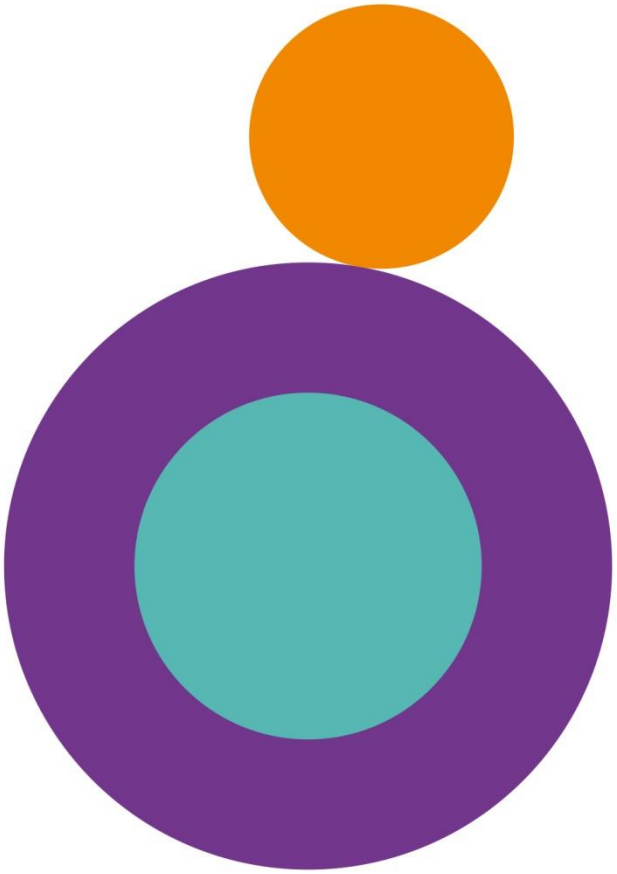
Opening the floodgates: Releasing IAG data to build safer communities



Disclaimer

Your access to and use of this information is governed by the Terms of Use. Your access to and use of the information constitutes your agreement to the Terms of Use. The Terms of Use are in Appendix B

Contents

Introduction

Who is IAG?

IAG is a general insurance company whose purpose is to make your world a safer place, whether you are a customer, partner, employee, shareholder or part of the communities IAG serves across Australia, New Zealand and Asia. Our businesses have helped people recover from natural disasters, accidents and loss since 1851.

We employ more than 15,000 people in our operations in Australia, New Zealand, Thailand and Vietnam, providing insurance under many leading brands, including NRMA Insurance, CGU, SGIO, SGIC, Swann, WFI and Lumley Insurance (Australia); NZI, State, AMI and Lumley Insurance (New Zealand); Safety and NZI (Thailand); and AAA Assurance (Vietnam). We also have interests in general insurance joint ventures in Malaysia, India and China.

Last year, we insured over $2 trillion worth of assets on behalf of millions of customers, and we paid $736 million in claims. Increasingly, we see our role extending beyond paying claims to increasing awareness of risk, and helping communities reduce and prevent risk. We believe it is our responsibility as an industry leader to use our influence and role as a major investor, purchaser and employer for the good of everyone.

Why we support GovHack 2016

IAG shares GovHack's belief that access to data drives the information age and the digital economy. IAG is a strong advocate of data openness and is excited by the opportunities the digital revolution is creating for our customers, communities and people.

We embrace the Spirit of GovHack – to demonstrate how we can utilise data to make it valuable and accessible like never before – as it opens up a world of possibilities to bring to life our purpose, to make your world a safer place.

As part of our support of GovHack, we are proud to be the International Prize Sponsor for GovHack 2016.

We wish all GovHack participants a very enjoyable three days of hackathon and look forward to the innovative concepts developed.

IAG Flood Risk Data

Data Product Identification

Title

IAG Flood Risk Data

Abstract

IAG are pleased to announce the release of a national dataset containing a view of flood risk across Australia. The released data is available at address and aggregated level. The address level data provides granular flood risk information and is based on the G-NAF February 2016. IAG's coordinates are used in areas where the IAG address database provides more appropriate information for defining primary dwelling flood risk. The second level of data is aggregated at admin boundaries level providing users with an overview of flood risk at a regional/national level.

IAG, through our brand NRMA Insurance, are already leading the way in opening the Australian insurance view of risk. The Safer Homes portal (http://saferhomes.nrma.com.au) is an educational interactive web-based tool that helps customers and non-customers better understand the risk profile of the suburb they live in and the real value of their assets. Practical tips and tools are offered to users to better protect their homes. Safer Homes is the first new home insurance innovation developed with the customer in mind, and has resulted in driving cultural change within the IAG organisation.

IAG recognises that the data behind the tool has numerous potential benefits outside of the insurance industry and this provides the impetus for opening up our data to the public.

The data used for identifying the degree of flood risk is a combination of IAG modelling, industry sourced data, and local and state government data. The insurance industry is uniquely positioned in its ability to relate a physical risk into a financial risk through its exposure to claims data and this dataset utilises these depth-damage functions to derive the risk categories supplied in the national dataset. It should be noted that the data represents a financial risk and consequently are not suitable to assess potential risk to life, public safety etc.

Purpose

In releasing the data, IAG hope to assist in informing the general public on flood risk at their property and surrounding communities. The data release ties in with IAG's core mission of making your world a safer place by:

1. Improving awareness of flood risk to property owners and communities.
2. Promoting self mitigation, especially in areas with significant flood risks.
3. Improving risk communication and coordination between general public, related government agencies and industries.
4. Promoting information sharing to benefit communities.
5. Attracting new innovations in communicating and visualisation of risks.

Summary of Dataset Coverage

The released dataset assesses flood risk for 88% of all addresses nationally. The remaining 12% contains third party modelling which cannot be released due to licencing constraints and the table below provides a breakdown of the source type for IAG's flood risk data.

Source of Data (Source_type)	Count	Percentage	Count of flood risk assessment included	Percentage of flood risk assessment included
GOVERNMENT FLOOD MAPPING	2,291,441	16%	2,291,441	16%
OTHER GOVERNMENT FLOOD DATA	1,054,220	8%	1,054,220	8%
IAG INTERNAL FLOOD MODEL	8,870,278	64%	8,870,278	64%
THIRD PARTY MODEL	1,674,696	12%	0	0%
TOTAL	13,890,635	100%	12,215,939	88%

Source_type categories are describe in data dictionary

Data Category Description

The High, Medium and Low risk bands describe the financial risk. The risk is based on two factors, how severe the flooding is and how often the property is likely to be flooded. You may have heard in the media that a flood is a 1 in 100 year event. This doesn't mean that the flood will happen exactly once every 100 year but that on average it will. A more appropriate way to think about it is that the event has a 1% chance of happening in any given year. A 1% event can be thought of as quite a rare event and there will be events that for example have a 10% chance of happening every year or even a 99% chance of happening every year. The consequences of a rare event are expected to be more severe and it is that relationship between how often an event is likely to occur and how severe (damage to property) an event is that informs the financial risk.

The insurance industry uses Average Annual Damage (AAD) which is the cost of potential flood events integrated across how likely these events are to occur in a given year (flood frequency). This can be thought of as a house that floods very often but with only minor damage versus a house that floods rarely but when it does significant damage is incurred. Both mechanisms may incur Low, Medium or High flood ratings but it is the relationship between the frequency of the event and the cost of the event that determines which risk band the house falls into.

Flood Risk AAD (average_annual_damage)	Count	Percentage
High (H)	388,993	2.8%
Medium (M)	609,402	4.4%
Low (L)	11,217,544	80.8%
NULL	1,674,696	12.0%
TOTAL	13,890,635	100.0%

Flood Risk AAD (average_annual_damage) categories are described in Data Dictionary

All null Values are from Third Party Models

The likelihood of a property being flooded (Flood Frequency) is also provided in banded form to assist the user

Frequency of over ground flooding (FLOOD_FREQUENCY)	Count	Percentage
High (H)	220,045	1.6%
Medium (M)	437,478	3.1%
Low (L)	11,558,416	83.2%
NULL	1,674,696	12.1%
TOTAL	13,890,635	100.0%

Frequency of over ground flooding (FLOOD_FREQUENCY) categories are described in Data Dictionary

All null Values are from Third Party Models

The three aggregated dataset includes values for “total AAD”. This value is calculated using the estimated building flood risk AAD of residential addresses. This doesn’t include home content AAD or the AAD of other assets that might be flood prone like commercial and industry assets. Therefore this value will not accurately estimate the whole community flood risk. The total AAD is best used to identify regions where residential communities are exposed to flooding.

Example of the IAG Flood Risk Data in Maps

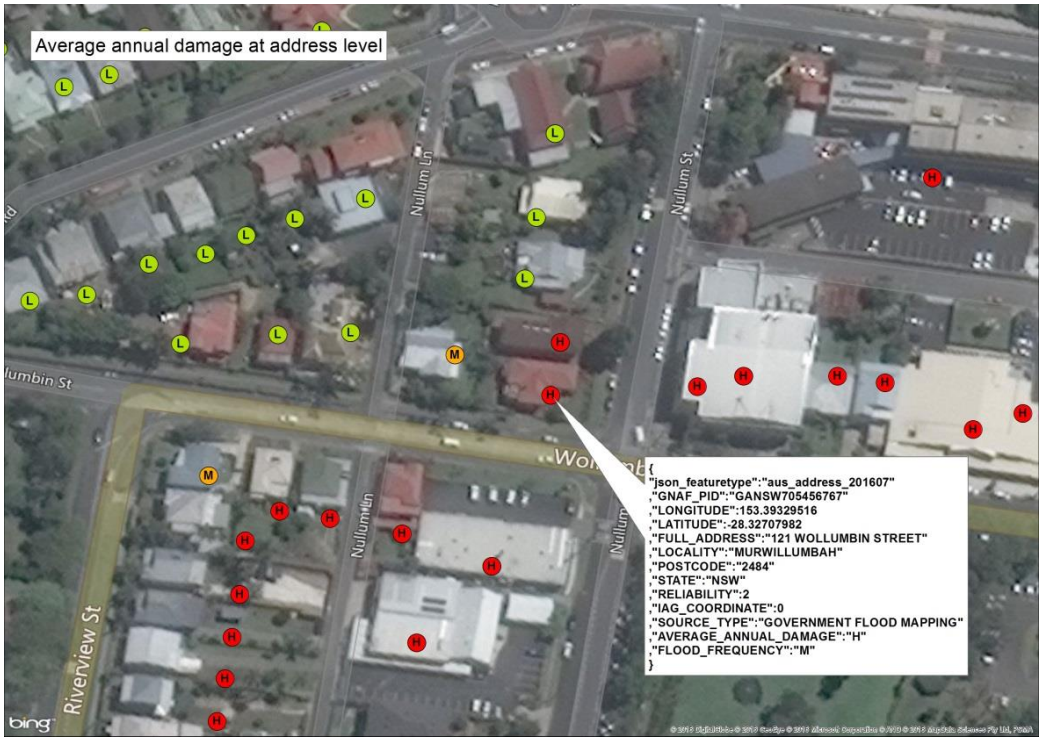


Figure 1, Sample of Address data from API service

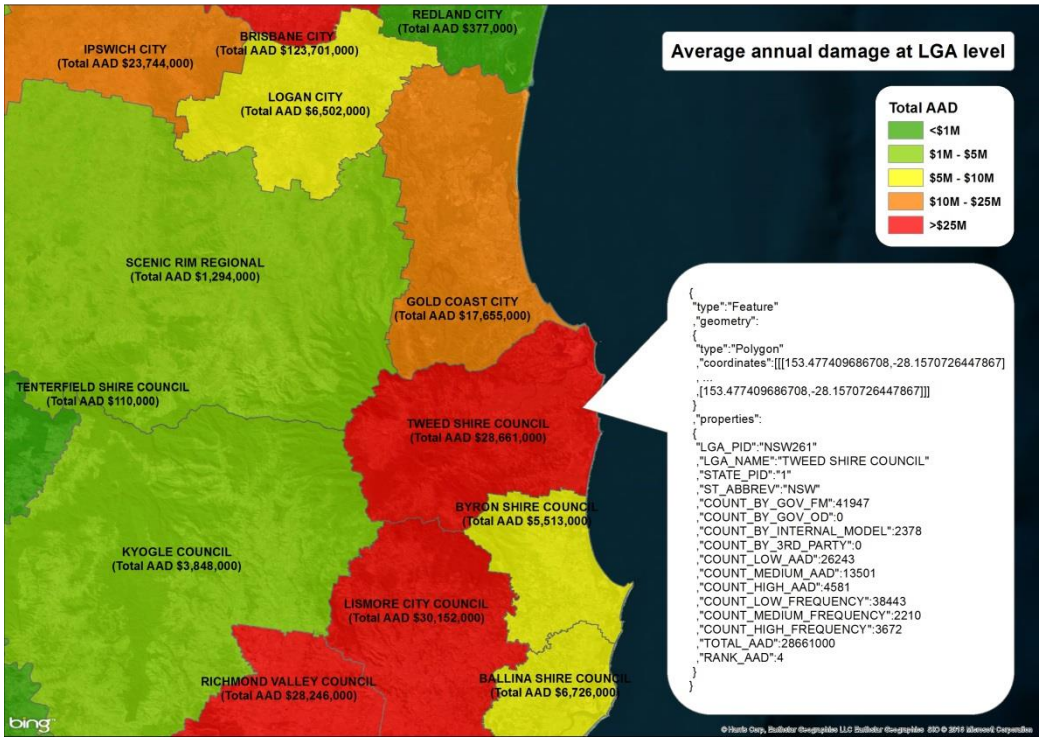


Figure 2, Example of Aggregated data at LGA level

IAG versus GNAF coordinates

GNAF coordinates are typically located at the centroid of the cadastral lot. The centroid is not always representative of the flood risk to a property especially when the asset sits on a large cadastral lot as illustrated on Figure 3. While the provided dataset is supplied in GNAF coordinates, it is important to note that the flood risk bands provided apply to where the major asset is located and not the GNAF point location. This shift affects approximately 2,000,000 points in total.

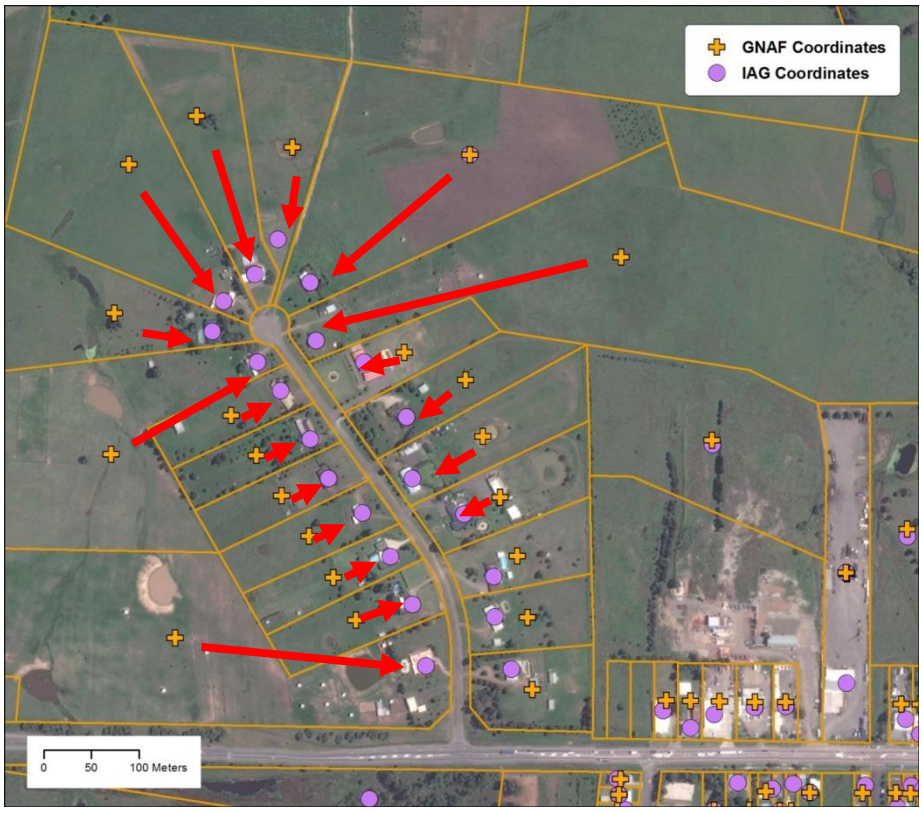


Figure 3, Example of adjustment of GNAF coordinate to better represent flood risk

Third Party Models

IAG uses several third party models from consultants to assess flood risk in areas where Government or IAG internal data is not readily available. This represents about 12% of all addresses nationally. *Due to licensing restriction covering third party models no derived data from these models is provided.* This means flood risk assessments are likely to be underestimated when it relies on a third party model. This will particular affect the accuracy of the total AAD and Rank in the aggregated data.

Reference Date

July 2016

Responsible party and Primary contact

IAG Flood Technical

Insurance Australia Limited

Level 10, 388 George St, Sydney NSW 2000

hydrology@iag.com.au

Terms and definitions

IAG Coordinates	IAG's coordinates are used in areas where the IAG address database provides more appropriate information for defining primary dwelling flood risk

Abbreviations and acronyms

AAD	The Average Annual Damage (AAD) measure the flood risk which combines the likely impacts and probabilities of all possible flood events at a given location, up to and including the Probable Maximum Flood (PMF)
FF	The Flood Frequency (FF) measure the flood average recurrence interval. Combination between AAD and FF provides answers to how much it going to cost individual property owner, and how often it is going to cost them.
G-NAF	Address data provided by PSMA. G-NAF License is available here http://data.gov.au/dataset/19432f89-dc3a-4ef3-b943-5326ef1dbecc/resource/09f74802-08b1-4214-a6ea-3591b2753d30/download/20160226---EULA---Open-G-NAF.pdf

Data Product Delivery

Delivery Medium Information

Your access to and use of the information constitutes your agreement to the Terms of Use. The Term of Use are available from GitHub

Address level flood risk data

This is available via an API with valid key

API is located here

http://flood-risk-api.app.skyops.io/

API key is available from GitHub location.

Aggregated Flood Risk Data

CSV and JSON format flood risk data is available from GitHub

https://github.com/iag-edge-labs/flood-data

CSV format contains flood risk data and no administrative boundary GIS data. Use the G-NAF permanent identifier to join CSV format data to GIS G-NAF administrative boundaries

JSON format contains flood risk data and GIS administrative boundaries.

Appendix A – Data Dictionary

IAG ADDRESS LEVEL FLOOD RISK

FIELD NAME	DATA TYPE	DESCRIPTION	POSSIBLE VALUES
GNAF_PID	Varchar(15)	GNAF's permanent identifier	
LONGITUDE	Numeric(11,8)	GNAF's longitude	
LATITUDE	Numeric(10,8)	GNAF's latitude	
FULL_ADDRESS	Varchar(255)	GNAF's address	
LOCALITY	Varchar(100)	GNAF's locality	
POSTCODE	Varchar(4)	GNAF's postcode	
STATE	Varchar(3)	GNAF's state	
RELIABILITY	Integer	GNAF's geocode reliability	
IAG_COORDINATE	Integer	Flag indicates which coordinates were used to assess the flood risk	• *1* = IAG coordinate was used to assess flood risk at this address. • *0* = GNAF coordinate was used to assess flood risk at this address.
SOURCE_TYPE	Varchar(29)	Type of flood hazard information	• *Government Flood Mapping* = State or local government flood mapping has been interpreted to risk assess flood risk at the site • *Other Government Flood Data* = State or local government information including flood gauge data or flood levee data has been interpreted to assess flood risk at the site. • *IAG Internal Flood Model* = In areas where government data is not readily available, IAG's internal flood modelling has been used to assess flood risk at the site. • *Third Party Model* = In areas where government data is not readily available, modelling by risk consultants has been used to assess flood risk at the site.

FIELD NAME	DATA TYPE	DESCRIPTION	POSSIBLE VALUES
AVERAGE_ANNUAL_DAMAGE	Varchar(1)	Expected financial average annual damage for flood risk	• L = Expected low average annual damage (≤ $100 for a typical Australian dwelling) • M = Expected medium average annual damage (> $100 AND ≤ $1,000 for a typical Australian dwelling) • H = Expected high average annual damage (> $1,000 for a typical Australian dwelling)
FLOOD_FREQUENCY	Varchar(1)	Frequency of any over ground flooding	• L = Low average recurrence interval (≥ 100 years) • M = Medium average recurrence interval (≥ 20 AND < 100 years) • H = High average recurrence interval (< 20 years)

RESIDENTIAL FLOOD RISK – AGGREGATED AT LOCALITY BOUNDARIES

Note: this will not assess flood risk impact to whole community as it only aggregates residential flood risk

FIELD NAME	DATA TYPE	DESCRIPTION
LOC_PID	Varchar(15)	GNAF's locality permanent identifier
NAME	Varchar(100)	GNAF's locality name
LOC_CODE	Varchar(1)	GNAF's locality code
LOC_DESC	Varchar(50)	GNAF's locality description
STATE_PID	Varchar(15)	GNAF's state permanent identifier
ST_ABBREV	Varchar(3)	GNAF's state
COUNT_BY_GOV_FM	Integer	Number of residential addresses assessed using Government Flood Mapping (GOV_FM).
COUNT_BY_GOV_OD	Integer	Number of residential addresses assessed using Other Government Flood Data (GOV_OD)
COUNT_BY_INTERNAL_MODEL	Integer	Number of residential addresses assessed using IAG's internal flood model
COUNT_BY_3RD_PARTY	Integer	Number of residential addresses assessed using third party models
COUNT_LOW_AAD	Integer	Number of residential addresses with expected low average annual damage (≤ $100 for a typical Australian dwelling).
COUNT_MEDIUM_AAD	Integer	Number of residential addresses with expected medium average annual damage (> $100 AND ≤ $1,000 for a typical Australian dwelling)
COUNT_HIGH_AAD	Integer	Number of residential addresses with expected high average annual damage (> $1,000 for a typical Australian dwelling)
COUNT_LOW_FREQUENCY	Integer	Number of residential addresses with low average recurrence interval (≥ 100 years)
COUNT_MEDIUM_FREQUENCY	Integer	Number of residential addresses with medium average recurrence interval (≥ 20 AND < 100 years)
COUNT_HIGH_FREQUENCY	Integer	Number of residential addresses with high average recurrence interval (< 20 years)
TOTAL_AAD	Big Integer	Total expected average annual damage for the only *residential* addresses in the locality. The figure is only for buildings, and uses typical building values for the locality (rather than the typical Australian dwelling figures reported in the per-address dataset).

FIELD NAME	DATA TYPE	DESCRIPTION
RANK_AAD	Integer	Rank of the average annual damage for *residential* addresses in this locality compared to other localities in Australia.
SHAPE	Geography	ONLY available in the JSON format data

RESIDENTIAL FLOOD RISK – AGRREGATED AT LOCAL GOVERMENT AREA BOUNDARIES

Note: this will not assess flood risk impact to whole community as it only aggregates residential flood risk

FIELD NAME	DATA TYPE	DESCRIPTION
LGA_PID	Varchar(15)	GNAF's LGA permanent identifier
LGA_NAME	Varchar(100)	GNAF's LGA name
STATE_PID	Varchar(15)	GNAF's state permanent identifier
ST_ABBREV	Varchar(3)	GNAF's state
COUNT_BY_GOV_FM	Integer	Number of residential addresses assessed using Government Flood Mapping (GOV_FM)
COUNT_BY_GOV_OD	Integer	Number of residential addresses assessed using Other Government Flood Data (GOV_OD)
COUNT_BY_INTERNAL_MODEL	Integer	Number of residential addresses assessed using IAG's internal flood model
COUNT_BY_3RD_PARTY	Integer	Number of residential addresses assessed using third party models
COUNT_LOW_AAD	Integer	Number of residential addresses with expected low average annual damage (≤ $100 for a typical Australian dwelling)
COUNT_MEDIUM_AAD	Integer	Number of residential addresses with expected medium average annual damage (> $100 AND ≤ $1,000 for a typical Australian dwelling)
COUNT_HIGH_AAD	Integer	Number of residential addresses with expected high average annual damage (> $1,000 for a typical Australian dwelling)
COUNT_LOW_FREQUENCY	Integer	Number of residential addresses with low average recurrence interval (≥ 100 years)
COUNT_MEDIUM_FREQUENCY	Integer	Number of residential addresses with medium average recurrence interval (≥ 20 AND < 100 years)
COUNT_HIGH_FREQUENCY	Integer	Number of residential addresses with high average recurrence interval (< 20 years)

FIELD NAME	DATA TYPE	DESCRIPTION
TOTAL_AAD	Big Integer	Total expected average annual damage for the only *residential* addresses in the LGA. The figure is only for buildings, and uses typical building values for the LGA (rather than the typical Australian dwelling figures reported in the per-address dataset).
RANK_AAD	Integer	Rank of the average annual damage for *residential* addresses in this LGA compared to other LGAs in Australia.
SHAPE	Geography	ONLY available in the JSON format data

RESIDENTIAL FLOOD RISK – AGRREGATED AT COMMONWEALTH ELECTORAL BOUNDARIES

Note: this will not assess flood risk impact to whole community as it only aggregates residential flood risk

FIELD NAME	DATA TYPE	DESCRIPTION
CE_PID	Varchar(15)	GNAF's Commonwealth Electoral (CEL) permanent identifier.
NAME	Varchar(100)	GNAF's CEL name
STATE_PID	Varchar(15)	GNAF's State permanent identifier
ST_ABBREV	Varchar(3)	GNAF's State
COUNT_BY_GOV_FM	Integer	Number of residential addresses assessed using Government Flood Mapping (GOV_FM)
COUNT_BY_GOV_OD	Integer	Number of residential addresses assessed using Other Government Flood Data (GOV_OD)
COUNT_BY_INTERNAL_MODEL	Integer	Number of residential addresses assessed using IAG's internal flood model
COUNT_BY_3RD_PARTY	Integer	Number of residential addresses assessed using third party models
COUNT_LOW_AAD	Integer	Number of residential addresses with expected low average annual damage (≤ $100 for a typical Australian dwelling)
COUNT_MEDIUM_AAD	Integer	Number of residential addresses with expected medium average annual damage (> $100 AND ≤ $1,000 for a typical Australian dwelling)
COUNT_HIGH_AAD	Integer	Number of residential addresses with expected high average annual damage (> $1,000 for a typical Australian dwelling)
COUNT_LOW_FREQUENCY	Integer	Number of residential addresses with low average recurrence interval (≥ 100 years)
COUNT_MEDIUM_FREQUENCY	Integer	Number of residential addresses with medium average recurrence interval (≥ 20 AND < 100 years)
COUNT_HIGH_FREQUENCY	Integer	Number of residential addresses with high average recurrence interval (< 20 years)
TOTAL_AAD	Big Integer	Total expected average annual damage for only *residential* addresses in the commonwealth electorate. The figure is only for buildings, and uses typical building values for the electorate (rather than the typical Australian dwelling figures reported in the per-address dataset).
RANK_AAD	Integer	Rank of the average annual damage for *residential* addresses in this commonwealth electoral compared to other commonwealth electoral in Australia.
SHAPE	Geography	ONLY available in the JSON format data

Appendix B – Terms of Use

Terms of Use – IAG Flood Risk Data

Your access to and use of this data is governed by these Terms of Use. Your access to and use of the data constitutes your agreement to these Terms of Use.

"This data" means all data provided to you by Insurance Australia Limited (ABN 11 000 016 722) (**IAL**).

If you use or reproduce this data, you must, amongst other things, identify the creator of this data in any reasonable manner as requested by Insurance Australia Limited.

Please note: **This data may not be used for insurance, reinsurance or comparative pricing purposes.**

Other Important Information

To the maximum extent permitted by law, IAL provides the data on an 'as is' basis and makes no warranties or representations, whether express or implied, as to the accuracy or reliability of the data or that the data is fit for a particular purpose. Further, you acknowledge that the data provided is an assessment of financial risk only and is not necessarily suitable for use in assessing any other kinds of risk, including risks regarding safety.

To the maximum extent permitted by law, IAL excludes liability of any kind (including for negligence) to any person arising directly or indirectly from accessing, using or being unable to access or use, the data.

If you contravene these Terms of Use, IAL reserves the right to terminate, modify or suspend your access to the data, in its sole discretion, without notice and without prejudice to any other rights that it may have.

Any failure by IAL to exercise any right, power or remedy under these Terms of Use shall not operate as a waiver of IAL's rights.

The law of New South Wales governs these Terms of Use.