of its kind to link socioeconomic factors to land-based sources of debris. Understanding the correlates or drivers for losses into the environment is a fundamental first step towards identifying potential solutions and targeting those solutions most effectively.

7 Information on ongoing waste diversion opportunities/outcomes as a result of the project.

7.1 Targeting of interventions to address debris hotspots

There are predictable contexts where debris loads are high. We found that socio-economic conditions are the strongest driver, followed by the typical activities in the location, followed by the land use context in which those activities occur. Specifically, debris loads are high in areas of socio-economic disadvantage, where activities are transitory (such as parking or shopping), and the area is surrounded by a natural or semi-natural environment (such as a creek or a grassy area). By taking account of these variables in a given location, we can predict the chance of high debris loads at any scale from meters up to whole suburbs. While these predictions do not guarantee high loads at these sites, they provide an estimate of the risk of inappropriate behaviours by people in these contexts.

The predictability of these high debris loads can be useful in targeting interventions, whether they are based on outreach and community engagement or increased surveillance and enforcement. A wide range of interventions related to enforcement could be increased in high load locations, ranging from signage and increased penalties, which alter the perception of consequences for littering and illegal dumping, to remote or direct surveillance, which increase the chance, and the perception, that violators will be caught. In a less putative approach, outreach to increase awareness, provide nudges, and encourage responsible behaviour can also be targeted to these high risk locations. This will be an area of focus for further research which specifically focuses on policy effectiveness.

7.2 The role of social context in the design of interventions

In addition to clear predictors for high debris load sites, we also found clear predictors for low load sites. While some, such as high levels of economic advantage, are just mirrors of predictors of high load sites, others are qualitatively different and can provide some clues as to interventions that could reduce debris levels. Three site types had particularly low debris levels, beaches, parks, and residential areas. In comparison with the transitory uses that are associated with high load sites, these three site types share two characteristics. First, users of these sites are likely present for much longer periods. Residential locations are a clear example, most people in these locations are present over long periods, and have a close association with the location. Second, these three site categories all have recognized positive aesthetic or cultural values associated with them. Beaches and parks are very likely to have positive associations in people’s minds, generating cultural value for them, which is affected by qualities such as perceived cleanliness. Interestingly, this pattern appears to be true for the larger land use context also. When the local land use context around the site was an official nature reserve it had lower debris levels than any other type of site, included natural, but unrecognized, habitats such as open grasslands and waterways.

This pattern suggests that changing the social context of high litter sites could be an effective intervention. Increasing the time people spend in what are otherwise transitory types of sites could change the social connection with them, possibly increasing stewardship of the sites. In addition, providing official recognition of the sites as locations that are typically associated with positive cultural value might also change user's behaviour at the site. For example, establishment of nature reserves, parks, and recreation facilities interspersed with parking or shopping areas may change user's littering or dumping behaviour in those contexts.

7.3 Intercepting debris effectively

While deposition (i.e. littering or illegal dumping) appears to be the dominant process driving the distribution of debris loads, based on our analysis debris is clearly transported among sites. Transport may even be higher than we were able to estimate in this analysis, given the absence of direct observations on transport. Transport via surface runoff and storm drains appears to be more important for debris on land, although wind transport does play some role. Debris loads at sites scale predictably with upstream load and watershed size. Given this, installation of debris traps and other infrastructure to reduce losses into the environment at targeted sites can make a significant contribution to reducing losses into the environment.

Our watershed scale analysis for the Brisbane region provides estimated load due to debris from other sites washing or blowing into the site. For water in particular, which flows down through watersheds into a decreasing number of drainages, we can estimate the relationship between the debris load that would be intercepted and the location and number of traps distributed throughout the Brisbane region. Thus we can evaluate the efficacy of the current system, and assess the additional benefit of installing additional traps or other infrastructure given their location. While we have not undertaken this analysis specifically in our current project, it would be straightforward to implement the analysis given information on the locations of traps at the current time.

7.4 Assessing feasibility, cost, and benefits of infrastructure investments

Installations should consider predicted loads in evaluating size and design of traps, and in establishing recurrence of maintenance and debris removal from traps. Our results can provide an approximate estimate of load, which can be used to evaluate both expected amount removed if infrastructure is installed at a given site, the likely rate of accumulation at the site, and the likely maintenance schedule for the site. Maintenance of debris traps has been a major failure point in a recent review of municipal solid waste inputs to marine and freshwater systems, driven by unanticipated loads which led to overburdened and ineffectual debris traps. Optimizing the efficiency of the infrastructure and operational investment in the context of a fixed budget for an urban area, such as the greater Brisbane area, again has not been done in the current project. However, it would be a relatively straightforward application of the project outputs.

8 Investment opportunities and ideas for future projects

8.1 Improving analysis of debris hotspots and drivers of hotspots

Our statistical modelling was able to clearly identify predictors for locations that are likely to have high and low levels of debris. This analysis is useful, both for targeting interventions and understanding drivers. However, we found that individual sites had loads that were sometimes consistently higher or lower than expected based on our modelling. This suggests that there are additional factors, related to those particular sites, which are affecting deposition at those sites. It would be useful to visit a number of these sites to evaluate possible variables that we could include in the models to improve their predictions and our understanding.

A second extension in this respect would be to disaggregate the debris loads from different sources. Debris in the environment is a result of a mix of littering of single items, groups of items, illegal dumping of household waste, and illegal dumping on a commercial scale. It is potentially possible to build separate models of loads for particular groups of materials which would be representative of each of these types of activities. For instance, modelling short-term use consumer items, such as disposable cups, could be indicative of littering, while building rubble might be indicative of commercial dumping. Understanding of the drivers for each of these types of materials could support more effective interdiction.

8.2 Extending the baseline data more widely to allow prediction of debris loads at the national scale

While we were able to make use of national scale data in this project, the data on debris loads on land were very concentrated around major cities. The KAB data, which provides the highest quality data on debris loads on land was even more concentrated, with very little sampling on urban margins, in smaller towns, and outside developed areas. It would be possible to identify the under-sampled contexts, and use a one-time sampling program to improve sampling effort in these areas. If these samples can be paired with sites that will be cleaned up by CUA volunteers, they could also be used to establish a relationship between debris at a site as reported to CUA and the absolute density found in a fixed area survey, such as the methods used by KAB. If it was possible to develop this relationship, it would allow a joint analysis of the KAB and CUA data which would substantially increase the data available for estimating loads and understanding the drivers for its variation over time and among sites.

8.3 Integrating data on transport with the existing landscape scale analysis

A number of the larger councils in Australia have major debris interception programs, ranging from daily removal of large debris in Sydney Harbor to litter traps on the lower Yarra River in Melbourne. Connecting our watershed scale load and transport modelling with data from these debris removal programs in waterways could provide an excellent additional data source, allowing significant refinement of our models. Furthermore, adding some detailed studies of transport rates, for instance by marking individual items could also improve the transport models significantly. Implicit in this would be extending our transport modelling from Brisbane to other major urban areas. Improvements to the transport modelling would be particularly useful, if it is applied to evaluate existing or proposed infrastructure investments, schedule maintenance, or optimize designs as the increased detail and accuracy would be useful.

9 Conclusions

The project successfully achieved its aims of improving our understanding of debris across Australia, particularly in coastal environments. We evaluated whether quantities of debris varied with different site types (e.g. beaches, highways, recreational park, etc.), finding that areas where people are more transient tended to have higher debris counts (carparks, retail strips and shopping centres). We identified a number of variables that affect debris loads, including socio-economic factors, population density, and accessibility. Understanding the factors that are linked to higher debris loads enables us to target limited resources more effectively. For example, identifying that human deposition was by far the most important in determining the load at a site means that programs that directly target behavioural change are likely to be more effective than reducing debris inputs. Similarly, litter traps over watercourses are more likely to reduce debris inputs to the coastal and marine environment than litter traps aimed at targeting wind movement.

As a result of the analyses, we can suggest opportunities or suggestions to reduce debris inputs to the environment. We suggest that addressing sites with high littering rates should be the first priority. The strongest effect on loads is due to direct deposition (i.e. littering behaviour), which contributed more than transport in determining the load (or debris count) at a site. Second, targeting particular population segments will likely result in better outcomes, given that the three strongest predictors of debris at a site had to do economic wealth and social disadvantage in the population near the site.

There are a number of opportunities for improvement, in both data and analysis concerning littering, transport, and debris loss to the marine environment. Transport modeling completed for this project is a useful first step in understanding the movement of debris, and identifying critical intervention points. Increasing the detail in this analysis would improve the capacity to identify and evaluate intervention points. For instance, it might be possible to estimate whether funds would be better spent installing and maintaining a litter trap in a storm drain, or clearing the existing debris load in the creek receiving the storm water. It is likely that limited funds and staff time force decisions on local governments responsible for waste management, thus there is a significant opportunity in providing information on the likely success or impact of such allocation decisions. Also, the load analysis in this project could be used in conjunction with interventions such as litter signage, bin installation, or community outreach programs to evaluate the impact of these programs on load. The current statistical model can be used to predict the expected load in these contexts. Thus, sites that have had or will have interventions can be compared to assess the effect of the intervention on the debris load.

10 References

Australian Bureau of Statistics (2006). 2033.0.55.001 - Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA), Australia - Data only.

Australian Bureau of Statistics (2011a). 1216.0 - Australian Standard Geographical Classification (ASGC), July 2011

Australian Bureau of Statistics (2011b). 2033.0.55.001 - Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA), Australia.

Australian Bureau of Statistics (2016). Estimated Residential Population by Statistical Area 2 1991 to 2015. Australia. http://stat.abs.gov.au/Index.aspx?DataSetCode=ABS_ANNUAL_ERP_ASGS.

Coe, J. M., Andersson, S., & Rogers, D. B. (1997). *Marine Debris* (pp. 25-33). Springer New York.

Department of Agriculture, Fisheries and Forestry (2015). Catchment Scale Land Use of Australia – Update March 2015, Canberra, Australia.

Derraik, J.G.B., (2002). The pollution of the marine environment by plastic debris: a review. Marine Pollution Bulletin. 44, 842–852.

Department of the Environment, W., Heritage and the Arts. (2009). Threat abatement plan for the impacts of marine debris on vertebrate marine life. Department of the Environment, Water, Heritage and the Arts, Canberra.

Gall, S. C., & Thompson, R. C. (2015). The impact of debris on marine life. Marine pollution bulletin, *92*(1), 170-179.

Gallant, J.C., Dowling, T.I., Read, A.M., Wilson, N., Tickle, P., Inskeep, C. (2011) 1 second SRTM Derived Digital Elevation Models User Guide. Geoscience Australia www.ga.gov.au/topographic-mapping/digital-elevation-data.html.

Galgani, F., Fleet, D., Van Franeker, J., Katsanevakis, S., Maes, T., Mouat, J., ... & Amato, E. (2010). Marine Strategy Framework Directive: Task Group 10 Report Marine Litter. Luxembourg: Office for Official Publications of the European Communities.

Geoscience Australia (2006). GEODATA TOPO 250K Series 3 Topographic Data, Canberra, Australia, ANZCW0703005458.

Hardesty, B, D. Lawson, T. van der Velde, T. Lansdell, M. and C. Wilcox (2016, in press). Estimating quantities and sources of marine debris at a continental scale.

Hardesty, B, D. Wilcox, C. Lawson, T. Lansdell, M. and T. van der Velde (2014). Understanding the effects of marine debris on wildlife. Final report to Earthwatch Australia.

Hothorn T., Zeileis, A. (2015). partykit: A Modular Toolkit for Recursive Partytioning in R. Journal of Machine Learning Research 16, 3905–3909. http://jmlr.org/papers/v16/hothorn15a.html.

Jang, Y. C., Hong, S., Lee, J., Lee, M. J., & Shim, W. J. (2014). Estimation of lost tourism revenue in Geoje Island from the 2011 marine debris pollution event in South Korea. Marine Pollution Bulletin, 81, 49–54.

Leggett, C., Scherer, N., Curry, M. & Bailey, R. (2014). Assessing the economic benefits of reductions in marine debris: A pilot study of beach recreation in Orange County, California. Final report: June 15, 2014, from National Oceanic and Atmospheric Administration. Cambridge, USA. pp. 44. http://marinedebris.noaa.gov/sites/default/files/MarineDebrisEconomicStudy.pdf.

McIlgorm, A., Campbell, H. F., & Rule, M. J. (2011). The economic cost and control of marine debris damage in the Asia-Pacific region. Ocean and Coastal Management*, 54*, 643–651.

Mouat, J., Lozano, R. L. & Bateson, H. (2010). Economic impacts of marine litter. Kommunenes Internasjonale Miljøorganisasjon, Lerwick, Shetland, UK.

Ofiara, D. D. & Brown, B. (1999). Assessment of economic losses to recreational activities from 1988 marine pollution events and assessment of economic losses from long-term contamination of fish within the New York Bight to New Jersey. Marine Pollution Bulletin 38, 990-1004.

NOAA website http://oceanservice.noaa.gov/facts/marinedebris.html; downloaded 01/08/2016.

OSPAR, 2007. OSPAR Pilot Project on Monitoring Marine Beach Litter: Monitoring of marine litter on beaches in the OSPAR region. OSPAR Commission, London.

Thompson, R. C., Swan, S. H., Moore, C. J., & Vom Saal, F. S. (2009). Our plastic age. Philosophical Transactions of the Royal Society B: Biological Sciences, 364(1526), 1973-1976.

UNEP (2005). Marine Litter an Analytical Overview. United Nations Environment Programme, Kenya, 47 pp.

UNEP (2009). Marine Litter: A Global Challenge. UNEP, Kenya. 232 pp.

11 Appendices

A. Australian Land Use and Management (ALUM) Classification version 7

CODE	TERTIARY	SECONDARY	PRIMARY
1.0.0	1.0.0 Conservation and natural environments	1.0 Conservation and natural environments	1 Conservation and natural environments
1.1.0	1.1.0 Nature conservation	1.1 Nature conservation	
1.1.1	1.1.1 Strict nature reserves		
1.1.2	1.1.2 Wilderness area		
1.1.3	1.1.3 National park		
1.1.4	1.1.4 Natural feature protection		
1.1.5	1.1.5 Habitat/species management area		
1.1.6	1.1.6 Protected landscape		
1.1.7	1.1.7 Other conserved area		
1.2.0	1.2.0 Managed resource protection	1.2 Managed resource protection	
1.2.2	1.2.2 Surface water supply		
1.2.3	1.2.3 Groundwater		
1.2.4	1.2.4 Landscape		
1.2.5	1.2.5 Traditional indigenous uses		
1.3.0	1.3.0 Other minimal use	1.3 Other minimal use	
1.3.1	1.3.1 Defence land - natural areas		
1.3.2	1.3.2 Stock route		
1.3.3	1.3.3 Residual native cover		
1.3.4	1.3.4 Rehabilitation		
2.1.0	2.1.0 Grazing natural vegetation	2.1 Grazing natural vegetation	2 Production from relatively natural environments
2.2.0	2.2.0 Production forestry	2.2 Production forestry	
2.2.1	2.2.1 Wood production		
2.2.2	2.2.2 Other forest production		
3.0.0	3.0.0 Production from dryland agriculture and plantations	3.0 Production from dryland agriculture and plantations	3 Production from dryland agriculture and plantations
3.1.0	3.1.0 Plantation forestry	3.1 Plantation forestry	3 Production from dryland agriculture and plantations
3.1.1	3.1.1 Hardwood plantation		
3.1.2	3.1.2 Softwood plantation		
3.1.3	3.1.3 Other forest plantation		
3.1.4	3.1.4 Environmental forest plantation		
3.2.0	3.2.0 Grazing modified pastures	3.2 Grazing modified pastures	
3.2.1	3.2.1 Native/exotic pasture mosaic		
3.2.2	3.2.2 Woody fodder plants		
3.2.3	3.2.3 Pasture legumes		
3.2.4	3.2.4 Pasture legume/grass mixtures		
3.2.5	3.2.5 Sown grasses		
3.3.0	3.3.0 Cropping	3.3 Cropping	
3.3.1	3.3.1 Cereals		
3.3.2	3.3.2 Beverage & spice crops		
3.3.3	3.3.3 Hay & silage		
3.3.4	3.3.4 Oil seeds		
3.3.5	3.3.5 Sugar		
3.3.6	3.3.6 Cotton		
3.3.7	3.3.7 Alkaloid poppies		
3.3.8	3.3.8 Pulses		
3.4.0	3.4.0 Perennial horticulture	3.4 Perennial horticulture	
3.4.1	3.4.1 Tree fruits		
3.4.2	3.4.2 Oleaginous fruits		
3.4.3	3.4.3 Tree nuts		
3.4.4	3.4.4 Vine fruits		
3.4.5	3.4.5 Shrub nuts, fruits & berries		
3.4.6	3.4.6 Perennial flowers & bulbs		

3.4.7	3.4.7 Perennial vegetables & herbs		
3.4.9	3.4.9 Grapes		
3.5.0	3.5.0 Seasonal horticulture	3.5 Seasonal horticulture	
3.5.3	3.5.3 Seasonal flowers & bulbs		
3.5.4	3.5.4 Seasonal vegetables & herbs		
3.6.0	3.6.0 Land in transition	3.6 Land in transition	
3.6.1	3.6.1 Degraded land		
3.6.2	3.6.2 Abandoned land		
3.6.3	3.6.3 Land under rehabilitation		
3.6.4	3.6.4 No defined use		
3.6.5	3.6.5 Abandoned perennial horticulture		
4.0.0	4.0.0 Production from irrigated agriculture and plantations	4.0 Production from irrigated agriculture and plantations	4 Production from irrigated agriculture and plantations
4.1.0	4.1.0 Irrigated plantation forestry	4.1 Irrigated plantation forestry	
4.1.1	4.1.1 Irrigated hardwood plantation		
4.1.2	4.1.2 Irrigated softwood plantation		
4.1.3	4.1.3 Irrigated other forest production		
4.1.4	4.1.4 Irrigated environmental forest plantation		
4.2.0	4.2.0 Grazing irrigated modified pastures	4.2 Grazing irrigated modified pastures	
4.2.1	4.2.1 Irrigated woody fodder plants		
4.2.2	4.2.2 Irrigated pasture legumes		
4.2.3	4.2.3 Irrigated legume/grass mixtures		
4.2.4	4.2.4 Irrigated sown grasses		
4.3.0	4.3.0 Irrigated cropping	4.3 Irrigated cropping	
4.3.1	4.3.1 Irrigated cereals		
4.3.2	4.3.2 Irrigated beverage & spice crops		
4.3.3	4.3.3 Irrigated hay & silage		
4.3.4	4.3.4 Irrigated oil seeds		
4.3.5	4.3.5 Irrigated sugar		
4.3.6	4.3.6 Irrigated cotton		
4.3.7	4.3.7 Irrigated alkaloid poppies		
4.3.8	4.3.8 Irrigated pulses		
4.4.0	4.4.0 Irrigated perennial horticulture	4.4 Irrigated perennial horticulture	
4.4.1	4.4.1 Irrigated tree fruits		
4.4.2	4.4.2 Irrigated oleaginous fruits		
4.4.3	4.4.3 Irrigated tree nuts		
4.4.4	4.4.4 Irrigated vine fruits		
4.4.5	4.4.5 Irrigated shrub nuts, fruits & berries		
4.4.6	4.4.6 Irrigated perennial flowers & bulbs		
4.4.7	4.4.7 Irrigated perennial vegetables & herbs		
4.4.8	4.4.8 Irrigated citrus		
4.4.9	4.4.9 Irrigated grapes		
4.5.0	4.5.0 Irrigated seasonal horticulture	4.5 Irrigated seasonal horticulture	
4.5.1	4.5.1 Irrigated seasonal fruits		
4.5.2	4.5.2 Irrigated seasonal nuts		
4.5.3	4.5.3 Irrigated seasonal flowers & bulbs		
4.5.4	4.5.4 Irrigated seasonal vegetables & herbs		
4.5.5	4.5.5 Irrigated turf farming		
4.6.0	4.6.0 Irrigated land in transition	4.6 Irrigated land in transition	
4.6.1	4.6.1 Degraded irrigated land		
4.6.2	4.6.2 Abandoned irrigated land		
4.6.3	4.6.3 Irrigated land under rehabilitation		
4.6.4	4.6.4 No defined use (irrigation)		
4.6.5	4.6.5 Abandoned irrigated perennial horticulture		
5.0.0	5.0.0 Intensive uses	5.0 Intensive uses	5 Intensive uses

5.1.0	5.1.0 Intensive horticulture	5.1 Intensive horticulture	
5.1.1	5.1.1 Shadehouses		
5.1.2	5.1.2 Glasshouses		
5.1.3	5.1.3 Glasshouses (hydroponic)		
5.1.4	5.1.4 Abandoned intensive horticulture		
5.2.0	5.2.0 Intensive animal husbandry	5.2 Intensive animal husbandry	
5.2.1	5.2.1 Dairy sheds and yards		
5.2.2	5.2.2 Cattle feedlots		
5.2.3	5.2.3 Sheep feedlots		
5.2.4	5.2.4 Poultry farms		
5.2.5	5.2.5 Piggeries		
5.2.6	5.2.6 Aquaculture		
5.2.7	5.2.7 Horse studs		
5.2.8	5.2.8 Stockyards/saleyards		
5.2.9	5.2.9 Abandoned intensive animal husbandry		
5.3.0	5.3.0 Manufacturing and industrial	5.3 Manufacturing and industrial	
5.3.1	5.3.1 General purpose factory		
5.3.2	5.3.2 Food processing factory		
5.3.3	5.3.3 Major industrial complex		
5.3.4	5.3.4 Bulk grain storage		
5.3.5	5.3.5 Abattoirs		
5.3.6	5.3.6 Oil refinery		
5.3.7	5.3.7 Sawmill		
5.3.8	5.3.8 Abandoned manufacturing and industrial		
5.4.0	5.4.0 Residential and farm infrastructure	5.4 Residential and farm infrastructure	
5.4.1	5.4.1 Urban residential		
5.4.2	5.4.2 Rural residential with agriculture		
5.4.3	5.4.3 Rural residential without agriculture		
5.4.4	5.4.4 Remote communities		
5.4.5	5.4.5 Farm buildings/infrastructure		
5.5.0	5.5.0 Services	5.5 Services	
5.5.1	5.5.1 Commercial services		
5.5.2	5.5.2 Public services		
5.5.3	5.5.3 Recreation and culture		
5.5.4	5.5.4 Defence facilities - urban		
5.5.5	5.5.5 Research facilities		
5.6.0	5.6.0 Utilities	5.6 Utilities	
5.6.1	5.6.1 Fuel powered electricity generation		
5.6.2	5.6.2 Hydro electricity generation		
5.6.3	5.6.3 Wind farm electricity generation		
5.6.4	5.6.4 Electricity substations and transmission		
5.6.5	5.6.5 Gas treatment, storage and transmission		
5.6.6	5.6.6 Water extraction and transmission		
5.7.0	5.7.0 Transport and communication	5.7 Transport and communication	
5.7.1	5.7.1 Airports/aerodromes		
5.7.2	5.7.2 Roads		
5.7.3	5.7.3 Railways		
5.7.4	5.7.4 Ports and water transport		
5.7.5	5.7.5 Navigation and communication		
5.8.0	5.8.0 Mining	5.8 Mining	
5.8.1	5.8.1 Mines		
5.8.2	5.8.2 Quarries		
5.8.3	5.8.3 Tailings		
5.8.4	5.8.4 Extractive industry not in use		
5.9.0	5.9.0 Waste treatment and disposal	5.9 Waste treatment and disposal	
5.9.1	5.9.1 Effluent pond		
5.9.2	5.9.2 Landfill		

5.9.3	5.9.3 Solid garbage		
5.9.4	5.9.4 Incinerators		
5.9.5	5.9.5 Sewage/sewerage		
6.0.0	6.0.0 Water	6.0 Water	6 Water
6.1.0	6.1.0 Lake	6.1 Lake	
6.1.1	6.1.1 Lake - conservation		
6.1.2	6.1.2 Lake - production		
6.1.3	6.1.3 Lake - intensive use		
6.2.0	6.2.0 Reservoir/dam	6.2 Reservoir/dam	
6.2.1	6.2.1 Reservoir		
6.2.2	6.2.2 Water storage - intensive use/farm dams		
6.2.3	6.2.3 Evaporation basin		
6.3.0	6.3.0 River	6.3 River	
6.3.1	6.3.1 River - conservation		
6.3.2	6.3.2 River - production		
6.3.3	6.3.3 River - intensive use		
6.4.0	6.4.0 Channel/aqueduct	6.4 Channel/aqueduct	
6.4.1	6.4.1 Supply channel/aqueduct		
6.4.2	6.4.2 Drainage channel/aqueduct		
6.4.3	6.4.3 Stormwater		
6.5.0	6.5.0 Marsh/wetland	6.5 Marsh/wetland	
6.5.1	6.5.1 Marsh/wetland - conservation		
6.5.2	6.5.2 Marsh/wetland - production		
6.5.4	6.5.4 Marsh/wetland - saline		
6.6.0	6.6.0 Estuary/coastal waters	6.6 Estuary/coastal waters	
6.6.1	6.6.1 Estuary/coastal waters - conservation		
6.6.2	6.6.2 Estuary/coastal waters - production		
6.6.3	6.6.3 Estuary/coastal waters - intensive use		

B. Covariate data field descriptions

Field name	**Description**
Global_ID	ID that carries across all data sources
Unique_ID	ID unique within each source dataset
Town	Name of the township that the survey site occurred within
Source	Name distinguishing the source of the dataset
RailNearID	The ID present in the GEODATA TOPO 250K Series 3 Topographic Data (Geoscience Australia, 2006) dataset of the nearest railway station to the survey site
RailDistKM	The distance in kilometers to the nearest railway station
RoadNearID	The ID present in the GEODATA TOPO 250K Series 3 Topographic Data (Geoscience Australia, 2006) dataset of the nearest road (regardless of type) to the survey site.
RoadDistKM	The distance in kilometers to the nearest road.
Landuse_code	The code that related to the Australian Land Use and Management (ALUM) Classification version 7
Area_m2	Area surveyed in meters squared (where available)
Date	The date the survey occurred
popXX	Total population at each segment distance from the survey site. XX will be either 1km, 5km, 10km, 25km or 50km.
Eco_advan_XX	Mean Index of Relative Socio-economic Advantage and Disadvantage that indicate either relative advantage or disadvantage. This index can be used to measure socio-economic wellbeing in a continuum, from the most disadvantaged areas to the most advantaged areas. XX will be either 1km, 5km, 10km, 25km or 50km.
eco_disadvXX	Mean Index of Relative Socio-economic Disadvantage that indicates relative disadvantage at each segment distance from the survey site. XX will be either 1km, 5km, 10km, 25km or 50km.
eco_resourXX	Mean Index of Economic Resources relate to the financial aspects of relative socio-economic advantage and disadvantage. These include indicators of high and low income, as well as variables that correlate with high or low wealth. XX will be either 1km, 5km, 10km, 25km or 50km.
Edu_occupaXX	Mean Index of Education and Occupation summarises variables relating exclusively to education, employment and occupation at each segment distance from the survey site. XX will be either 1km, 5km, 10km, 25km or 50km.
DualCarriageRdXX	Total length of Dual Carriageway type roads within each segment distance from the survey site. XX will be either 1km, 5km, 10km, 25km or 50km.
MinorRdXX	Total length of Minor type roads within each segment distance from the survey site. XX will be either 1km, 5km, 10km, 25km or 50km.
PrincialRdXX	Total length of Principal type roads within each segment distance from the survey site. XX will be either 1km, 5km, 10km, 25km or 50km.
SecondaryRdXX	Total length of Secondary type roads within each segment distance from the survey site. XX will be either 1km, 5km, 10km, 25km or 50km.
TrackXX	Total length of Track type roads within each segment distance from the survey site. XX will be either 1km, 5km, 10km, 25km or 50km.

CONTACT US

t 1300 363 400
+61 3 9545 2176
e enquiries@csiro.au
w www.csiro.au

FOR FURTHER INFORMATION

Oceans and Atmosphere/Marine Resources and Industries
Britta Denise Hardesty
t +61 03 6232 5276
e denise.hardesty@csiro.au
w www.csiro.au/flagship-businessunit