Automated generation of a nationally consistent surface water digital network

Prepared for Regional Software Holdings Limited
Ministry for the Environment

July 2026

Prepared by:
Doug Booker, Matt Wilkins, Rachel Smith, Clare Wilkinson, Paul Franklin

For any information regarding this report please contact:

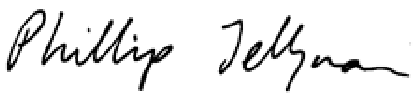
Doug Booker
Principal Scientist - Hydro-ecological modelling
Freshwater Modelling Group
+64 3 343 7848
doug.booker@earthsciences.nz

New Zealand Institute for Earth Science Limited
PO Box 8602
Riccarton
Christchurch 8440

Phone +64 3 348 8987

Client Report No: 2026189CH
Report date: July 2026
Project No: RSH25501

Revision	Description	Date
Version 1.0	Final version sent to client	1 July 2026

Quality Assurance Statement		
	Reviewed by:	Michelle Greenwood
	Formatting checked by:	Terry Smith
	Approved for release by:	Phillip Jellyman

Contents

Tables

Figures

Executive summary

Conceptually, Digital Networks (DNs) represent connections across freshwater-land systems, including rivers, lakes, wetlands, estuaries, and land. Technically, DNs comprise representations of surface flow pathways (segments or polygons), areas contributing to each surface flow pathway (watersheds), and connections between surface flow pathways (routing). DNs are essential for a variety of freshwater research, management, and policy development purposes because they provide spatial frameworks for water quantity modelling, hydrological modelling, species distribution modelling, and river ecosystem classification. Different purposes require different characteristics for DNs in terms of their level of spatial detail (node density), accuracy of local positioning (segment alignment), and spatial extent (watershed coverage).

Three main versions of a national-coverage DN were produced between 1998 and 2018. The previous national DNs have often been referred to as digital river networks because they explicitly represent river channels while not explicitly representing lakes or overland flow pathways. The previous DNs were widely used by various institutions for many different purposes. However, they were produced using now out-of-date topographic data, were generated using a spatially uniform node density, and did not explicitly represent lakes or artificially engineered flow pathways such as diversions between catchments. They were also not easily updatable and expensive to produce due to high manual input requirements.

The current work improved upon previous national DNs by devising and testing a DN generation process that maximised automation and minimised manual input requirements. The process is repeatable to allow efficient DN generation following requests for changes from users or updated topography, lake polygons, artificial channel, or headwater/sinkhole data. The process ingested a hydrologically-corrected Light Detection and Ranging (LiDAR) derived Digital Elevation Model (DEM) and lake polygons obtained from Land Information New Zealand (LINZ). A method was devised for extracting headwater and sinkhole locations from national 1:50,000 scale topographic maps (Topo50).

The new DEM-derived DN improves upon previous national DNs by explicitly representing interactions between surface flow pathways, sinkholes, artificial flow pathways, lakes, river channels, and the coast. The new DN is suitable for a broad range of purposes because it extends to a broad coverage, distinguishes river channels from overland flow pathways, and utilises high resolution information on spatial alignment.

A set of formal tests was devised and applied to validate the correctness of DN functionality following DN generation. Representative correctness was checked by end users who trialled a process of making requests for changes or improvements. Results indicated that, while most of the DN matched user expectations, some misalignment of segments and poorly represented routing were present in areas without LiDAR, on the border between LiDAR datasets, around lakes, or in flatter areas near to the coast.

This report assessed progress against a previously devised set of principles that a DN should meet. When assessed against these principles, it was concluded that the new DEM-derived DN is:

- reproducible because it is accompanied by this report which describes technical steps applied in its generation;

- automatically updatable because the DN generation process can be reapplied following improvements to input data;
- version controlled because DN versions are stored and labelled so that data and methods used for DN generation are traceable and users know whether DN-derived products would be expected to match with each other;
- spatially consistent and spatially complete because the same process was applied using nationally available data;
- functionally correct because it passed a set of checks that confirmed mathematical correctness of its routing behaviour;
- functionally informative because a labelling system was applied that allowed the DN to be sub-sampled (e.g., river channels only) for different purposes; and
- appropriate because its capabilities have been set out and the envisioned purposes have been clearly explained.

It was demonstrated that future work could allow the DEM-derived DN to be:

- made available via a data sharing portal as was demonstrated using the Ministry for the Environment (MfE) data service; and
- bespoke updatable if a method to collate, vet, and approve user requests for DN alterations is developed and operationalised.

The findings highlight the need for an operational mechanism for change requests provided by DN users to be collated, vetted, and resolved to maintain a DN that is fit for user purposes.

1 Introduction

1.1 General background

Digital Networks (DNs) are virtual representations of connections across freshwater-land systems, including rivers, lakes, wetlands, estuaries, and land. DNs must contain segments representing surface flow pathways, nodes representing connections between surface flow pathways, and watersheds representing areas contributing to each segment. A DN is characterised by its level of spatial detail (node density), accuracy of local positioning (segment alignment), and spatial extent (watershed coverage). DN characteristics are determined by a combination of input data, technical methods, and developer decisions.

DNs provide spatial frameworks for water quantity modelling, hydrological modelling, species distribution modelling, and river ecosystem classification. They are necessary for regional-scale planning, environmental reporting, and landscape-scale conservation planning. DNs are therefore essential for a variety of freshwater research, management, and policy development purposes.

1.2 Previous national DNs

The National Institute for Water and Atmospheric Research (NIWA) produced three main versions of a national-coverage DN between 1998 and 2018. The previous DNs have often been referred to as digital river networks because they explicitly represent river channels while not explicitly representing lakes or overland flow pathways. They exhibited some valuable qualities.

- Segments that were assumed to represent river channels.
- Nodes that represented confluences which are the joining segments in the downstream direction.
- Watersheds that represented the area draining to each segment.
- The downstream end of each segment joined with another segment or terminated at the coast.
- All connected segments routed downstream towards a single terminal segment thus each segment belonged to a single catchment that could be identified by its terminal segment.
- Relatively simple DN conceptualisation facilitated straightforward calculation of network attributes such as catchment area, slope, Strahler stream order, catchment membership, etc.

However, they suffered from some disadvantages.

- They were time consuming to produce due to manual input requirements.
- They were generated from now out-of-date topographic data.
- They were generated using a spatially uniform node density that was selected at the time of DN production for a particular purpose (e.g., lower density for river

classification purposes for version 1, higher density for hydrological modelling purposes for version 3).

- Lakes were represented by collections of segments rather than being represented by polygons with inflowing and outflowing segments.
- It was possible for segments to cross lake boundaries.
- Artificial channels such as canals or drains were not represented.
- Braided channels, islands, and diversions between catchments were not represented because the DNs did not include bifurcations that split when routing in the downstream direction.

Several advancements have occurred over time since previous national DNs were generated that impact how a national DN might be generated for Aotearoa New Zealand (NZ).

- DN users have clarified that different DN purposes may require DNs with differing levels of spatial resolution (e.g., flood modelling purposes may need much finer detail than river management purposes) (see Section 1.3.1 and Booker 2023).
- The Ministry for the Environment (MfE) have indicated a desire to produce ecosystem classifications that integrate across physical domains (e.g., rivers, lakes, wetlands, terrestrial, estuarine environments) (Sprague et al. 2026).
- High resolution topographic data have been collected for most of the country (Pearson et al. 2023).
- Access to geographic data comprising cartographic maps has improved (e.g., https://data.linz.govt.nz/).
- Improved computing resources and availability of open-source software have aided automated generation of DNs (e.g., GRASS or SAGA with QGIS).
- A method for incorporating and distinguishing between overland flow pathways and river channels within river networks has been devised (Booker et al. 2025).
- A method that extends the traditional Strahler ordering system to use zero and negative numbers has been published (Booker et al. 2025). The method can be used to distinguish mapped river channels from overland flow pathways and consistently compare digital networks. The method relies on using mapped locations to initiate first order streams rather than relying on an upstream contributing area threshold for stream initiation.
- NIWA has been merged with Geological and Nuclear Sciences (GNS) to form Earth Sciences New Zealand (Earth Sciences NZ). Bringing surface water and groundwater researchers together within the same institution may facilitate integration of surface water routing and groundwater routing within a DN.

1.3 How the current report follows from previous reports

This report is the third in a series of reports about national DNs. The first report outlined user needs (Booker 2023). The second report outlined challenges and solutions to meet user needs

(Booker et al. 2024). This report outlines methods designed to implement some of the previously proposed solutions. Further details of how the three reports follow each other are described in the remainder of this section.

1.3.1 2023: future pathways for stewardship, maintenance, and upgrading

Canvasing of DN users reported in Booker (2023) indicated that applications of DNs largely fall into two broad categories. The first category required DNs to comprise detailed representation of flow routing along overland flow pathways and river channels for biophysical modelling such as nutrient or flood inundation modelling. The second category required DNs to only represent river channels (or distinguish between overland flow pathways and river channels) for river management purposes such as targeting sites for river restoration. Contrasting needs for DN applications demonstrated that a multi-purpose DN must simultaneously extend to a broad coverage, distinguish rivers from overland flow pathways, and include high resolution information on spatial alignment.

Booker (2023) clarified that biophysical modelling purposes include hydrological or water quality modelling that simulate fate and flux of substances within watersheds and between segments. Depending on the specific application and the data available for model parameterisation, biophysical modelling can benefit from detailed spatial representation of surface flow pathways both inside and outside of river channels. Detailed representation of surface flow pathways is particularly beneficial for flood modelling (Blöschl 2022) and is also important when intercepting river channels with land ownership boundaries for consenting purposes. River management and policy development purposes include landscape-scale river classification that require information on river channels and their interconnections but do not require information on surface flow pathways outside of river channels (Snelder and Biggs 2002). Landscape-scale representation of river channels and their catchments is particularly beneficial for regional- and national-scale conservation planning (Stoffels and White 2025). The report recognised that biophysical modelling often feeds into river management purposes.

Booker (2023) outlined a future vision for stewardship, maintenance, and upgrading of DNs to support both biophysical modelling and river management from a national perspective. The report proposed a framework to operationalise ongoing development and support of national DNs, river classifications, and their dependent products. The proposed framework included five types of activity.

1. Project oversight group to steer overall aims, intermediate objectives, and timing delivery.
2. Project management to prioritise tasks, distribute resources to complete tasks, and report on delivery.
3. Technical tasks to advance methods and deliver outputs and services (see next paragraph for details).
4. Trial users to test outputs and services.
5. Communication pathways to receive user feedback, and to ensure users are aware of outputs and services.

Technical tasks that the proposed framework could facilitate and co-ordinate fall into two categories.

1. Operational tasks where methods are known but not implemented.
   a. Workflow to produce DN using DEM as input.
   b. Tool for users to submit recommended changes to DN alignment or routing.
   c. Procedure to ingest, assess, verify, and apply recommended changes.
   d. Procedure to calculate REC classes on any DN and using any input environmental data.
   e. Tool for users to apply routing (e.g., downstream accumulation to calculate conditions in the area upstream of each segment).
2. Research tasks where methods are currently unknown or untested.
   a. Procedure to classify flow pathways as being inside or outside of river channels.
   b. Procedure to identify river channels from remotely sensed images.

Booker (2023) suggested that, if enacted, the proposed framework might produce six outcomes.

1. Produce DNs that are consistent with each other and fit for purposes ranging from local-scale flood modelling to national-scale river classification by producing parent DNs comprising highly detailed surface flow pathways and child DNs comprising only surface flow pathways that are classified as river channels.
2. Encourage consistency in use of DNs across applications by ensuring that segment alignment and identifiers are comparable between parent and child DNs.
3. React to improvements in data availability and methodological advancements by developing an automated procedure to produce DNs and DN-derived outputs following updates to inputs.
4. Produce DNs in a replicable and transparent way by applying best practices for version control.
5. Reduce unnecessary duplication of effort and confusion at the national level by allowing various agencies to provide input about their needs for DNs and river classifications.
6. React to user needs by ingesting and prioritising requests for amendments to segment alignment or network routing.

1.3.2 2024: technical challenges and solutions

Booker et al. (2024) described a sequence of technical steps required for DN generation, regardless of technical details relating to data or algorithms. Booker et al. (2024) also described some technical advancements designed to increase DN utility, including a multi-coloured labelling system for DN objects. A multi-coloured labelling system could be used to generate a single network that can be sub-sampled to represent various resolutions, and then consistently applied for purposes with varying needs. Blue objects within a network would represent river channels, purple objects can be added to represent ephemeral flow pathways downstream of river channels, grey objects can be added to represent engineered flow pathways, and red objects can be used to represent surface flow pathways upstream or river channels at the highest possible level of detail. Booker et al. (2024) proposed a negative Strahler stream order scheme to facilitate consistent comparisons of network connectivity between DNs regardless of their resolution. The extended Strahler ordering system was published in Booker et al. (2025).

Booker et al. (2024) outlined several challenges for DN generation associated with various landscape phenomena. For example, it is desirable to represent lakes, wetlands, estuaries, braided rivers, springs, and sink holes in DNs but they are not explicitly represented in DEMs and flow directions within these landscape phenomena may not be easy to extract from a DEM. Existing nationwide data and potential solutions to improve representation of these phenomena were outlined and demonstrated for two case study catchments.

Booker et al. (2024) clarified that multi-channel rivers and artificial channels require bifurcations to be included in DN routing. The report proposed to improve DN functionality by representing bifurcating channels so that islands, braided rivers, and artificial channels can be represented. The report proposed that lakes should be included as polygons in DNs to allow representation of processes such as storage and mixing in lakes. Various solutions were set out to meet the challenge of routing through lakes.

Booker et al. (2024) also set out some checks that can be applied to a DN after it has been produced to verify its routing is behaving correctly. Approaches for testing DN alignment, including subjective visual comparisons against remotely sensed images and objective measures of difference with digitised maps were discussed. However, both approaches for testing segment alignment are challenged by a lack of independent ground-truth data for DN validation.

Booker et al. (2024) proposed the following principles for DNs to guide their future generation in view of the breadth of user needs, likely ongoing improvements to input data, and the need for spatial consistency.

1. To be reproducible, DNs should be accompanied by a description of technical steps applied in their generation.
2. To be automatically updatable, DN generation should be as automated as possible so that DNs can be efficiently updated following improvements to input data.
3. To be version controlled, DN generation code should be stored and labelled so that it can be reapplied and amended when needed, and DN products should be stored and labelled so that data and methods used for DN generation are traceable and users know whether DN-derived products would be expected to match with each other.
4. To be spatially consistent, landscape phenomena (e.g., lakes) should be delineated within DNs so that separate procedures can be consistently applied within their bounds.

5. To be bespoke updatable, a method to collate, vet, and approve user requests for DN alterations should be developed and operationalised.
6. To be spatially complete, DN generation procedures should be applied within an agreed coastline, and any blank areas that have been filled should be identifiable.
7. To be functionally correct, DNs should pass a set of checks that confirm mathematically correct routing behaviour before they are released.
8. To be functionally informative, a labelling system should be applied so that a DN can be sub-sampled to be utilised for different purposes.
9. To be appropriate, DN capabilities and the envisioned purposes should be clearly explained alongside released DNs.
10. To be available, DN users should be able to obtain DNs, including previous versions, from a stable source.

1.3.3 2026 (current report): automated generation for an improved national DN

This report documents work conducted between July 2025 and June 2026 by Earth Sciences New Zealand (Earth Sciences NZ). The work was designed to fulfil many, but not all, issues raised in the two previous reports (Booker 2023; Booker et al. 2024) as described in Table 1-1. The main aim of the work was to devise a process for automated generation of a national-coverage DN that incorporates several advancements compared to previous national DNs. The work was carried out for the collective of Regional and Unitary Councils (Te Uru Kahika) using funding provided by the Ministry for the Environment (MfE). The work was designed to improve upon previous national DNs in four main ways.

1. Previous national DNs required considerable manual input during their generation. We set out to devise a DN generation process that maximised automation and minimised manual input requirements. Thus, the process was designed to be repeatable to allow a new DN to be generated following updates to input data (e.g., topographic data) or requests for changes from users.
2. Previous national DNs were not necessarily subjected to a set of formal tests. We set out to apply some formal tests to validate the correctness of DN functionality following DN generation.
3. Previous national DNs were not trialled with DN users prior to their wider release. We set out to share a pilot DN generated from our automated process with a selection of DN users. We also set out to trial methods for sharing instructions about the DN with users and receiving their requests for changes or improvements.
4. Previous national DNs explicitly represented river channels but did not represent the effects of sinkholes, artificial flow pathways, or lakes on routing through river channels. We set out to devise a DN that explicitly represented interactions between various surface water entities (Table 3-1).

We assessed progress against the six outcomes proposed by Booker (2023) and the 10 criteria for DNs proposed by Booker et al. (2024).

Table 1-1: Connections between previously proposed actions and the current work.

Category	Previously proposed action	Relevance to the current work
Activities	Project oversight group to steer aims and objectives.	A project oversight group convened monthly to discuss progress (Section 3.1).
	Project management.	Standard Earth Sciences NZ project management principles were applied.
	Technical tasks and deliver.	See operational and research development tasks in rows below (Section 3.2 and 3.3).
	Trial users to test outputs and services.	Members of the project oversight group acted as trial users (Section 3.5.2).
	Communication of outputs and services.	Instructions about outputs and services provided by our pilot DN were posted on the MfE data service (Appendix C). This report provides further details of DN outputs (Section 4.2).
Operational tasks	Workflow to produce DN using DEM as input.	The core contracted objective of this work. See further details provided in this report (Section 3.2 and 3.3).
	Tool for users to recommend changes.	Not a contracted requirement for this work, but a dedicated email address was established to receive recommended changes during this work (Section 3.5.2).
	Procedure to ingest, assess, verify, and apply recommended changes.	Not a contracted requirement for this work, but a method for ingesting recommended changes was trialled during this work (Section 3.5.2).
	Procedure to calculate REC classes.	Not contracted for this work.
	Tool to apply routing.	Not contracted for this work, but code identifying segments upstream and downstream of a selected segment was developed and distributed to end users (Appendix C).
Research tasks	Procedure to classify segments as being either inside or outside of river channels.	A contracted objective of this work. See further details provided in this report (Section 3.2 and 3.3).
	Procedure to identify river channels from remotely sensed images.	Not contracted for this work.

2 Input data

2.1 Regional data

The work was guided by an oversight group consisting of representatives from regional councils, LINZ, MfE, and Te Uru Kahika. The oversight group met monthly during the course of the work. We used a formal survey to canvass the project oversight group for data that council staff thought might be useful as input to the DN generation process. Our survey results are included in Appendix B. We found that some useful data were available (e.g., locations of culverts, ditches, etc), but the coverage was different between regions, and the data type/definitions were very inconsistent between data providers. We did not use any regionally-specific input data to maintain national consistency of the DN. We used data that were available for the entire country excluding outlying islands as described in the remainder of this section.

2.2 LiDAR-derived DEM

We used a hydrologically-corrected 8 m resolution nationwide DEM that had been previously generated from high-resolution Light Detection and Ranging (LiDAR) data where it was available, and contour data where LiDAR data were not available. The process for generating the DEM is fully described by Pearson et al. (2023). The DEM covered the main islands of NZ but did not cover outlying islands such as the Chatham Islands and the Auckland Islands. The DEM incorporated ocean bathymetry data that extended beyond the immediate coastline. “Daylighting” had previously been applied to remove the artificial flow-blocking effect of bridges and culverts from waterways but “burning” (Lindsay 2016) had not been applied to force the DEM to follow cartographic mapping of waterways. We used the 8 m grid DEM to generate our DN to maintain consistency within our work despite an equivalent DEM with 4 m resolution becoming available during the work.

2.3 LINZ topographic map layers

Land Information New Zealand (LINZ) provides topographic mapping for the main islands of NZ, Chatham Islands, and most offshore islands, at 1:50,000 scale[1]. LINZ 1:50k topographic maps consist of several data layers that we refer to collectively Topo50. We used Topo50 data as inputs because they provide a nationally consistent data source defining the coastline, river and lake locations, and other features relevant to DN generation. We considered applicability of all layers within the LINZ Topo50 hydrography dataset (LINZ Set ID 4781) and three additional LINZ layers (NZ Dam Centrelines, NZ Coastlines and Islands Polygons, NZ Geographic Names) for DN generation. Definition of Topo50 data layers that all appear as blue lines in maps were of particular interest (Table 2-1). We included layers representing surface water flow pathways (e.g., river, lake, lagoon) and excluded layers not representing surface water flow pathways (e.g., snow, ice, pond, swamp, reef). A description of each layer that we considered using in our process and the rationale for its inclusion or exclusion is given in Appendix A.

[1] Improvements to the LINZ Topo50 series are continuously being released. See [Topo50 release notes and update history | Toitū Te Whenua - Land Information New Zealand] for update and version histories.

Table 2-1: LINZ data layers appearing as blue lines on 1:50k topographic maps. Note we provide a suggested alternative definition for the drain data layer.

Data layer	Topo50 definition	Expected characteristics
River	A natural, flowing body of water emptying into an ocean, lake or other body of water and usually fed along it's course by converging tributaries.	Often not straight because meandering or braiding of the river channel creates bends when water interacts with sediment and vegetation within a valley or floodplain.
Drain	A channel which carries surface water. Earth Sciences NZ suggested alternative: An artificially constructed channel that typically drains water from predominantly flat land.	Often have ambiguous flow directions, but when a drain does have a flow direction it slopes towards a river, lake, or the coast. Appear as a series of interconnected straight channels orientated parallel or perpendicular to each other to form a regular pattern.
Water race	A man-made water channel, often running parallel to the contours, built specifically for the distribution of water to crops, pasture, mining activities etc.	Typically, single channel with direction of flow away from a natural feature supplying water toward a location associated with a use of water.
Canal	An artificial open waterway used for transportation, waterpower or irrigation.	Share characteristics of water races, except are larger and can be straight rather than follow contours. Are often named.

2.4 Lake polygons

We used a database describing the outline of lakes obtained from LINZ (LINZ lake polygons). The database contained lakes that were represented by multi-polygons and lakes that touched or overlapped with each other. We post-processed lake polygons so that each lake was represented by a single isolated polygon. After our post processing, lake areas ranged from 2 m^2 to 614 km^2 while median lake area was 1,347 m^2.

2.5 Coastline

We used a database describing the coastline of NZ including islands obtained from LINZ (LINZ coastline). The database contained 9,057 curved polygons. Each polygon represented the coast of a separate landmass. We did not post-process the coastline database.

3 Methods

3.1 Overall approach

We devised the approach depicted in Figure 3-1 to generate a DN that represented the various surface water entities described in Table 3-1. We first used an existing hydro-corrected LiDAR-derived DEM to extract the point locations of river headwaters and sinkholes from topographic maps. A routed DN that resembled the topographic map but did not have watersheds or catchment areas was produced as a byproduct of this initial step. We then generated a multi-coloured DN from the LiDAR-derived DEM, headwaters, sinkholes, lakes and coastline. We then tested the DN for functional correctness, distributed a pilot version to users, received feedback, and iteratively amended our DN generation process. We further tested DN functionality when conducting an associated piece of work requested by LINZ and funded by MfE to route river names along DN segments (Booker et al. 2026). Because we set out to reduce manual steps in DN generation, we tested the hypothesis that a DN could be solely generated using automated processes. However, we included the ability to fix the position of segments using a process known as burning which lowers the DEM along a line (e.g., a Topo50 river line) to the coast.

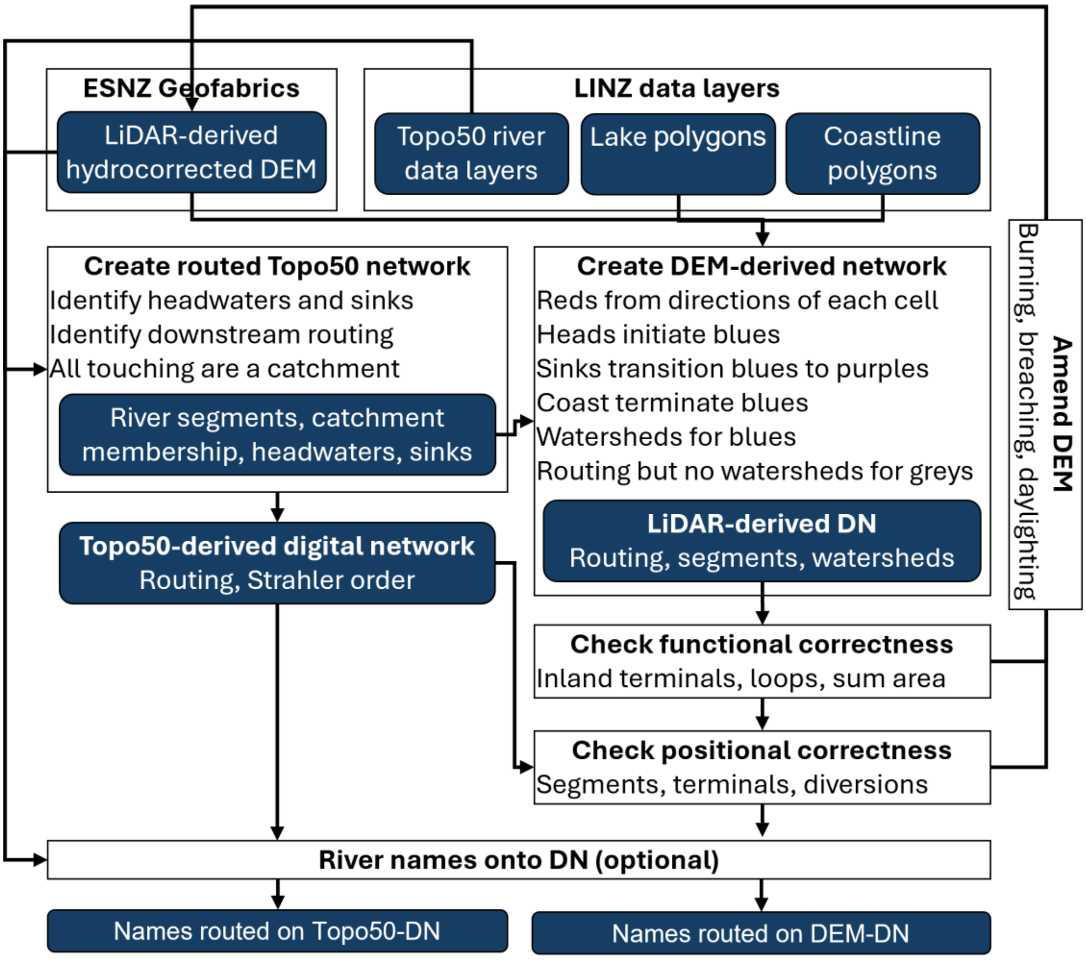


Figure 3-1: Overview of DN generation process. ESNZ = Earth Sciences New Zealand. LINZ = Land Information New Zealand. LiDAR = Light Detection and Ranging. DEM = Digital Elevation Model. DN = Digital Network. Topo50 = 1:50k topographic maps

Table 3-1: Entities relevant to digital network generation.

Name	Description	Reason
Rivers	Body of water with unidirectional flow direction and an associated draining watershed.	Main pathway for transport of water through surface freshwater systems.
Watersheds	Areas draining directly to each surface flow pathway.	Used to calculate local rainfall or local contaminant contributing to each segment.
Catchments	Collection of watersheds and segments upstream of a common terminal segment.	Can be used to identify a set of segments those upstream/downstream states are interdependent.
Headwaters	Location at which a river channel is initiated.	Determines positioning of headwater segments and network density.
Sinkholes	Inland location at which a river channel ends.	Recognition that river channels may terminate inland of the coast.
Lakes	Polygon-shaped body of water with no flow direction and potentially multiple inlets and outlets.	Allows the effects of lakes on flow routing to be represented.
Overland flow pathways	Flow directions upstream of river channels.	Improved routing of water, contaminants, and sediment outside of river channels.
Artificial flow pathways	Engineered flow pathways that do not have catchment area, whose flow direction is not necessarily reflected in DEM slopes, and may result in inter-basin transfer of water.	Allows the effects of diversions to be represented in flow routing.
Coastline	Boundary between land and ocean outside of which DN segments are not meaningful.	Needed to identify terminal locations where river channels and overland flow pathways will discharge to the ocean.

3.2 Topo50-derived digital network

Our processes for extracting headwaters and sinkholes from Topo50 topographic map data is depicted in Figure 3-2. Our process for subsequently generating a routed river network from Topo50 topographic map data is depicted in Figure 3-3. Some parts of the process were particularly noteworthy.

- The Topo50 data objects included in DN generation (see Appendix A) were assumed to be connected as part of the same catchment if they were touching or near enough to appear to be touching on a printed map. This approach was necessary because some of the objects plotted very near to each other but were not intersecting (touching).
- Drains (as defined by objects in the Topo50 Drain Centrelines data layer, see Appendix A) were only part of a catchment if they served to link catchments to the coast that would otherwise be isolated from each other. This approach was

necessary because we did not want drains to initiate river headwaters, but objects contained within the Topo50 Drain Centrelines data layer (possibly representing straightened rivers) were positioned between two river objects that would not be connected if the drain was absent.

- Polygons representing rivers were converted to lines using a skeletonization algorithm. This approach was necessary because some Topo50 river objects were polygons that have no ends to represent headwaters or sinkholes.
- Unconnected ends of river lines were designated as either headwaters, sinkholes, or terminals. This approach was necessary to subsequently identify points that could be headwaters or sinkholes.
- The lowest isolated end of a river line in a catchment was designated a sinkhole if it was located inland and designated a terminal if it was located within a threshold of distance to the coastline. This approach was necessary to extract sinkholes.
- All isolated ends of lines that were not a catchment sinkhole or terminal were headwaters. This approach was necessary to extract headwaters.

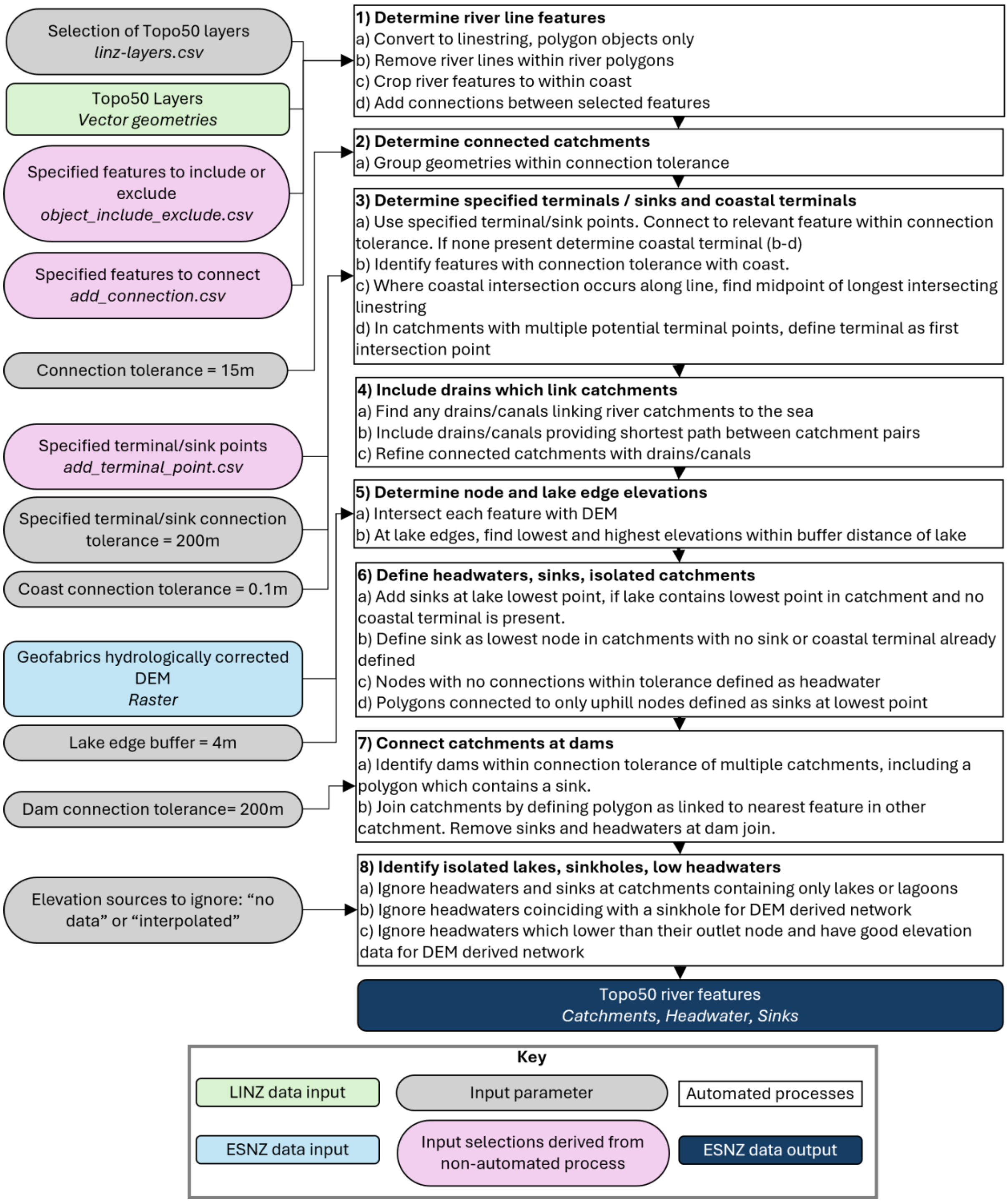


Figure 3-2: Process for identifying headwaters and sinkholes from Topo50 data.

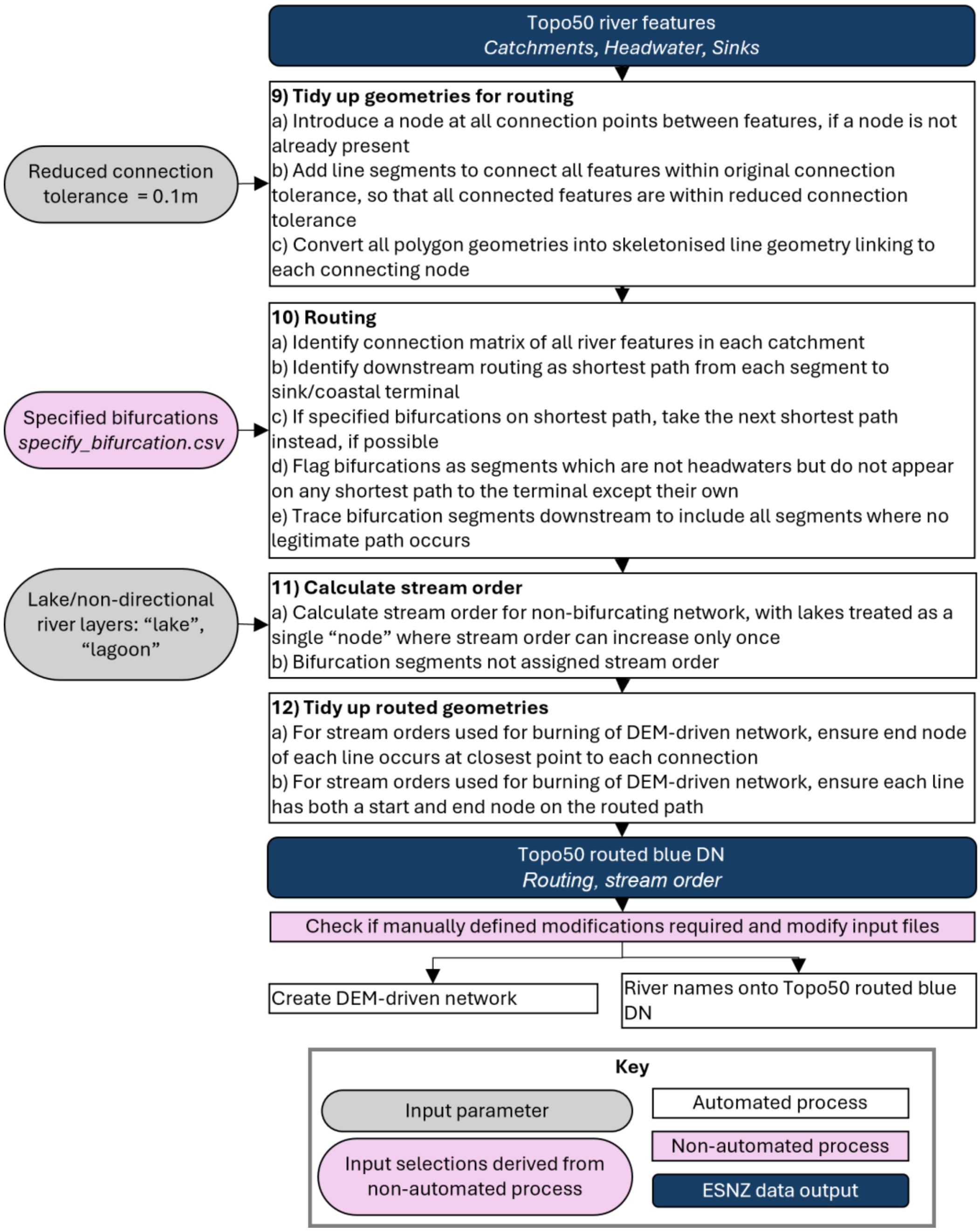


Figure 3-3: Process for generating a routed network from Topo50 data. We used text input files to allow a user to alter the data included in the Topo50-derived DN generation process (Table 3-2).

Table 3-2: Description of input files used to generate a Topo50-derived DN.

File	Description
linz-layers.csv	Specifies the LINZ layers to be used for each part of the process e.g. as the coast, as surface water, for naming.
object_include_exclude.csv	Specify which objects comprise the 'blue network' when the default rules in linz-layers.csv should not be followed. LINZ feature identifiers are used to include or exclude specific features.
add_connection.csv	Used to add connecting line features between pairs of LINZ features. These connecting lines will be added into the 'blue network'. Features will be joined by a straight line between their nearest points.
add_terminal_point.csv	Used to specify a terminal point associated with a specified LINZ feature. The terminal point will be set as the terminal for all connecting features (i.e. the entire catchment) of the specified LINZ feature.
specify_bifurcation.csv	Used to specify where routing should not follow a specified feature and this feature should be labelled as a 'bifurcation,' if this is possible.

3.3 DEM-derived digital network

3.3.1 DN generation

Our process for generating a DN from a LiDAR-derived DEM, LINZ lakes, LINZ coastline, and Topo50-derived headwaters and sinkholes is depicted in Figure 3-4. Figure 3-5 and Figure 3-6 provide further technical details of the DN generation process. Table 3-3 provides further details about interactions between segments with sinkholes. Table 3-4 sets out user inputs for the DEM-derived DN generation procedure. We initially tested our burning process by burning all segments from our Topo50-derived DN with Strahler order greater than four. The initial results produced by systematic burning were poor due to inconsistencies between burnt locations (from Topo50 segments) and the DEM. We therefore removed systematic burning from the generation process but kept the ability to burn in specified locations.

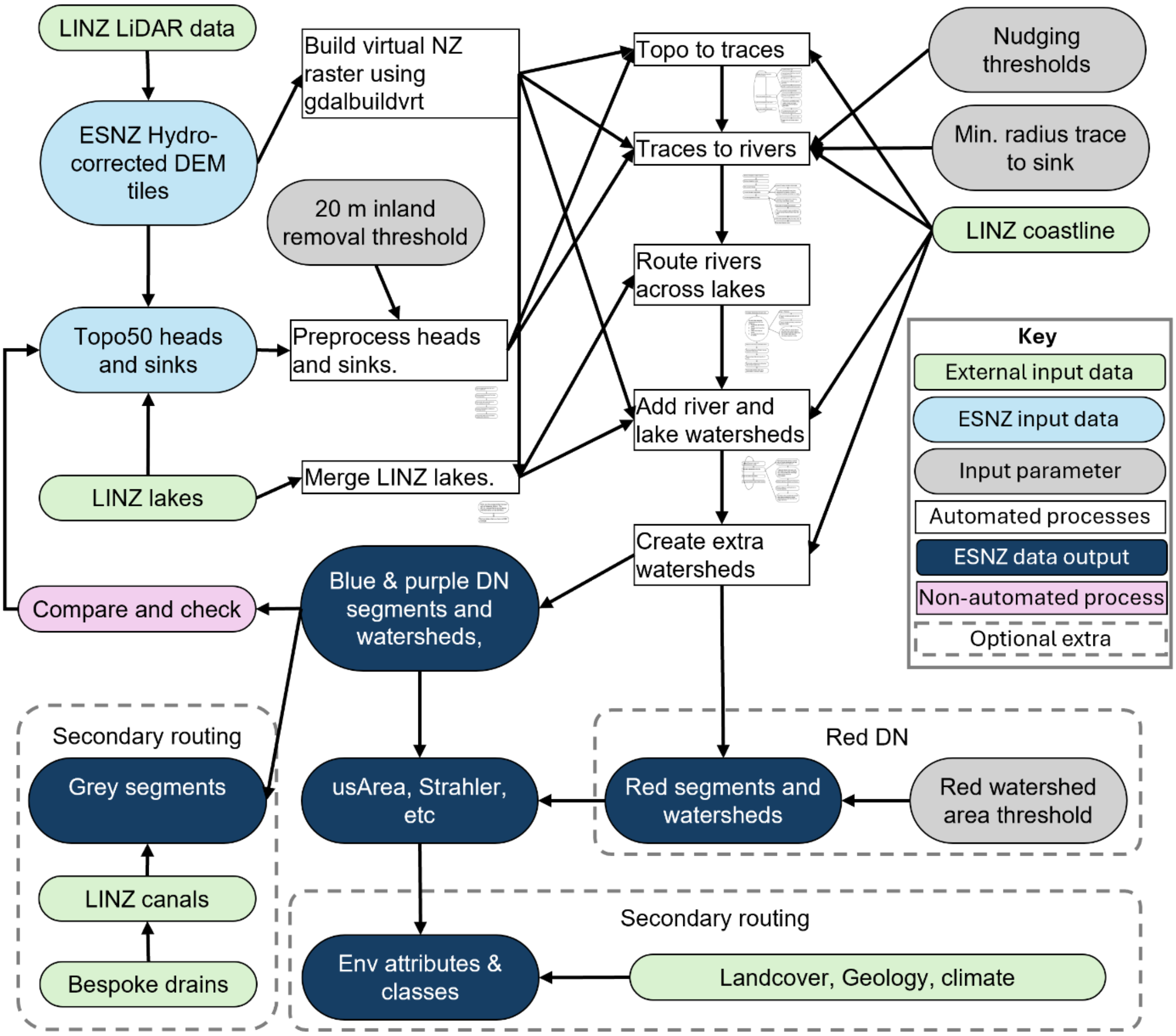


Figure 3-4: Overview of process for generating blue (river channel) and purple (overland flow downstream of river channels) parts of a DEM-derived DN. Smaller inset diagrams indicate further sub-component process shown in Figure 3-4 and Figure 3-5.

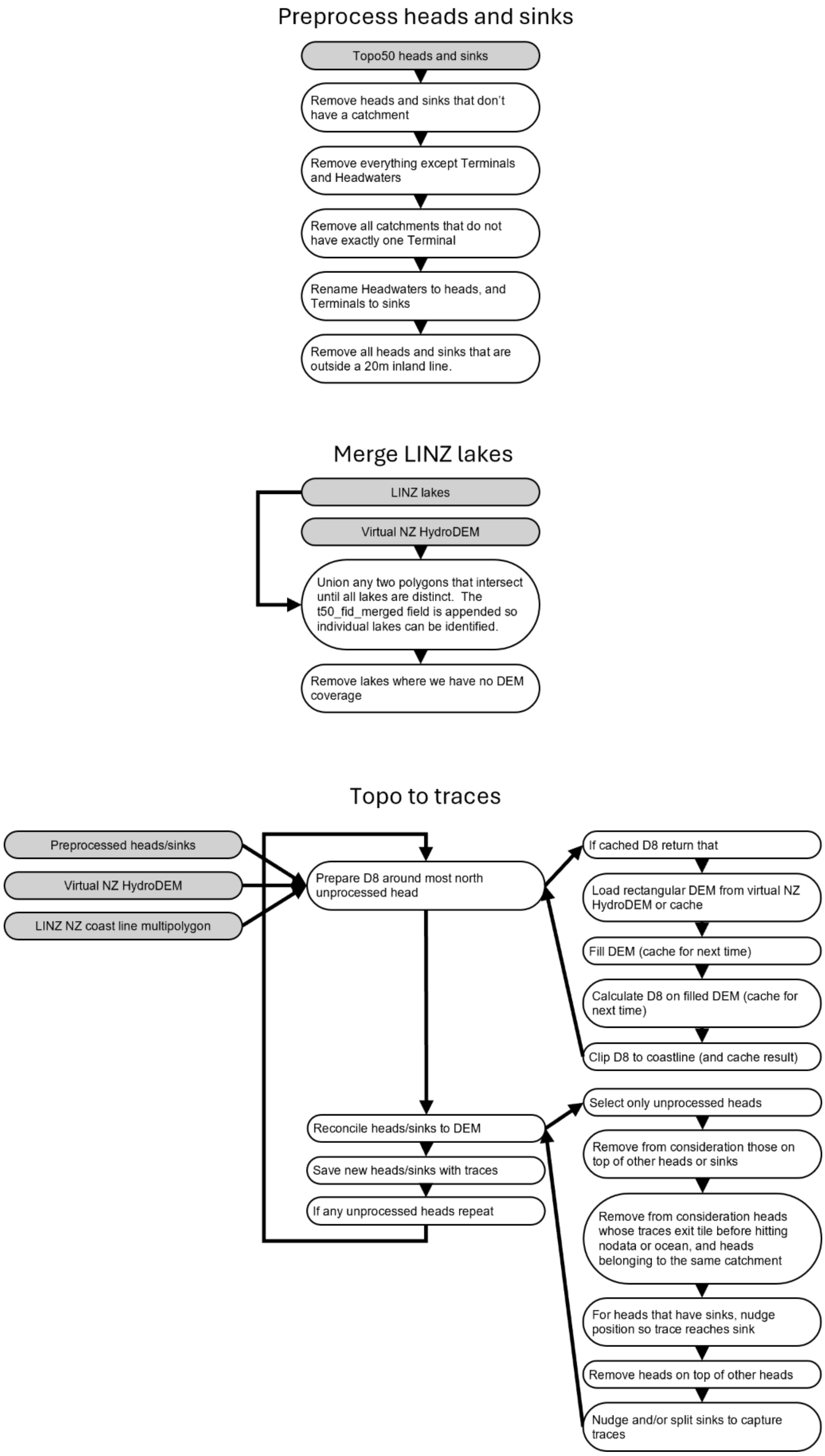


Figure 3-5: Details for generating a DEM-derived DN. Part 1 of 2. D8 is eight-directional flow index.

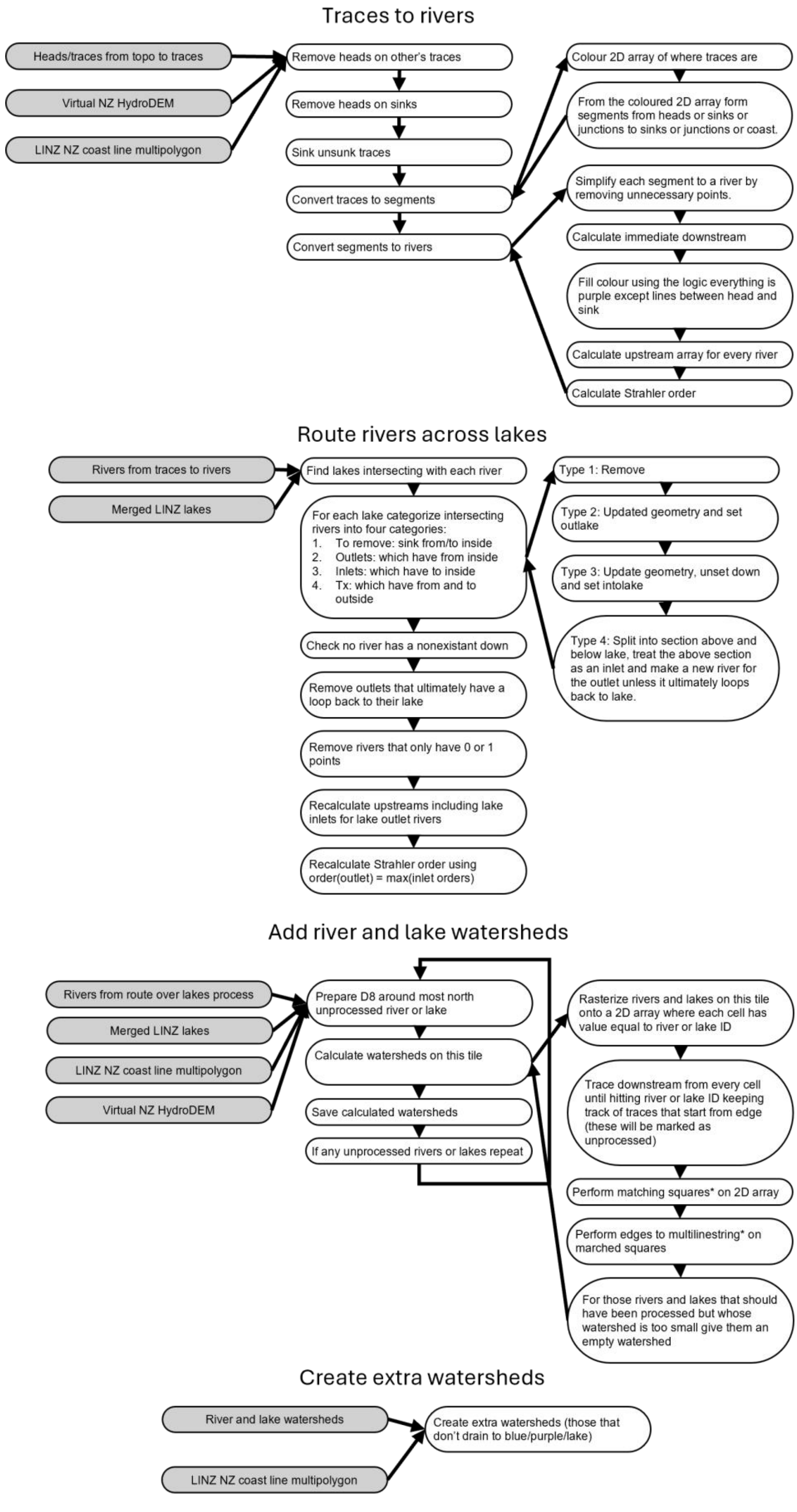


Figure 3-6: More details for generating a DEM-derived DN. Part 2 of 2.

Table 3-3: Methods for associating network segments with cartographically mapped sinkholes (after Booker et al. 2026). The cartographically mapped information is coloured grey: open circle is headwater; filled dot is sinkhole; line is river channel. The dashed circle has a specified radius (labelled the interception radius). The blue line is the network trace. The red cross is a prematurely terminated trace. The dashed blue line indicates the downstream trace had it not been prematurely terminated (i.e., a purple segment).

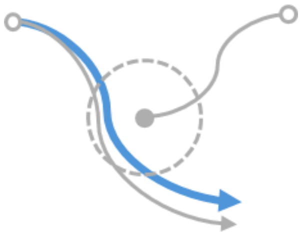
Picture	Description	Solution
	Traces whose cartographically mapped headwater was not associated with a downstream cartographically mapped sinkhole.	Traces not connected to mapped sinkholes were allowed to continue to the coast because they were not terminated regardless of proximity to any cartographically mapped sinkholes.
	There was only a single headwater associated with the sinkhole, but the downstream trace bypassed the sinkhole by more than the interception radius.	The trace was not prematurely terminated and continued to the coast.
	There was only a single headwater associated with the sinkhole, and the downstream trace came within the interception radius of the sinkhole.	The trace was terminated at its closest point to the sinkhole.
 	There were multiple headwaters associated with the sinkhole, the headwaters traced to a common node, and the downstream trace came within the interception radius of the sinkhole.	The trace was terminated at the closest point to the sinkhole downstream of the intersection of the traces.
	There were multiple headwaters associated with the sinkhole, the headwaters did not trace to a common node within the interception radius of the sinkhole.	More than one termination point was applied. A greedy algorithm was employed to minimise the proliferation of new termination points by grouping together traces.

Table 3-4: User inputs for the DEM-derived DN generation process. Value shows default values.

Parameter	Description	Value
Headwater nudging threshold	Allows head to move this far in an attempt to get downstream trace to get within sink interception radius of supposed sink.	20 cells
Sinkhole interception radius	Sinks can be moved this far to land on heads trace.	20 cells
Headwater/sinkhole near coast removal threshold	Removes all heads and sinks that are outside of a line that is a specified distance inland of the coast.	20 m
Maximum breach distance	How far to search for a breach point in DEM so a tunnel can be dug to allow water to flow to coast.	200 cells
Watershed connection proportion	Initially, watersheds can have knots (i.e., contain cells that only touch at a corner). Connection splits knots apart into four points, each point is moved this proportion of cell size along each side of the cell.	0.1

The process generated a DN that conformed to a set of rules.

1. A DN must contain segments and can contain lakes.
    a. All parts of all lakes and segments must be enclosed within, or intersect with, a coastline. Thus, lakes and segments (and therefore headwaters and sinkholes) must sit on land.
    b. Every segment and lake has an associated watershed defining the area of land draining directly into it.
    c. Every DEM grid cell must belong to a single watershed. Thus, watersheds are comprised of DEM cells.
    d. Some watersheds drain directly to the coast. Thus, a watershed does not have to be associated with a segment or a lake.
    e. Every segment and every lake must have a unique identifier (id) which is used to identify its associated watershed and to specify routing between segments and lakes.
    f. Headwaters, sinkholes, and coastlines are inputs to the DN. They can be extracted from the DN output, but they are not explicitly part of the DN.
2. Segments must be labelled with a type.
    a. Blue = surface flow pathways along rivers channels.
    b. Purple = overland flow pathways downstream of river channels that have disappeared into a sinkhole.

c. Grey = artificial flow pathways along secondary routing (see Section 3.3.2 for rules governing grey segments).

d. See Section 3.3.3 for explanation of red segments within a red DN.

3. Blue and purple segments must flow downhill.

   a. Every blue or purple segment must have a watershed.

   b. Blue and purple segments can join to form confluences but cannot split to form bifurcations.

   c. Every segment has a length but does not have an area. Thus, segments are lines. Segments are not polygons.

   d. Blue and purple segments cannot cross themselves or each other.

   e. Every segment has an upstream end and a downstream end.

   f. The upstream end of each blue segment must flow from either a headwater, a lake, or another blue segment.

   g. The downstream end of each blue or purple segment must flow into either the upstream end of another segment, a lake, or the coast.

   h. Sinkholes must exist at the downstream end of a blue segment and at the upstream ends of purple segments that are not connected to any other purple segments. Thus, sinkholes exist where a blue segment flows into a purple segment.

   i. Headwaters exist at the upstream ends of blue segments whose upstream end is not connected to a lake or any other segments.

   j. A segment whose downstream trace does not intercept any further segments is labelled a terminal segment.

4. Lakes must be polygons that are isolated from all other lakes.

   a. Segments cannot cross lake boundaries.

   b. Every lake has a positive area.

   c. A DEM cell contained entirely inside a lake boundary must belong to that lake's watershed. Thus, DEM cells that intersect a lake boundary do not to belong to that lake's watershed unless their flow direction points toward the lake (e.g. it can belong to the watershed of a nearby segment).

   d. A lake can have no, one, or multiple inflowing segments.

   e. A lake that has no inflowing segments can have no outflowing segments. Thus, a lake can be isolated from all segments.

   f. A lake that has at least one inflowing segment must have at least one outflowing segment.

g. A lake can have multiple outflowing segments. Thus, lakes with multiple outflows can trace downstream along segments to multiple terminal segments at the coast. Therefore, the presence of lakes with multiple outflows can cause segments upstream of the lake to be associated with multiple terminal segments.

5. Downstream traces can pass through blue segments, purple segments, and lakes, but must terminate at the coast.

   a. Breaching of topography is required to reach the coast from within inland depressions.

   b. No loops in routing can exist. Thus, all downstream traces should only pass by a segment or a lake once.

   c. All surface flow pathways for blues, purples, and lakes must trace to the coast.

Our DN generation process fused raster data (flow directions calculated on DEM cells) with vector data (lake and coast polygons). We used coast polygons to apply a land mask to DEM cells. All cells within or intercepting the coast were designated as land and therefore had surface flow directions and belonged to a watershed. We experimented with using lake polygons to apply a similar mask to distinguish between DEM cells inside or outside of lake polygons. We found rasterization of lakes to be unsatisfactory due to lake polygons being inconsistent with DEM topography because: a) DEM cells outside lake polygons could be lower than adjacent DEM cells inside lake polygons; and b) the elevations of the DEM cells inside lake polygons were not always constant or gradually decreasing towards the lake outlet. We therefore applied an alternative approach to explicitly incorporate lake polygons into our DEM-derived DN. Our approach involved initially generating a DN without lakes, then removing all segments within lake polygons, and finally removing segments that caused loops in routing because they previously flowed between several lake boundaries. One consequence of the approach was that it allowed lakes to have multiple outflowing segments.

3.3.2 Secondary routing

We extended our process to optionally add grey segments to a blue/purple/lake DN to represent flow pathways along artificially engineered channels. Grey segments and their secondary routing effects on routing can be superimposed on a blue/purple DN as described in Appendix C. Superposition of grey segments does not impact the alignment, flow directions, or watersheds of blue/purple segments. Secondary routing does impact the routing of a blue/purple DN by influencing which blue/purple segments are touched when tracing either upstream or downstream. We allowed a high level of flexibility in the rules associated with grey segments, rules for which follow on from the rules stated in Section 3.3.1.

6. Grey segments and their effects on routing must be superimposed onto a blue/purple DN.

   a. Each grey segment has a length but does not have an area. Thus, grey segments are lines not polygons.

   b. Grey segments have no watersheds.

c. Grey segments can cross any other object (blue segment, purple segment, lake, coast) without being connected to it.

7. Grey segments must flow in one direction.

   a. Flow direction and line positioning of grey segments are independent of slope. Thus, they can flow uphill or downhill.

   b. The upstream end of grey segment can be isolated or flow from a blue segment, a purple segment, a lake, or one or more other grey segments.

   c. The downstream end of a grey segment can be isolated or flow toward a blue segment, a purple segment, a lake, the coast, or one or more other grey segments.

   d. An isolated downstream end of a grey segment can terminate inland, and an upstream end of a grey segment can exist in isolation. Furthermore, grey segments can join to form confluences or bifurcations.

The high level of flexibility within the rules associated with grey segments allow secondary routing to represent several landscape phenomena.

- Diversions between river catchments.
- Diversions within river catchments.
- Abstractions from rivers that supply water to one or more locations on land or at the coast.
- Drains or pipes that are either connected or unconnected to the river network.

Although grey segments were designed to represent artificial flow pathways, they can also be used to represent natural phenomena that result from bifurcations.

- Anastomosing or braided rivers channels that flow around islands.
- Deltas that split before entering the coast in multiple locations (e.g., Lower Clutha River).

We populated the information required to insert South Island canals from the LINZ Canal Centrelines and Canal Polygons data layers described in Appendix A. We did not insert canals for the North Island. We did not insert any grey segments using information in the LINZ Drain Centrelines data layer.

3.3.3 Red DN

We devised a process for generating a red DN that was consistent with an existing blue/purple DN but represented higher resolution information describing surface flow pathways, including overland flow pathways upstream or river channels. Segments of a red DN followed all rules for blue segments stated in Section 3.3.1 except for the effects of headwaters and sinkholes. Our process for generating a red DN had one input; a threshold for red watershed areas. If the threshold for red watershed area was the DEM grid resolution, then the red DN would represent the finest possible level of detail. A red DN is consistent with a blue/purple DN if they are generated from the same DEM. All blue/purple segments can be mapped to one or more red segments.

3.4 Outputs

Our process to generate a DEM-derived DN produced two main output files. The first output file described segment geometries and attributes (Table 3-5). The second output file described watersheds geometries and their id's. Our process to generate a Topo50-derived DN produced an output file that followed the format set out in Table 3-5. The Topo50-derived DN did not have watersheds. All output files were GeoPackage (.gpkg) files, which is an open, non-proprietary file format for Geographic Information Systems (GIS) defined by the Open Geospatial Consortium (OGC). GeoPackage files allow basic geometric variables such as segment length and watershed area to be easily calculated from object geometries.

Table 3-5: Attributes for each segment within our DEM-derived DN. Value shows default values.

Attribute	Description	Notes
id	A unique identifier for the segment.	A number (six digits in our DN).
type	Distinguishes between blue, purple, grey, and red segments.	"blue", "purple", "grey" or "red".
down	The segment id (or id's) that the local segment flows into.	Empty for none ("[]"). May contain multiple id's for bifurcations ("[111111,222222]").
hid	Identifier for headwater that the segment flows away from.	-1 for segments that are not immediately downstream of a headwater.
sid	Identifier for sinkhole that the segment flows into.	-1 for segments that are not immediately upstream of a sinkhole.
Intolake	Identifier for lake that the segment flows into.	A number (less than five digits in our DN).
outlake	Identifier for lake that the segment flows away from.	A number (less than five digits in our DN).
Terms	The terminal segment(s) at the downstream end(s) of a trace from the local segment.	Cannot be empty. May contain multiple id's for bifurcations. Is amended by secondary routing.
order	Strahler stream order.	Positive values for blue segments. Negative values for purple segments.

We created a Python script that allowed interactive upstream and downstream tracing from a selected segment in QGIS ("traceupanddown.py"). See Appendix C for further details. We also devised an R-shiny app which allowed interactive upstream and downstream tracing from a selected segment. We used the app for internal purposes such as checking the routing of our Topo50-derived and DEM-derived DNs.

3.5 Validation

3.5.1 Checks for functional correctness

Booker et al. (2024) suggested several checks could be applied to verify that a DN is behaving correctly after it has been produced. The checks suggested by Booker et al. (2024) assumed that the network does not represent bifurcations and used a different method to express routing than shown in Table 3-5. We reviewed the checks suggested by Booker et al. (2024). We then devised and applied several tests to be applied to a blue/purple DN before secondary routing is imposed.

1. Blue and purple segments.
   a. All segments should have length greater than zero.
   b. No segments should cross themselves, other segments, lakes, or the coast.
   c. All segments should have a watershed.
   d. All segments should route downstream to a terminal segment (i.e., no indefinite loops should occur).
   e. All segments should have one downstream terminal segment unless they are downstream of a multi-outflow lake that routes to the coast at more than one location.
2. Lakes
   a. All lakes that are connected to a segment should have at least one outflowing blue or purple segment.
   b. All lakes should have a watershed.
3. Terminal segments
   a. All terminal segments should plot on the coast, or at a lake with no outflowing segments.
4. Watersheds
   a. All watershed areas should not be negative.
   b. All watershed areas should be divisible by the resolution of the DEM grid.
   c. The sum of all watershed areas should be equal to the area of the DEM.
   d. The sum of upstream cumulated areas for all terminal segments and all watersheds that drain directly to the coast (are not associated with a segment) should be equal to the area of the DEM.
   e. All watersheds that are not associated with a segment should intersect with the coast.

3.5.2 Checks for representative correctness and suggestions for secondary routing

We compared parts of our Topo50-derived and DEM-derived networks that we expected to be particularly difficult to derive purely from a DEM (e.g., lowland areas with many artificial channels, alpine areas with many springs and braided rivers).

We arranged for a pilot version of our automatically DN to be uploaded onto the MfE data service. We arranged for access to the uploaded data to be provided to all members of the project oversight group via their koordinates accounts. Alongside the pilot DN, we posted descriptions, instructions for creating secondary routing, some code that facilitated DN routing, and an example of how to provide feedback (see Appendix C). Alongside the pilot DN, we also posted instructions for adding grey segments onto a DN (see Appendix C) to a selection of DN users (see Appendix C). We created an email address to allow users to provide feedback about the pilot DN (river.network@earthsciences.nz).

4 Results

4.1 Topo50-derived digital network

The main purpose for generating a Topo50-derived DN was to extract headwater locations to subsequently generate a DEM-derived DN. Nevertheless, we inspect the Topo50-derived DN to assess its validity. The Topo50-derived DN conformed to our expectations about river routing in most locations because it matched dendritic connections between rivers depicted in Topo50 maps (e.g., Figure 4-1). However, the Topo50-derived DN did not match with our expectations about river routing in a few locations. These results were mainly associated with blue lines that contained complex sets of loops because they represented braided rivers (e.g., the larger braided rivers in Figure 4-2) or because blue lines on Topo50 maps associated with rivers and/or drains repeatedly connected to each other (e.g., some smaller rivers with angular pathways in the centre of Figure 4-2).

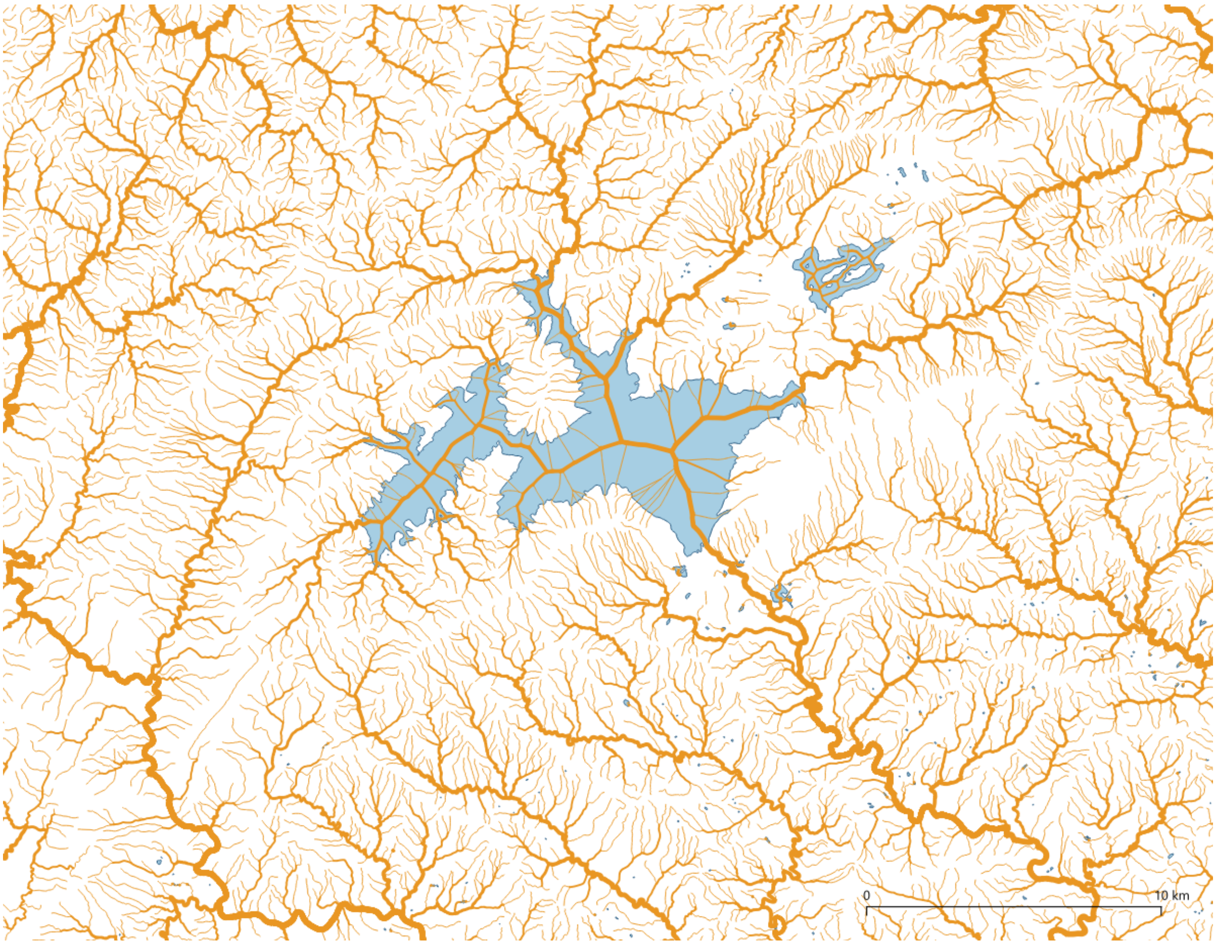


Figure 4-1: Topo50-derived network around Lake Waikaremoana near the border between Hawke's Bay and Gisborne regions. Line width is proportional to Strahler order.

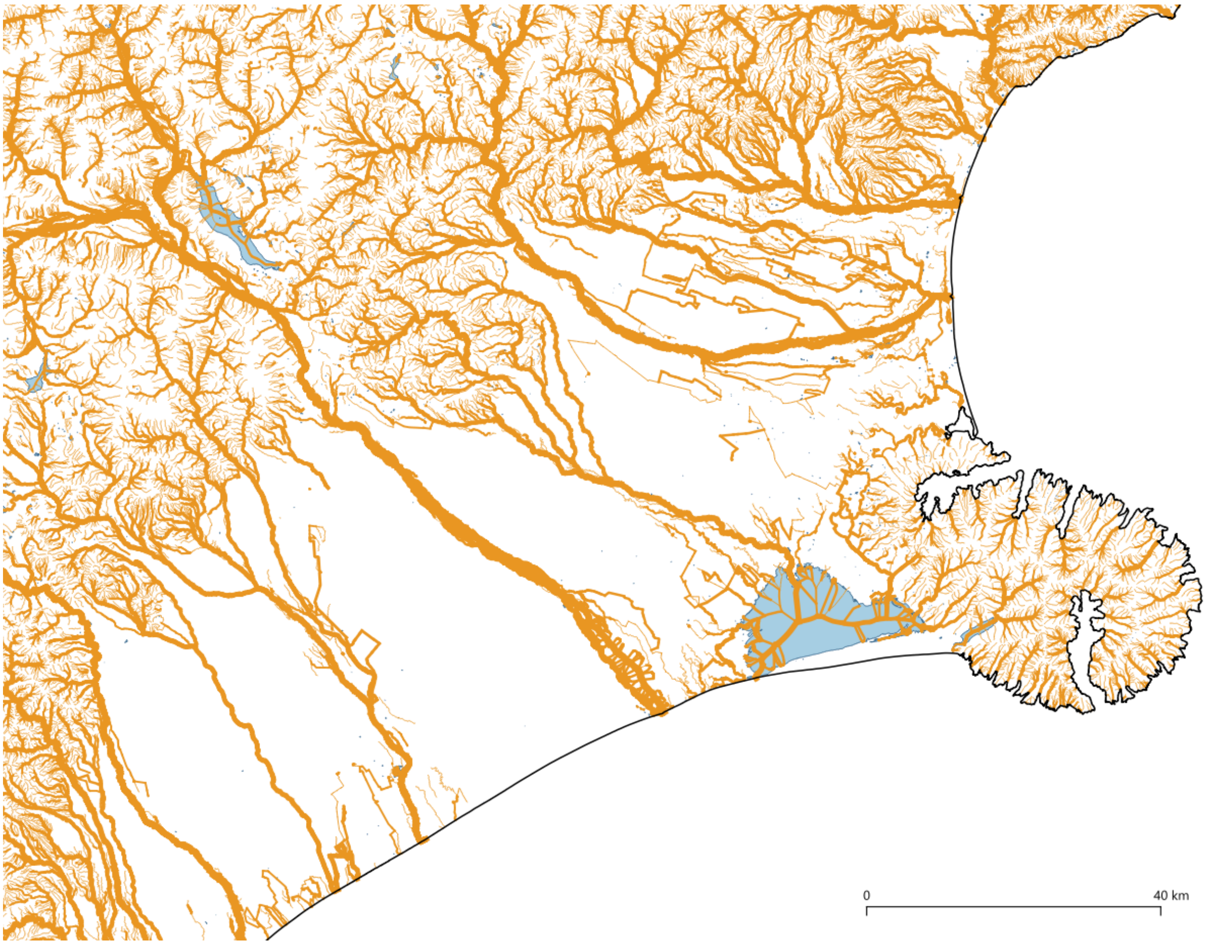


Figure 4-2: Topo50-derived network around Banks Peninsula in the Canterbury region. Line width is proportional to Strahler order.

The Topo50-derived DN contained approximately 945,000 segments. Most segments (87%) in the Topo50-derived DN were derived from objects in the LINZ river data layers. Less than 1% of segments in the Topo50-derived DN were derived from lagoons, rapids, or waterfalls. Approximately 1% of segments were derived from LINZ drains or canals that caused river segments to connect to the coast. Some segments were not directly derived from a single Topo50 data layer (4%). Some segments were derived from skeletonization of lakes (8%). Approximately 2.7% of segments flowed out of lakes and 3.3% of segments flowed into lakes. Approximately 2.6% of segments were terminal segments.

The Topo50-derived DN produced many more lower Strahler order segments than higher Strahler order segments (Figure 4-3). Strahler order could not be calculated for approximately 10% of Topo50-derived DN segments because they experienced loops within their downstream routing.

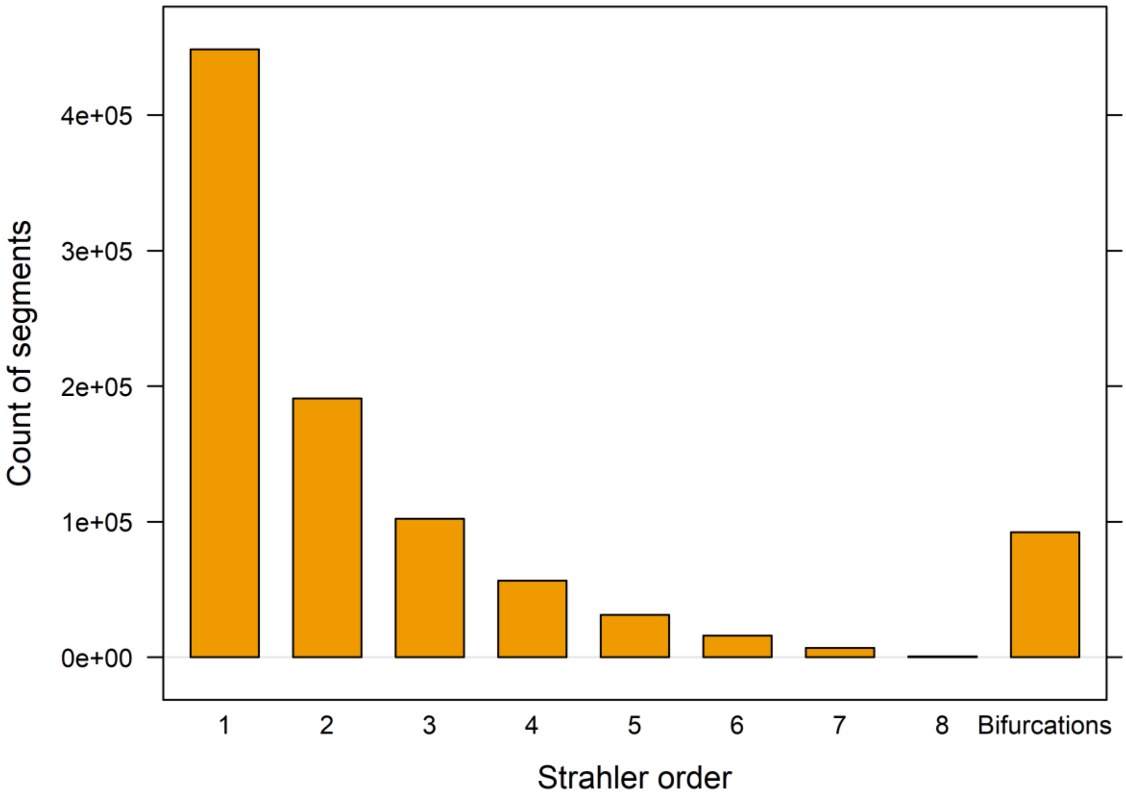


Figure 4-3: Count of segments in the Topo50-derived network by Strahler order. Strahler order could not be calculated for bifurcating segments.

The Topo50-derived DN produced approximately 362,000 headwater locations. The headwaters were not uniformly spread across the country (Figure 4-4). Headwaters were absent from areas with large lakes, permanent snow or ice, and areas with more porous geology (e.g., Canterbury Plains). Headwaters were denser in steeper areas and denser in areas where geological/hydrological conditions produced more springs than elsewhere. Square patterns in headwater density in a few locations on the North island indicated that standard methods for mapping headwaters may not have been applied across all Topo50 map tiles. Nevertheless, spatial patterns in headwater density demonstrated that it would be inappropriate to apply a global area threshold to initiate DN segments to represent river channels.

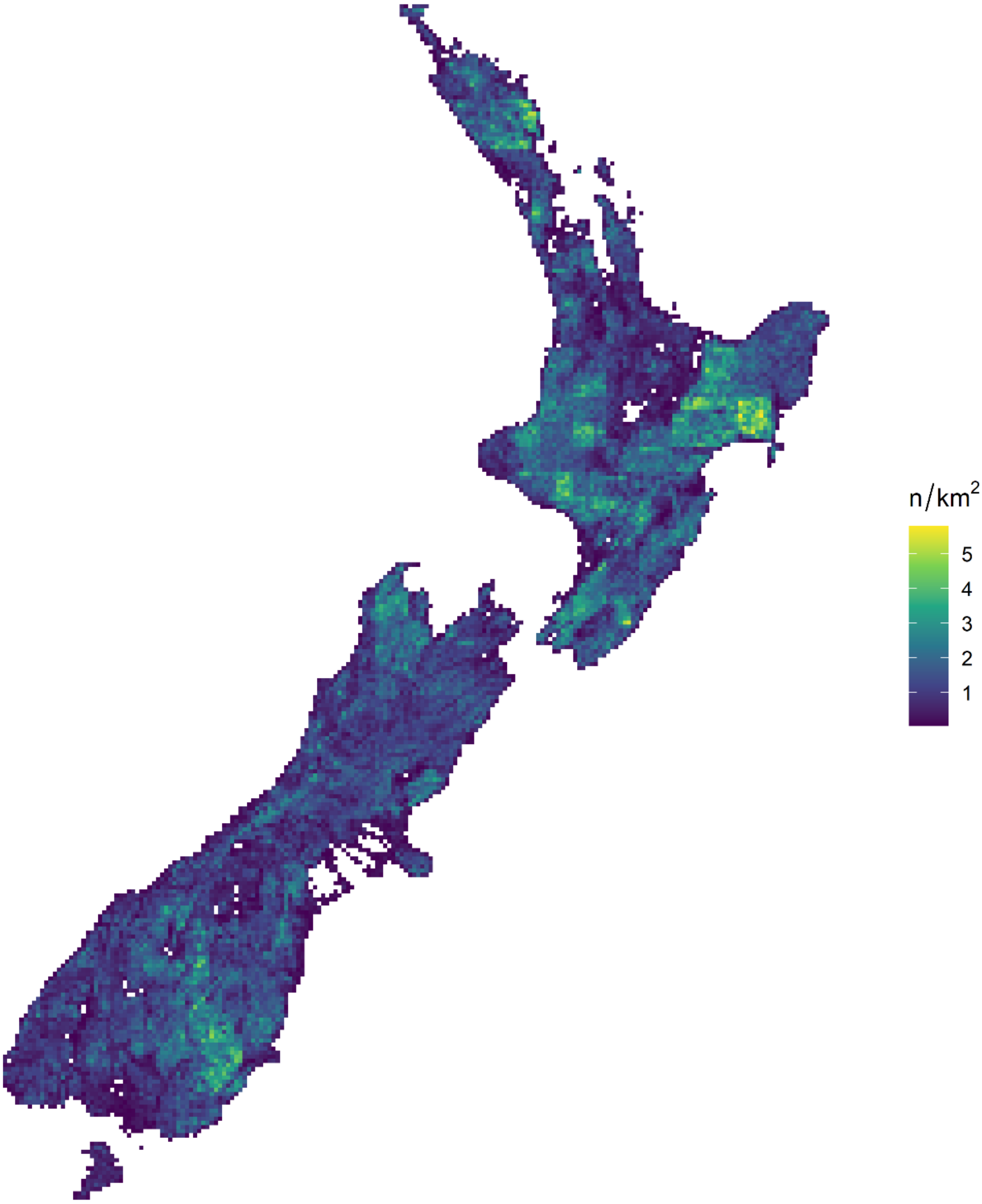


Figure 4-4: Density of Topo50 river headwaters in 5 km² cells across the country. Transparent cells contain no headwaters.

The Topo50-derived DN produced approximately 23,600 terminal locations. Terminal locations are sinkholes or locations where a river terminates at the coast. Intricate areas of coastline were associated with more dense terminal locations (e.g., Northland coastline). The inland sinkholes were not uniformly spread across the country (Figure 4-5). Many 5 km² cells contained no sinkholes (Figure 4-5). Sinkholes were denser in some areas, which was possibly related to proximity to porous geology (e.g., central North Island), upstream ends of alluvial fans in alpine settings (e.g., central South Island), braided rivers (e.g., inland South Island), and wetlands (e.g., various flat areas), or coastal lagoons/hapua (e.g., east coast South Island).

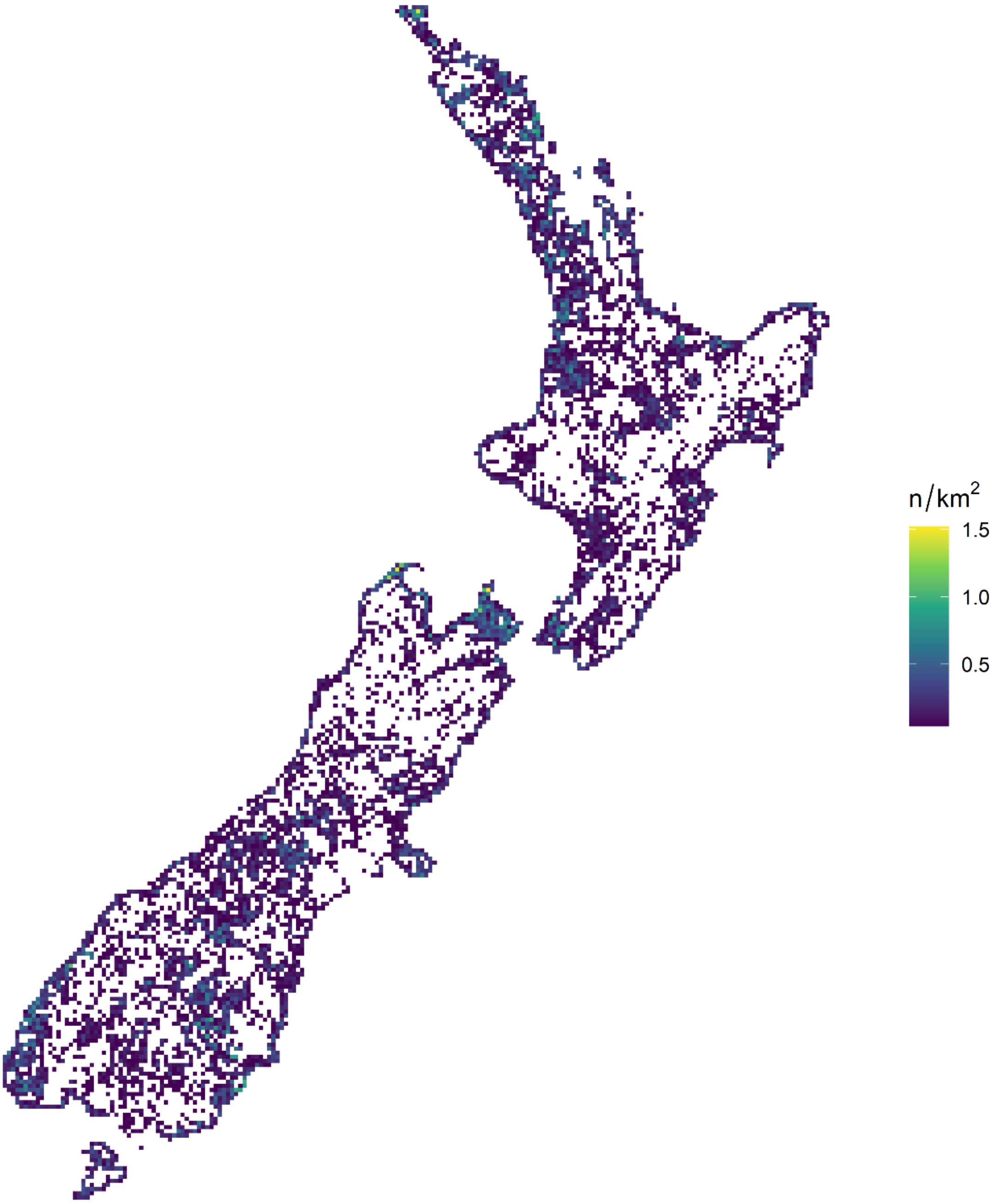


Figure 4-5: Density of Topo50 river terminal locations in 5 km^2 cells across the country. Terminal locations are sinkholes or locations where a river terminates at the coast. Transparent cells contain no terminal locations.

4.2 LiDAR DEM-derived digital network

4.2.1 Segments

Our DEM-derived DN was generated using headwater locations extracted from our Topo50-derived DN. Approximately 360,000 segments within the DEM-derived DN were headwater segments. Thus, approximately 2,000 headwater nodes generated from the Topo50-derived DN did not eventuate to produce segments within the DEM-derived DN. Possible reasons for Topo50-derived headwaters not producing segments include being too near to the coast, plotting outside of the DEM, and having an upstream-generated segment passing exactly through their DEM cell.

Our DEM-derived DN contained approximately 740,000 segments. Thus, the DEM-derived DN contained around 20% fewer segments than the Topo50-derived DN. The Topo50-derived DN contained more segments because it included segments within lakes (Figure 4-1) and segments that formed loops (Figure 4-2). Approximately 98% of DEM-derived DN segments were blue and 2% were purple which allowed representation of sinkholes in areas such as the Canterbury Plains (Figure 4-6) and where steep alpine rivers disappeared as they entered flatter valley floors (e.g., Figure 4-7).

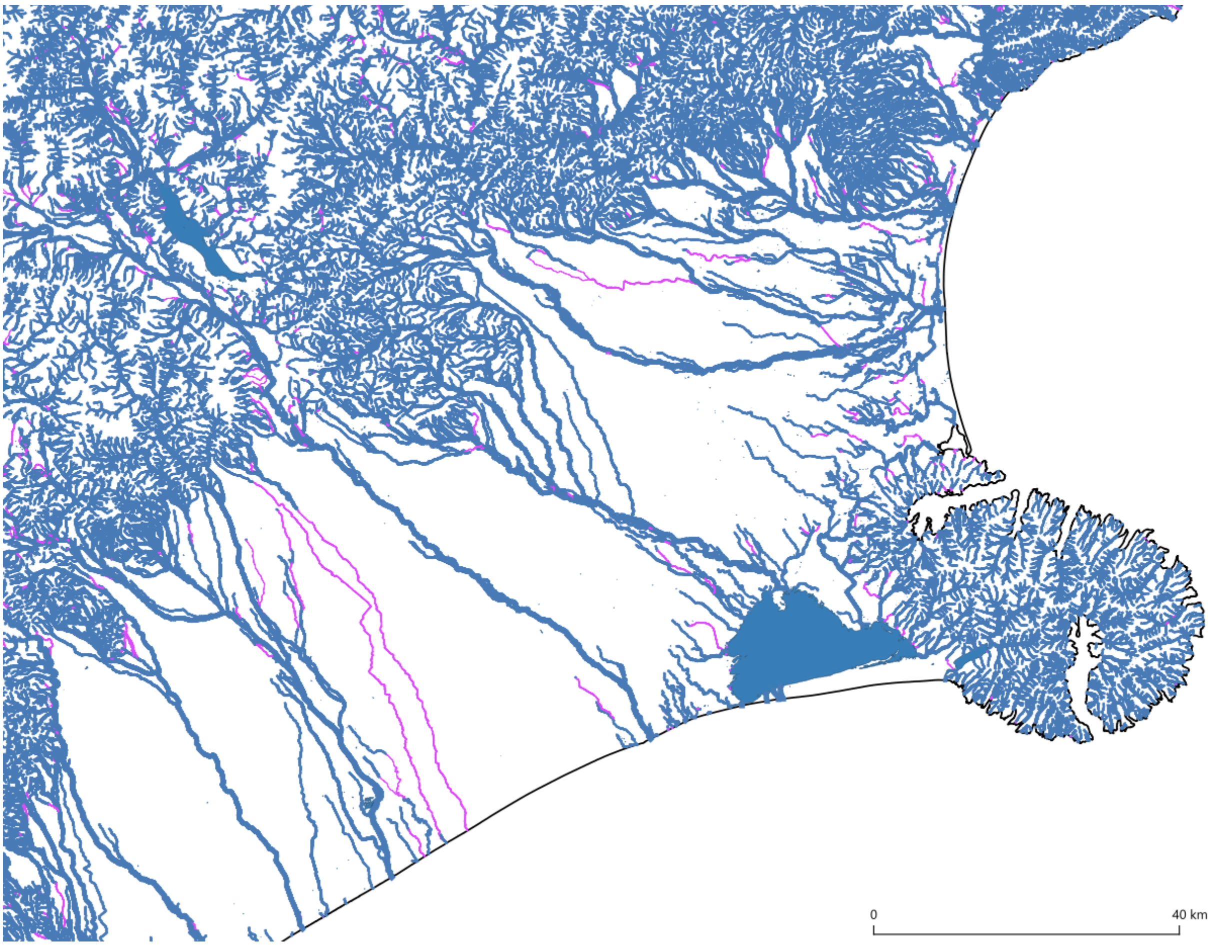


Figure 4-6: DEM-derived network around Banks Peninsula in the Canterbury region. Line width is proportional to Strahler order. Purple lines are downstream of sinkholes.

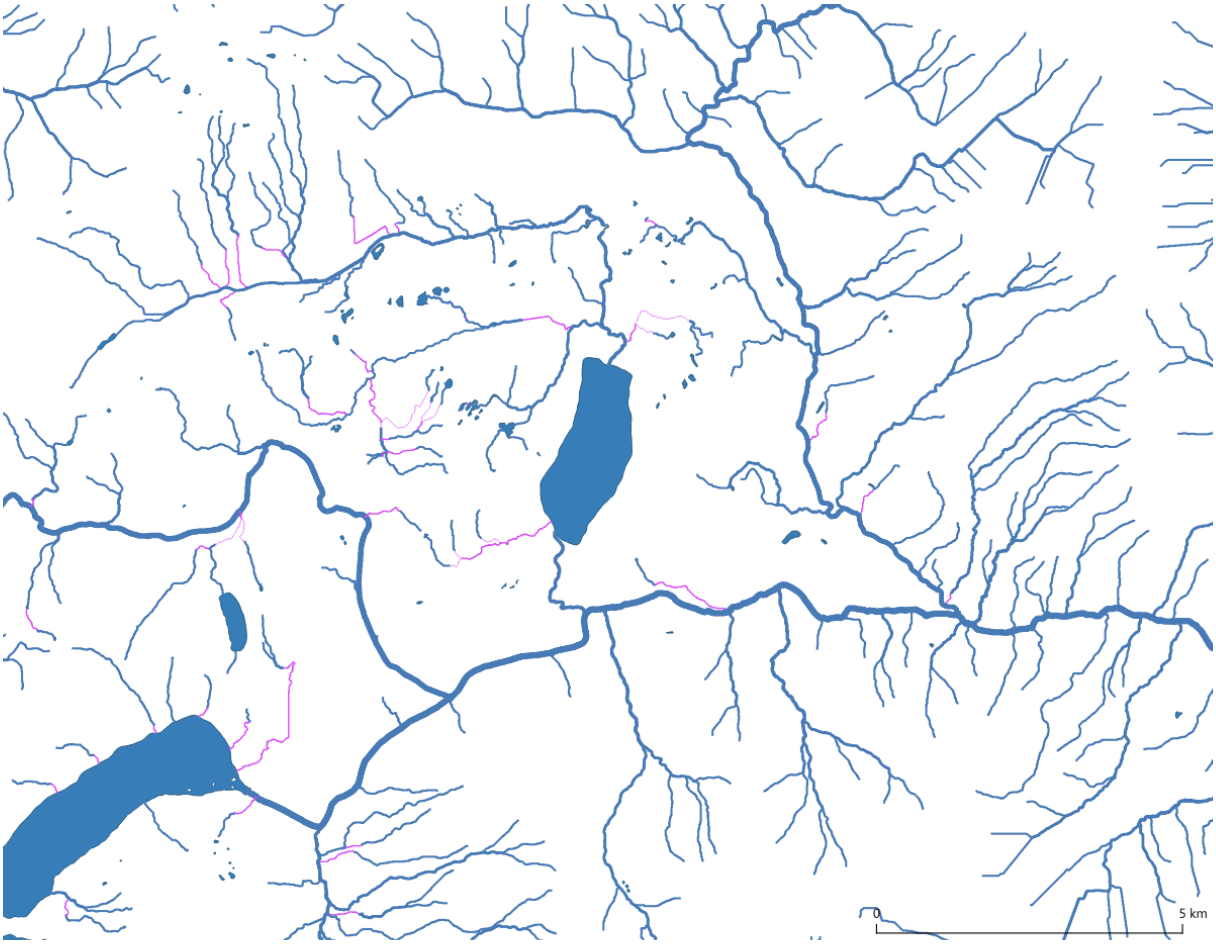


Figure 4-7: DEM-derived network around Lake Hayes in the Otago region. Line width is proportional to Strahler order. Purple lines are downstream of sinkholes.

All DEM-derived DN segments had lengths greater than zero. Median length was 617 m. Mean length was 757 m. Maximum length, which was for a purple segment, was 70,291 m. Only 4% of segments terminated at the coast. Only 1% of segments flowed into lakes. Fewer segments flowed out of lakes than flowed into lakes.

The number of segments belonging to each Strahler order within the DEM-derived DN (Figure 4-8) matched well with those found for the Topo50-derived DN (Figure 4-3). Similar Strahler orders for the DEM-derived DN and Topo50-derived DN was expected because the DEM-derived DN was generated using headwaters extracted from the Topo50-derived DN.

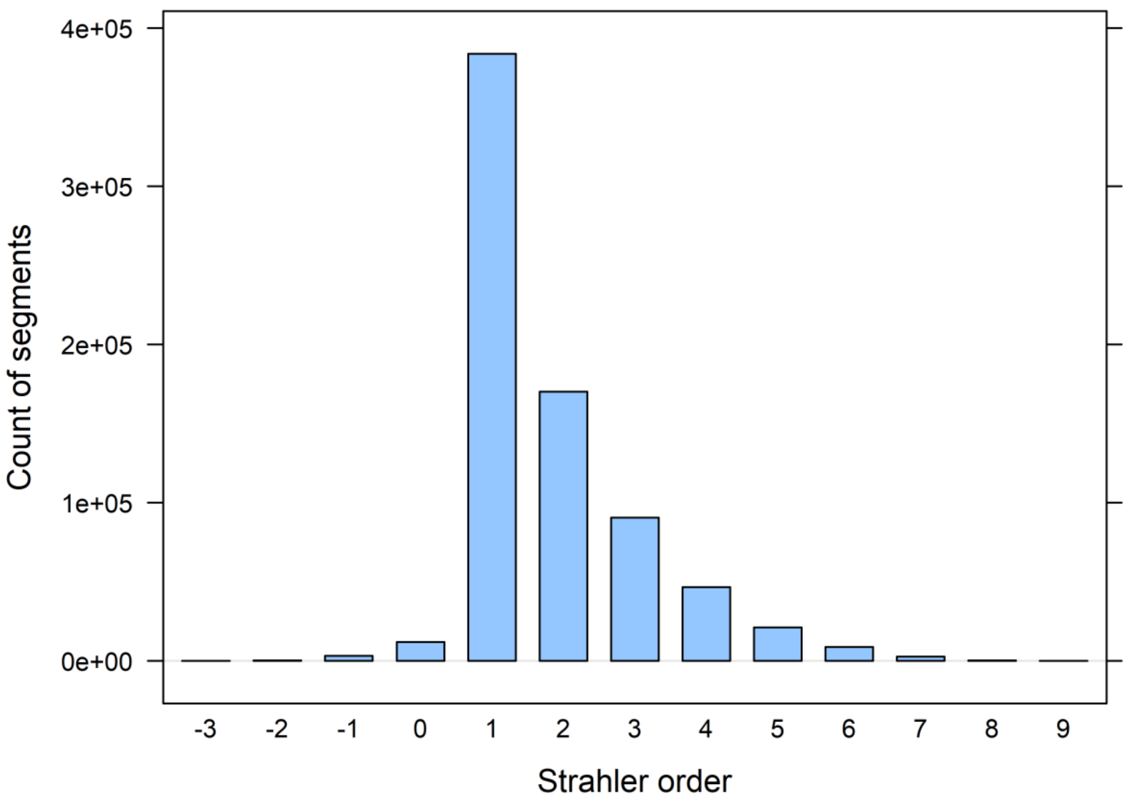


Figure 4-8: Count of segments in the DEM-derived network by Strahler order. Negative Strahler orders denote overland flow pathways (purple segments).

4.2.2 Watersheds

Every blue or purple segment had a watershed. Every lake had a watershed. Thus approximately 740,000 watersheds had segments, and 61,000 watersheds had lakes. A further 2,974,000 watersheds had no segments or lakes because their flow direction led directly to the coast (Figure 4-9).

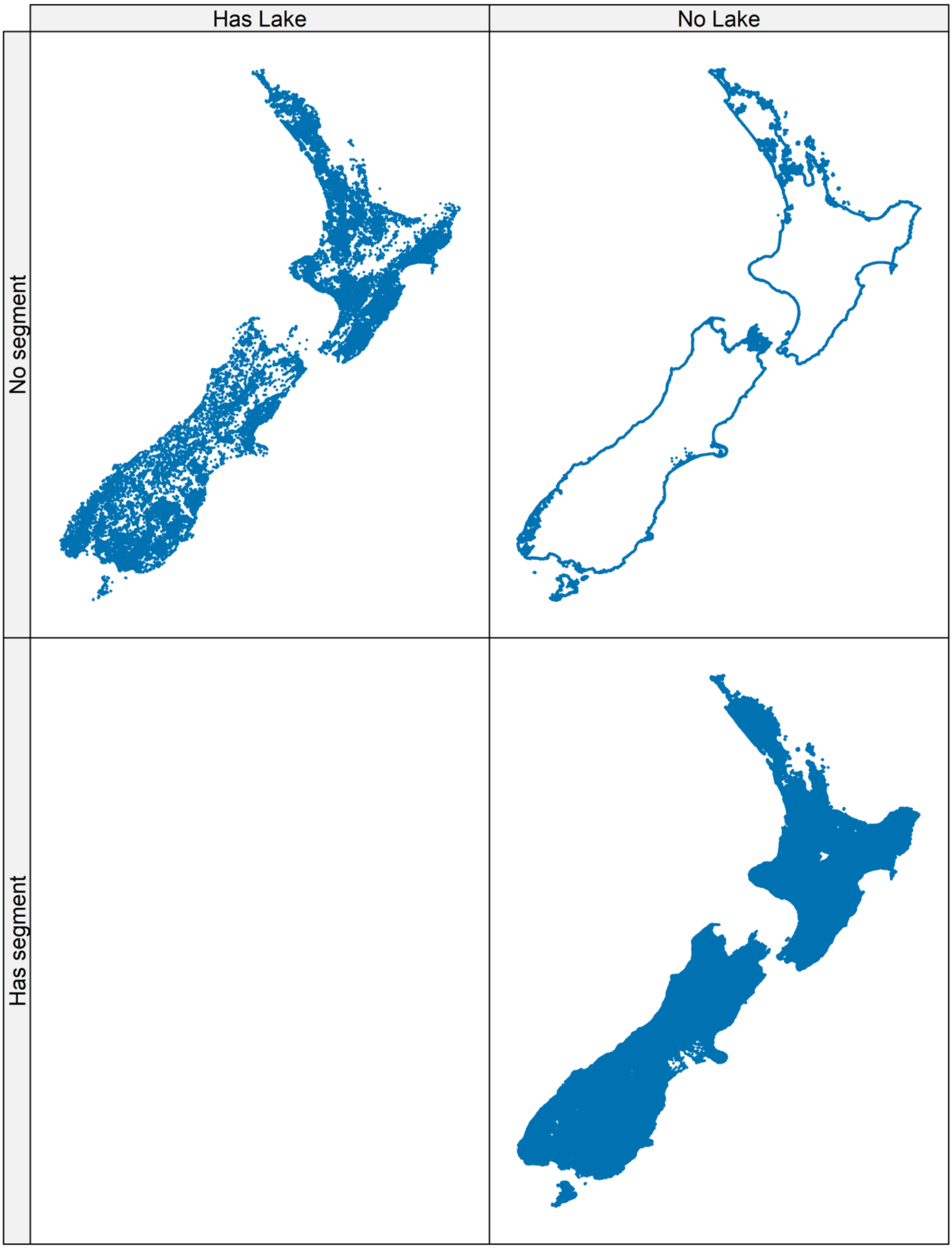


Figure 4-9: Locations of watershed centroids and their relationship to segments and lakes.

Watersheds flowing directly to the coast tended to have much smaller areas than watersheds containing segments or lakes (Figure 4-10). Watersheds containing lakes had a wide range of areas. The largest watershed was 685 km^2 and was associated with Lake Taupō (Taupōmoana), which has an area of 613 km^2. Lake watersheds could be larger than their associated lake because some land drained directly into the lake. Lake watersheds could also be smaller than their associated lake because our rules dictated that cells crossing lake boundaries could belong to watersheds associated with nearby segments (or other lakes or flow directly to the coast) as determined by their flow direction. Some watersheds containing segments had very small areas, but most tended to be between 0.1 and 1 km^2. Median area for watersheds containing segments was 0.19 km^2.

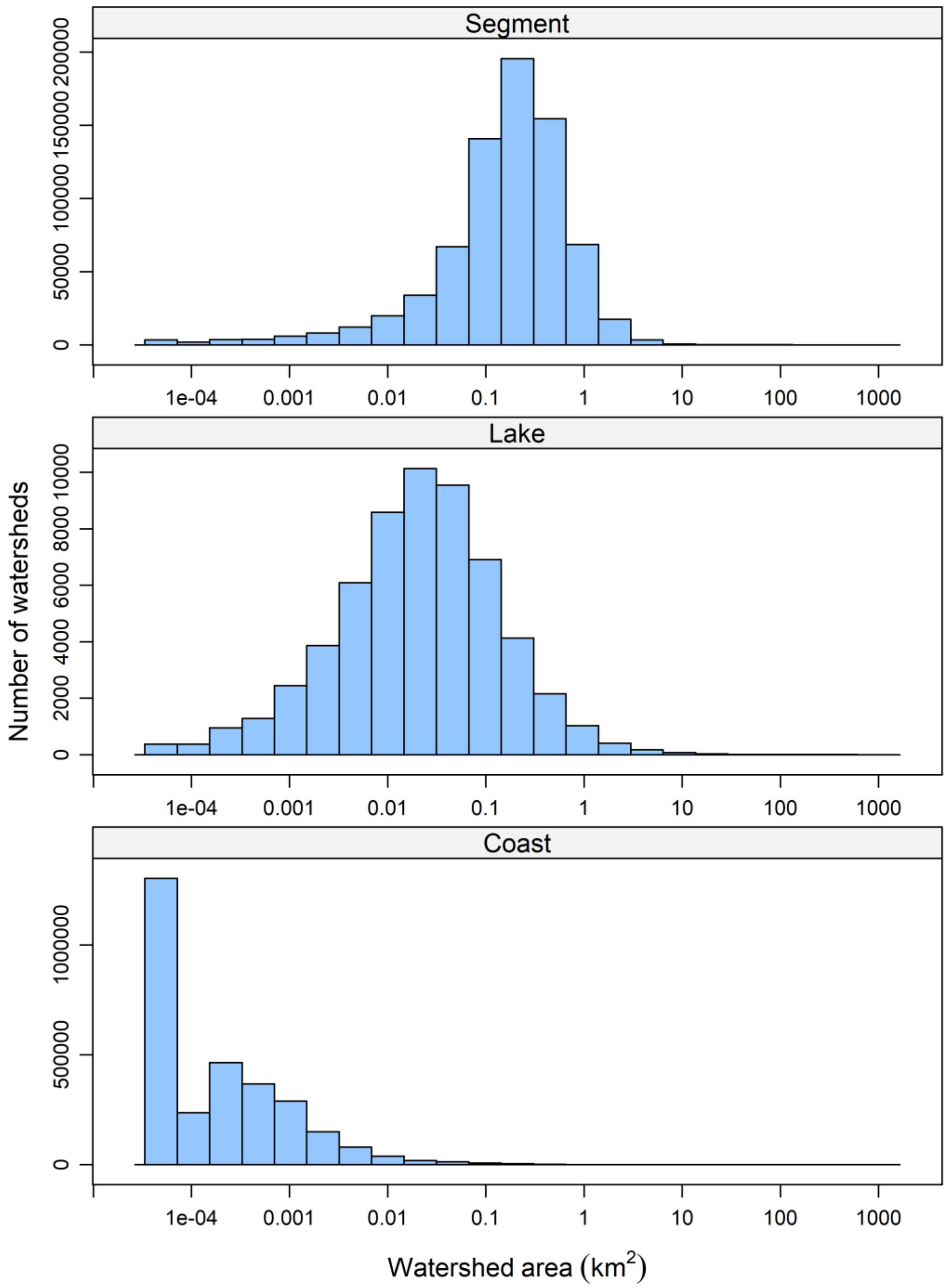


Figure 4-10: Areas for watershed containing blue or purple segments, containing lakes, or flowing directly to the coast.

4.2.3 Interaction with lakes

Our DEM-derived DN explicitly represents all lakes in the LINZ lake database. Of approximately 61,000 lakes in the LINZ database, 24,000 were connected to the DN (Figure 4-11). Lakes connected to the DN tended to be larger than those not connected, but some large lakes were not connected to the DN. For example, Lake Rotokawau and Lake Waikaramu in Northland were not connected to DN segments. Approximately 21,000 lakes had one inflowing segment and one outflowing segment. Approximately 200 lakes had no inflowing segments. Approximately 260 lakes had multiple outflowing segments, but many of their outflowing segments converged within a short distance of leaving the lake. The maximum number of outflowing segments from a lake was four. Approximately 96% of segments flowing into lakes were blue while 4% were purple.

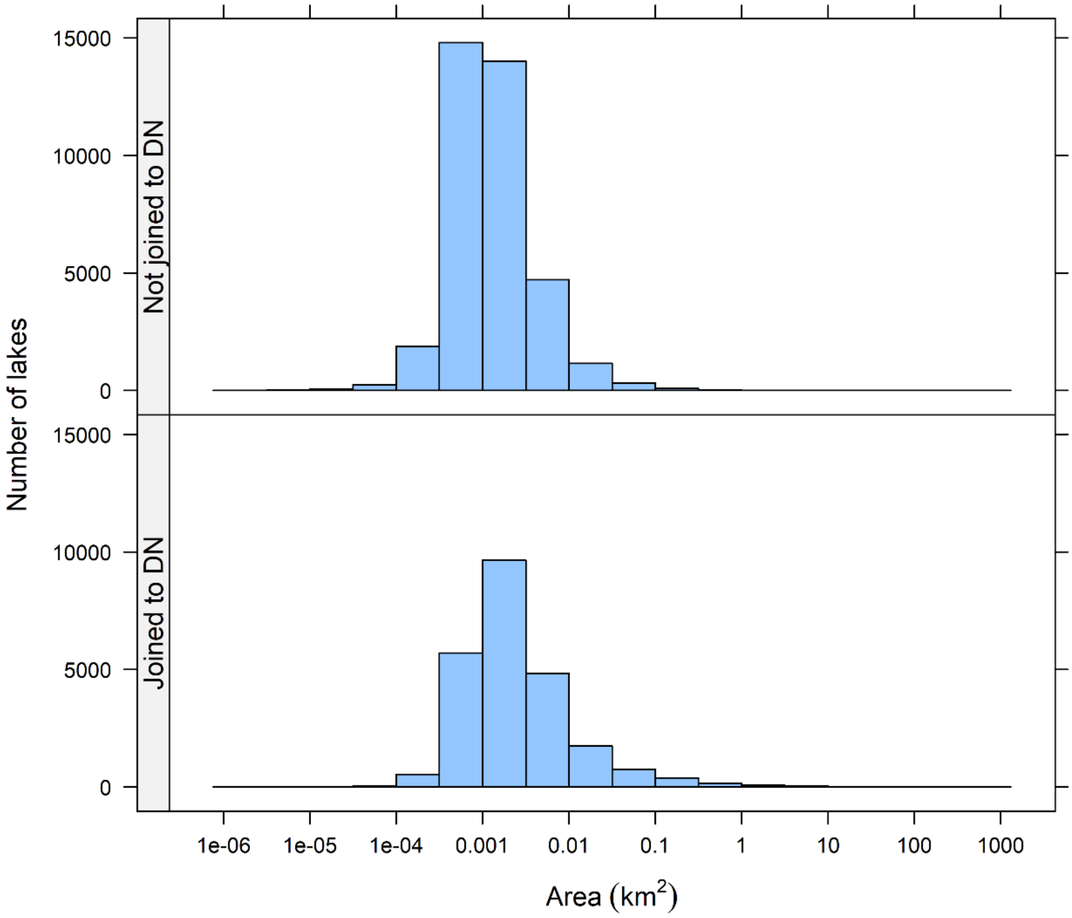


Figure 4-11: Count of lakes of different lake areas that were joined to the network or not joined to the network. Lakes were not joined to the network because they were isolated away from any river channel segments.

We found 12 lakes that had no outflowing segments within a pilot iteration of our DEM-derived DN. Further inspection indicated that the outflowing segments had been removed from ten of these lakes by our algorithm which conformed to our rules about exclusion of loops when downstream routing. Two lakes had no outflowing segments because they intersected with the coast (Lake Widgeon, Southland and an unnamed lake polygon that borders Kaipara Harbour).

Table 4-1: Number of segments flowing into and out of lakes that had no outflowing segments.

Lake identifier	In blue	In purple	Out	Touches Coast
2	1	1	0	FALSE
6	2	0	0	FALSE
2235	5	1	0	TRUE
2912	1	1	0	FALSE
2922	1	0	0	FALSE
10834	2	0	0	FALSE
10835	1	0	0	FALSE
12618	0	1	0	FALSE
21518	2	0	0	FALSE
21981	1	0	0	FALSE
53231	1	0	0	FALSE
53359	1	0	0	TRUE

4.2.4 Secondary routing

We used LINZ Topo50 canal data for the South Island to incorporate secondary routing onto our DEM-derived DN. Grey segments were added that crossed blue and purple segments (Figure 4-12). Our secondary routing included the effects of the Rangitata Diversion Race, Central Plains Water, and the Manapouri Power Scheme. Secondary routing therefore demonstrated that the effects of engineered diversion to river flows can be incorporated into a DN.

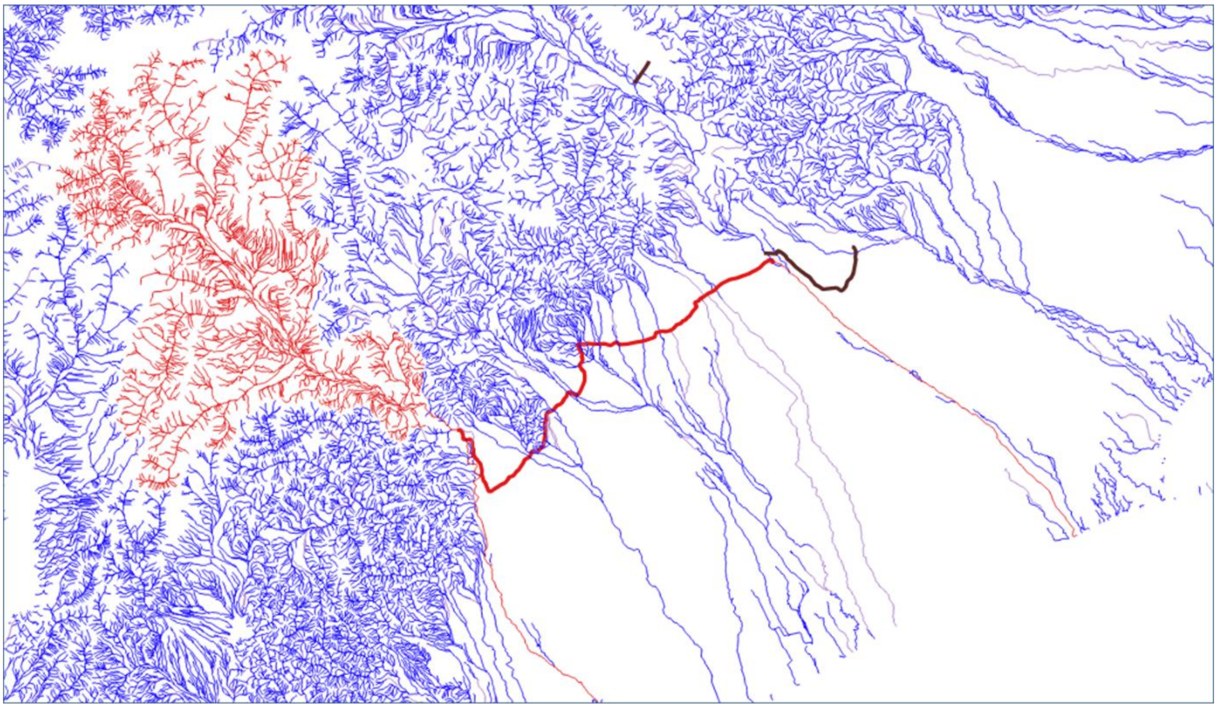


Figure 4-12: Example of secondary routing. Red lines show segments upstream and downstream of the Rangitata Diversion Race in central Canterbury. Purple segments are downstream of sinkholes.

We demonstrated that secondary routing could be added from user supplied data. A member of our oversight group supplied a GIS file (.gpkg) representing line objects for some example drains from their region. The secondary routing covered a small area as an example, but it included connections to river channels and loops in the form of a ring-ditch.

4.3 Testing functional correctness

We applied tests for the rules set out in Section 3.5.1. The DEM-derived DN passed all tests. Although the test stating that watershed areas should not be negative was passed, we found some instances of zero watershed areas. Zero watershed areas eventuated because our approach followed our rules about DEM cells belonging to a single watershed (Rule 1c in Section 3.3.1) and only DEM cells intersecting lakes not necessarily having to be within the lake watershed (Rule 4c in Section 3.3.1). Approximately 900 segments had watersheds with zero area because they were very short (<8 m long). Approximately 8,600 lakes had watersheds with zero area because they were small enough to not contain any whole DEM cells.

4.4 Testing representative correctness

We received feedback from regional council, LINZ and MfE staff by email and during our monthly oversight group meetings after we shared a pilot version of our DEM-derived DN with users who were part of our project oversight group.

In general, users agreed that most of the DN represented the alignment and connectedness of river channels well. River channel alignment matched with user expectations particularly well in steeper locations. However, the pilot DN did not match with expectations about alignment or connectedness in some locations, which were often flat locations near to the coast.

Recommended improvements to our pilot fell into two main categories. The first category consisted of disconnects or misalignments that caused major downstream effects on catchment areas. We found less than ten of these major issues after having inspected the DEM-derived DN across the North Island. Examples included the outflow of Lake Rotoiti not flowing via the Kaituna River and Okere Falls and a break in the Whanganui River which caused it to flow to the coast in the wrong location. The second category consisted of minor misalignments or effects of sinkholes that caused unexpected Strahler ordering. We found approximately 20 of these minor issues after having inspected the DEM-derived DN on the North Island. Examples included the Tutaekuri River turning into a purple segment because it passed close to a sinkhole of a small tributary.

5 Discussion

5.1 Progress against proposed principles

It is important to reflect on intended end-goals when assessing a product or service. Booker et al. (2024) proposed that a national DN and the data and procedures used to develop it should meet 10 principles. Table 5-1 provides our assessment of how the current work and our pilot DN have contributed to meeting the predefined principles.

Table 5-1: Proposed principles and associated success measures for generation and operationalisation of a national digital network. Answers are self-reflections informed by workshops with project oversight group.

Principle	Success measure	Assessment
Reproducible	The methods and data used to generate the DN are clearly described such that the generation process could be replicated independently.	Partially met. This report describes the process for DN generation. Code constituting the process is recorded on Earth Sciences NZ internal GitLab. Running the code requires access to high-performance computing facilities.
Spatially consistent	Consistent methods and procedures were applied to generate all parts of the DN regardless of location within the country.	Met. The same algorithms are used in all locations. Headwater and sinkhole locations come from a nationally consistent source (Topo50). However, LiDAR coverage is not complete for the entire DEM. Parts of the DEM are populated from digitized contours rather than LiDAR.
Automatically updatable	The DN can be re-generated relatively easily following an update to input data or to a change to an automated generation algorithm or parameter.	Met. The Topo50-derived DN generation process can be re-run following changes to Topo50. The DEM-derived DN generation process can be re-run following changes to Topo50 and/or LiDAR-derived DEM.
Version controlled	The user is given information that allows them to understand that there may be different versions of the DN and why one version is different to another.	Met. All outputs are date-labelled for version control.
Bespoke updatable	Users can provide feedback on DN behaviour so that future versions can incorporate potential improvements.	Partially met. We demonstrated a mechanism for users to provide feedback on DN functionality and to provide input for proposed grey segments and secondary routing. The method requires ongoing funding to be fully operationalised.

Principle	Success measure	Assessment
Spatially complete	The DN covers the whole country.	Partially met. Outlying islands (e.g., Chatham Islands) are not included in our pilot DEM-derived DN due to those locations not being included in the input datasets. We note that separate Topo50 layers are available for the Chatham Islands (LINZ Chathams Topo50).
Functionally correct	The routing behaviour produced by the DN (e.g., downstream routing and accumulation) has been verified and can be applied by a user following instructions released alongside the DN.	Met. We applied formal tests to assess functional correctness (e.g., locations of terminal segments, sum of watershed areas). However, the current pilot DN requires burning of segment locations in some (<20) locations to improve routing or segment alignment. We note that, if burning is conducted, the standard procedure is to burn a path to the coast.
Functionally informative	The DN is accompanied with basic information required for subsequent analysis.	Met. Description of DN and instructions for routing code provided to example users via the MfE data service.
Appropriate	There is a list describing the purposes that the DN was intended to be used for at the time of production.	Partially met. We provided a list of suggested purposes for our pilot DN but testing of the appropriateness of the DN for these purposes was outside the scope of this work.
Available, including version history	The DN can be obtained by its intended users. Users can obtain previous versions.	Partially met. The pilot DN was made available via the MfE data service for testing purposes by the project oversight group.

5.2 How can the outputs be used?

Topo50 map data are produced for cartographic purposes. The line and polygon objects that plot in blue on Topo50 maps (streams, lakes, drains, water races, canals etc) can be used to make maps of freshwater features. Topo50 map data do not contain information such as river length, Strahler order, or catchment membership because it is not a DN. We found some spatial inconsistencies in the labelling of different data Topo50 layers (drains on one map tile versus water races on another map tile), which is understandable because these data layers all appear as blue lines on Topo50 maps. We generated a Topo50-derived DN and a DEM-derived DN that can be used for several purposes beyond making visual maps.

Our Topo50-derived DN can be used for applications that require position, length, and river connectivity but do not require watersheds, catchment characteristics, or overland flow pathways. Examples include:

- membership of river channels to management zones;
- fish passage assessment;
- length of river to be fenced for stock exclusion; and
- position of river segments relative to lakes.

Our DEM-derived DN can be used for the above applications and those that require information describing watersheds, catchment characteristics, or overland flow pathways. Examples include:

- membership of land parcels to management zones;
- water accounting between rivers, land, lakes and the coast, especially when engineered diversions are present;
- hydrological and hydraulic/flood modelling;
- sediment and nutrient (contaminant) modelling;
- modelling spread of invasive species;
- extracting local and catchment characteristics for monitored sites; and
- calculation and display of river classification/typologies.

The functional improvements to our DEM-derived DN over previous national DNs were generally well received by members of our project oversight group. Users agreed that LiDAR-generated object alignment and explicit representation of lakes, sinkholes, and secondary routing increased the usefulness of a DN for the above listed purposes. Users also agreed that using colours to label segments (blue, purple, grey) was preferable to using subjective terms without careful definitions (river, drain, canal).

5.3 Automation and future change requests

We demonstrated that a DN could be solely generated using automated processes because our pilot DN was initially generated without any burning or manual adjustments. Uniform application of automated processes to nationwide data fosters generation of a nationally consistent DEM-derived DN. An automated process has the advantage that it can be repeated to generate an updated DN following changes to input data (e.g., topographic data from LiDAR or changes to Topo50). However, checks applied by us and the trial users indicated that the automated processes did not generate a DN that was perfect in all locations. Some misalignments of segments and poorly represented routing were present.

The DEM-derived DN could be copied and altered by users. We understand that users may wish to alter a DN for bespoke purposes or to correct misrepresented segment alignment or connectivity. We consider this to be acceptable for local bespoke purposes. It should be noted that altering the alignment of blue/purple segments may cause inconsistencies with routing and watersheds if they are not altered accordingly. Bespoke alteration or addition to secondary routing is more viable because grey segments can be iteratively and independently added to an existing blue/purple DN. Furthermore, grey segments have no watersheds. However, proliferation of multiple altered versions of a parent DN would cause inconsistency. Inconsistency between DN versions is not desirable for national applications such policy development, or collation of datasets required for national environmental reporting.

We trialled methods for receiving change requests from users to incorporate them into a centrally held version-controlled national DN. We demonstrated that our automatically generated DN could be amended to meet a change request by manually altering the DEM or by burning. However, the infrastructure required for users to submit change requests and for the requested to be vetted, accepted and applied is not currently enabled or funded. We demonstrated that this utility would be possible in the future if funding is provided and a suitable governance arrangement can be put in place to build a nationwide DN user community.

It is important to note that a change request (e.g., a burn to correct a segment misalignment) may become outdated when a DN is regenerated with improved topographic data because the new DN may be better than the previous version. Alternatively, a change request may still be valid if the topographic data has not been altered in the area that requires burning. Therefore, methods are required for recording previously accepted change requests and assessing whether solutions to previously accepted change requests need to be applied to a newly updated DN.

5.4 Technical advancements for future improvements

We suggest that several technical advancements could be applied to improve the DEM-derived DN.

- The DN represents overland flow pathways, lakes, headwater locations, rivers, sinkhole locations, and the coast. Functionality of our DN would be improved by additional representation of wetland and aquifers because these are notable features of freshwater systems that are important for the accounting, modelling, and management of water, nutrients, and ecological conditions.
- Our rules for DN generation allowed lakes to intersect with the coastline. The process meant that two lakes had inflows but not outflows because they intersected the coast. The process could be altered so that lakes cannot intersect the coast and therefore must have outflows. Thus, guaranteeing that all lake outflows (including those located near the coast) could be extracted from models produced using the DN.
- Our rules for DN generation allowed breaching of topography to reach the coast from within inland depressions. The process meant that lakes sitting within inland basins that should not have legitimate outflows may exhibit an outflow in the DN. The process could be altered so that a terminal segment could form inside a lake. From a technical perspective this outcome could be achieved by setting the DEM elevation at the centre of the lake to -999, which was our default value for the coast in our process.
- We recommend development of a process for checking long-profiles of segments which are created by breaching or filling of topography. The process could then be altered such that segments would not be allowed to breach over a threshold of height.
- We applied our process to an 8 m DEM. The process could be applied to a finer resolution DEM (e.g., 4 or 1 m), and the results assessed for increases in accuracy of alignment and correctness of routing.
- Our process produced many small watersheds around the coast that contain neither a segment nor a lake. Some of these watersheds could be amalgamated to reduce modelling burden and the size of the watershed file without any meaningful reduction in functionality.

- Our process allowed small lakes to be disconnected from all segments and have zero watershed area. These lakes could be removed from future DNs because they serve no functional purpose within the DN.
- Our overall strategy involved production of a very detailed red network which was consistent with the blue/purple network. We created a pure version of the red network at the finest possible spatial resolution by calculating and storing surface flow directions for each DEM cell. However, the file size of these data was very large. We recommend that the trade-offs between file size, resolution, and functionality of the red network be assessed.
- LINZ could revisit the application of their data layer definitions, particularly for drains and water races.

5.5 General discussion

Our main technical advancements incorporated more realistic representations of flow pathways in our DN than would otherwise have been the case (e.g., initiating river channels at mapped headwaters, representing sinkholes, lakes with inflows and outflows, and secondary routing along artificial flow pathways). However, incorporating these phenomena into a DN created some data needs that were not present for previous DN versions. For example, our processes for initiating blue segments, initiating purple segments, and burning relied on Topo50 to accurately represent headwaters, sinkholes, and river channels respectively. Similarly, we developed the ability to represent artificial flow pathways using grey segments. We did not populate drains/ditches/pipes into the grey network because there is no well-defined and nationally consistent dataset representing these features. We demonstrate the utility of secondary routing by using canals from Topo50 to represent diversions. An example of improved functionality with secondary routing is calculation of the locations that a fish can reach with and without the effects of artificial channels by comparing the blue network against the blue plus grey network. However, full use of secondary routing does present two major challenges. First, spatially consistent and locally accurate input data are needed for small artificial flow pathways such as drains, water races, and pipes to be included in a national DN. Second, additional information is required about the proportional splitting of flow by bifurcations associated with grey segments for applications such as hydrological or nutrient modelling.

We generated outputs for a blue/purple network. We then added secondary routing and re-generated the outputs for a blue/purple/grey network using the same formatting (Table 3-5) to ensure that all code or models could be applied to either network version. An alternative option would have been to change “down” to “downPrimary” and “downSecondary” and to change “terms” to “termsPrimary” and “termsSecondary”. The advantage of our method was that all code could be applied to all networks without any amendments because “down” was present in all versions. The disadvantage of our method was that two versions of the same network were generated; one without secondary routing and the other with secondary routing.

Our DN generation process did not apply any smoothing when generating segments or watersheds. All parts of all blue and purple segments must follow one of eight directions. Blue and purple segments are therefore not smooth curves. Smoothing of segments may have produced slightly shorter segment lengths. Similarly, our watersheds had regularly shaped boundaries because our method allocated each DEM cell to a watershed. We judged that the DEM grid was fine enough that we did not need to smooth segments or watersheds.

It was particularly noteworthy that lake polygons were independent of DEM elevations. Lake polygons were not consistent with DEM topography because: a) DEM cells were lower than DEM cells inside adjacent lake polygons; and b) DEM cells inside lake polygons were not flat or near gradually sloping towards the lake outlet.

6 Conclusion

A DN generation process that maximised automation and minimised manual input requirements was devised. The process is repeatable so that a new DN can be generated following requests for changes from users or updated LiDAR topography, lake polygons, or headwater/sinkhole locations.

The new DEM-derived DN improves upon previous national DNs by explicitly representing interactions between various surface water entities, including surface flow pathways, sinkholes, artificial flow pathways, lakes, river channels, and the coast. The ability to distinguish river channels from overland flow pathways means that the new DN is suitable for a broad range of purposes.

A set of formal tests was devised and applied to validate the correctness of DN functionality following DN generation. Informal tests were used to test representative correctness. Results indicated that, while most of the DN was fit for user purposes, some misalignment of segments and poorly represented routing was present. Sub-optimal representation usually occurred in areas without LiDAR, on the border between LiDAR datasets, around lakes, or in flatter areas near to the coast.

A pilot DN generated from our automated process was shared with a selection of DN users. Methods for sharing instructions about the DN with users and receiving their requests for changes or improvements were trialled. A dedicated email address was found to be more suitable than a dedicated form because it provided necessary flexibility in user feedback while also being traceable.

7 Acknowledgements

Thanks to all members of the oversight group for their feedback and valuable contributions to this work. Thanks to Christina Robb for organising and chairing oversight group meetings. Thanks to James Rodgers from MfE for assistance with uploading the pilot DN onto the MfE data service.

8 References

Blöschl, G. (2022) Flood generation: process patterns from the raindrop to the ocean. Hydrology and Earth System Sciences, 26(9), pp.2469-2480.

Booker, D.J. (2023) National river digital networks and the River Environment Classification: future pathways for stewardship, maintenance, and upgrading of products and services for national benefit. NIWA client report 2023272CH prepared for the Ministry for the Environment, 39pp.

Booker, D.J., Wilkins, M.C, Wilkinson, C.E., Smith, R.G. (2025) Extended Strahler ordering to distinguish mapped river channels from ephemeral pathways and consistently compare digital networks. River Research and Applications, 42(4), pp.869-885.

Booker, D.J., Wilkinson, C.E., Smith, R.G., Wilkins, M.C. (2026) Projecting river names onto a national digital network. Earth Sciences New Zealand client report 2023272CH prepared for the Ministry for the Environment, 45pp.

Booker, D.J., Wilkinson, C.E., Wilkins. M.C. (2024) Digital Networks: challenges, solutions, and case studies to inform nationwide integrated freshwater-land mapping. NIWA client report 2024138CH prepared for the Ministry for the Environment. 90pp.

Lindsay, J.B. (2016) The practice of DEM stream burning revisited. Earth Surface Processes and Landforms, 41(5), pp.658-668.

Pearson, R.A., Smart, G., Wilkins, M., Lane, E., Harang, A., Bosserelle, C., Cattoën, C., Measures, R. (2023) GeoFabrics 1.0.0: An open-source Python package for automatic hydrological conditioning of digital elevation models for flood modelling. Environmental Modelling & Software, 170, p.105842.

Snelder, T.H., Biggs, B.J. (2002) Multiscale river environment classification for water resources management. JAWRA Journal of the American Water Resources Association, 38(5), pp.1225-1239.

Sprague, R., McCarthy, J., Burge, O., Etherington, T., Franklin, P., Booker, D., Lundquist, C., Leunissen, E., Houghton, K., Bolton, A., Weaver, L., Wood, S., Schallenberg, L. (2026) A current ecosystem map for Aotearoa New Zealand. BSI report to Ministry for the Environment. 74pp.

Stoffels, R.J., White, R.S. (2025) Quantifying regulatory limits for multiple stressors in an open and transparent way. Conservation Biology, 39(2), p.e14375.

Appendix A Description of LINZ topographic layers

Table A-1: Description of LINZ topographic layers considered for inclusion in the digital network generation procedure.

LINZ Layer ID	LINZ Name	LINZ Description	Used for digital network generation procedure?
50327	NZ River Centrelines (Topo, 1:50k)	A natural, flowing body of water emptying into an ocean, lake or other body of water and usually fed along its course by converging tributaries.	Yes. Represents surface water flow pathways.
50328	NZ River Polygons (Topo, 1:50k)	A natural, flowing body of water emptying into an ocean, lake or other body of water and usually fed along its course by converging tributaries.	Yes. Represents surface water flow pathways.
50320	NZ Rapid Centrelines (Topo, 1:50k)	An area of broken, fast flowing water in a stream, where the slope of the bed increases (but without a prominent break of slope which might result in a cascade or waterfall), or where a gently dipping bar of harder rock outcrops.	Yes. Represents surface water flow pathways.
50321	NZ Rapid Polygons (Topo, 1:50k)	An area of broken, fast flowing water in a stream, where the slope of the bed increases (but without a prominent break of slope which might result in a cascade or waterfall), or where a gently dipping bar of harder rock outcrops.	Yes. Represents surface water flow pathways.
50370	NZ Waterfall Centrelines (Topo, 1:50k)	A sudden descent of water over a step or ledge in the bed of a river.	Yes. Represents surface water flow pathways.
50371	NZ Waterfall Edges (Topo, 1:50k)	A sudden descent of water over a step or ledge in the bed of a river.	Yes. Represents surface water flow pathways.
50373	NZ Waterfall Polygons (Topo, 1:50k)	A sudden descent of water over a step or ledge in the bed of a river.	Yes. Represents surface water flow pathways.
50292	NZ Lagoon Polygons (Topo, 1:50k)	A sheet of salt water separated from the open sea by sand, shingle banks or a road embankment. The sheet of water between an offshore reef, esp. of coral, and mainland. The sheet of water within a ring or horseshoe shaped atoll.	Yes. Represents surface water flow pathways.
50293	NZ Lake Polygons (Topo, 1:50k)	Any standing body of fresh inland water.	Yes. Represents surface water flow pathways.

LINZ Layer ID	LINZ Name	LINZ Description	Used for digital network generation procedure?
50353	NZ Soakhole Points (Topo, 1:50k)	The point at which a river or stream percolates into or through the soil.	Yes. Used to exclude potential headwaters coinciding with soak hole points.
50250	NZ Canal Centrelines (Topo, 1:50k)	An artificial open waterway used for transportation, waterpower or irrigation.	Yes. Included where connecting river systems to the coast.
50251	NZ Canal Polygons (Topo, 1:50k)	An artificial open waterway used for transportation, waterpower or irrigation.	Yes. Included where connecting river systems to the coast.
50262	NZ Drain Centrelines (Topo, 1:50k)	A channel which carries surface water.	Yes. Included where connecting river systems to the coast.
50260	NZ Dam Centrelines (Topo, 1:50k)	A barrier constructed across a watercourse to control the flow or raise the level of water.	Yes. Used to identify places where a dam may join two sections of river.
50288	NZ Island Polygons (Topo, 1:50k)	Area of dry or relatively dry land surrounded by water or low wetland.	Yes. Used to identify islands within lakes and rivers, which are needed for conversion of river and lake polygons to lines.
51153	NZ Coastlines and Islands Polygons (Topo: 1:50k)	This topographic coastline is the line forming the boundary between the land and sea, defined by mean high water. Islands from the NZ Island Polygons layer that lie within the NZ Coastline and Chatham Islands areas (i.e., islands in lakes, rivers and estuaries) have been removed.	Yes. Used to ensure rivers intersect with New Zealand land area.

LINZ Layer ID	LINZ Name	LINZ Description	Used for digital network generation procedure?
50280	NZ Geographic Names (Topo: 1:50k)	Geographic name holds the proper noun of certain objects within the Topo databases. Geographic name does not apply to roads, railway lines, or descriptive text – e.g., 'Brunner Mine' would be a Geographic name but 'mine disused' is held in the descriptive_text feature.	Yes. Used to name rivers.
50372	NZ Waterfall Points (Topo, 1:50k)	A sudden descent of water over a step or ledge in the bed of a river.	No. A point object is not useful for representation of surface water flow pathways.
50258	NZ Coastlines (Topo, 1:50k)	The line forming the boundary between the land and sea, defined by mean high water.	No. Not representative of natural flow pathways.
50310	NZ Pond Polygons (Topo, 1:50k)	An area of shallow water with walls or banks created for a specific purpose.	No. Not representative of natural flow pathways.
50324	NZ Reservoir Polygons (Topo, 1:50k)	Reservoir; receptacle in which a large quantity of water is stored.	No. Not representative of natural flow pathways.
50369	NZ Water Race Centrelines (Topo, 1:50k)	A man-made water channel, often running parallel to the contours, built specifically for the distribution of water to crops, pasture, mining activities etc.	No. Not representative of natural flow pathways.
50356	NZ Spring Points (Topo, 1:50k)	The place where water issues from the ground naturally.	No. Many springs are not near a river or lake so do not represent initiation of natural flow pathways.
50358	NZ Swamp Points (Topo, 1:50k)	A wet or moist region with water standing on or just below the surface of the ground, and usually covered by a growth of vegetation.	No. Not representative of surface water flow. Swamp regions can span headwaters.
50359	NZ Swamp Polygons (Topo, 1:50k)	A wet or moist region with water standing on or just below the surface of the ground, and usually covered by a growth of vegetation.	
50352	NZ Snow Polygons (Topo, 1:50k)	A region of permanent snow.	

LINZ Layer ID	LINZ Name	LINZ Description	Used for digital network generation procedure?
50287	NZ Ice Polygons (Topo, 1:50k)	A region of permanent ice.	No. Not representative of surface water flow. Permanent snow and ice regions can span many headwaters.
50323	NZ Reef Polygons (Topo, 1:50k)	A ridge, usually of rocks or coral, lying near the surface of the sea, which may be visible at low tide, but is usually covered with water.	No. Not a water feature.
50331	NZ Rock Points (Topo, 1:50k)	An isolated rock or rock formation between high and low tides.	No. Not a water feature.
50332	NZ Rock Polygons (Topo, 1:50k)	Coastal rocks; rocks between high and low tide.	No. Not a water feature.
50343	NZ Shoal Polygons (Topo, 1:50k)	A submarine object over which the water is shallow, which is detached from the shore and whose bottom is visible at low water; may be a menace to navigation on the surface.	No. Not a water feature.

Appendix B Regional council input data survey responses

Table B-1: Regional Council survey responses: dataset metadata. Responses shown include only key survey fields.

Authority	Name	Description/ Purpose	Date	Spatial coverage	Format	Privacy considerations	License/ Usage rights	Version control
Tasman District Council	Hydrology Rivers LiDAR Derived Streams	Summary The Tasman District Council Digital Rivers Network is region-wide geometric stream and drainage network dataset. Flow paths have been extracted from a hydro-corrected terrain model using a minimum contributing catchment area threshold of 2 ha. This dataset may serve as a foundational resource for various analyses, including hydrology, stormwater management, ecology, and planning. For full method details, limitations, assumptions and recommendations please refer to the supplied technical memo: TDC Digital Rivers Network – Technical Memo.pdf	1/01/2023	Tasman District, TasmanExtent West 172.048091 East 173.305004 North -40.496995 South -42.302164	Enterprise Geodatabase Feature Class	NO	CC BY 4.0 License View license details	internal versioning
Hawke's Bay Regional Council	Hawkes Bay LiDAR Stream Network 2020	Digital streamlines representing the drainage network with a 10-ha upstream area threshold for channel initiation. Created to support Regional SedNet Model for Hawke's Bay.	1/06/2024	Hawke's Bay region	ESRI filegeodatabase	Public Dataset	CC BY INT 4.0	Static Dataset
Hawke's Bay Regional Council	Hawkes Bay LiDAR Subwatersheds 2020	Subwatersheds draining to segments within the digital stream network (10-ha threashold). Created to support Regional SedNet Model for Hawke's Bay.	1/06/2024	Hawke's Bay region	ESRI filegeodatabase	Public Dataset	CC BY INT 4.0	Static Dataset

Authority	Name	Description/ Purpose	Date	Spatial coverage	Format	Privacy considerations	License/ Usage rights	Version control
BOPLASS Ltd	Bay of Plenty LiDAR 1m DEM 2019–2022	2019-2022 is the latest region wide coverage flown over this year range.	27/10/2019	Bay of Plenty Region	DEM from LINZ data service - ascii, geotiff	None	Creative Commons 4.0	See LDS
Regional Software Holdings Limited (RSHL)	Bay of Plenty LiDAR 1m DEM (2024)	Latest BOP LiDAR for part of Western Bay and the coast to Opotiki.	1/02/2024	See LDS	ascii or geotiff	None	Creative Commons 4.0	See LDS
Bay of Plenty Regional Council	Catchments Streams_DTM_2011/Model 1 inputs/zx_01_burn	We have previously developed a stream and catchment network from our original region-wide LiDAR dataset which was capture in 2011. For this we divided our region into nine zones - basically large sea-draining catchments so all our input data is divided this was as well. So we have z1_01_burn, z2_01_burn, z3_01_burn etc. These are all the features that were burnt into the DEM and includes river lines, bridges, culverts etc. Unfortunately, there is no attribution saying what line is what feature and there is no guarantee of completeness. P.S - below is an arbitrary date in 2014 and 2016 for created and updated. I've no idea when in 2014 it was made, or when we made updates.	30/06/2014	Covers the Bay of Plenty region in separate datasets for each of 9 zones/catchments	File geodatabase line feature class	None	Creative Commons 4.0	No changes have happened since 2016-ish and no changes are planned so no version control. It is what it is.

Authority	Name	Description/ Purpose	Date	Spatial coverage	Format	Privacy considerations	License/ Usage rights	Version control
Bay of Plenty Regional Council	Catchments Streams_DTM_2011/Model 1 inputs/z2_03_sink, z3_02_sink	Describes legitimate sinks in zones 2 and 3 for lakes and wetlands/depressions without out flow. TYPE attribute field describes if it is a lake or land sink.	30/06/2014	Both datasets together cover the Rotorua lakes area.	file geodatabase polygon feature class	None	Creative Commons 4.0	No version control
Bay of Plenty Regional Council	Catchments Streams_DTM_2011/Model context/BOPRC_lakes	Significant lakes in the Bay of Plenty Region digitised from 2011 aerial photography so they matched the LiDAR data. Does not include all the small ponds etc. that is in the topo data that can create confusion in the resulting stream network and catchments.	30/06/2014	Bay of Plenty region but there aren't any up the east coast. I think this is correct. There aren't any significant lakes in that area.	File geodatabase polygon feature class	None	Creative Commons 4.0	No version control
Tasman District Council	Stormwater Assets	Operational Stormwater Drains, Stormwater Features, Stormwater Pipes derived from Councils Asset Management System. Used in the stream delineation project for hydro correction, in addition to manual burnlines, the piped stormwater network was used to identify and route piped streams.	7/01/1997	Tasman District urban areas	ESRI Map Service	N/A	Creative Commons 4.0 International CC BY 4.0	N/A

Authority	Name	Description/ Purpose	Date	Spatial coverage	Format	Privacy considerations	License/ Usage rights	Version control
Hawke's Bay Regional Council	HBRC Asset Lines	Spatial representation of Hawke's Bay Regional Council (HBRC) flood protection assets that are represented as line in HBRC's asset management system (point and polygon layers also exist). Asset type includes the likes of drainages, stopbanks, access tracks, streams, rope rail, culvert etc.	1/01/2000	Hawke's Bay Region	Vector - ESRI feature service	Public Dataset	CC BY INT 4.0	Internal Versioning at source data level
Hawke's Bay Regional Council	HBRC Asset Points	Spatial representation of Hawke's Bay Regional Council (HBRC) flood protection assets that are represented as points in HBRC's asset management system (point and polygon layers also exist). Asset type includes the likes of floodgate and pump station.	1/01/2000	Hawke's Bay Region	Vector - ESRI feature service	Public Dataset	CC BY INT 4.0	Internal Versioning at source data level
Hawkes Bay Regional Council	Hawkes Bay Surface Catchments	This layer contains definitions for the main catchment that form the basis of Regulatory Environmental Management Areas (EMA) as well as some sub-catchments across the region. A surface catchment is an area of land where water collects and drains off into a common outlet, such as a river, lake, or ocean.	15/08/2024	Hawkes Bay region plus surface catchments that flow into Hawke's Bay	Vector - ESRI feature service	Public Dataset	CC BY INT 4.0	Updated from Source Dataset that has version control on

Authority	Name	Description/ Purpose	Date	Spatial coverage	Format	Privacy considerations	License/ Usage rights	Version control
Auckland Council	Permanent Streams	Permanent streams are the predicted primary flow paths of permanent continuous surface water flow throughout the year in seasons of normal rainfall for the Auckland Region. Where flow paths are natural and unmodified the flow paths were derived from the Overland Flow Path layer. Where the flow path is modified because of streams being piped or channels being constructed, the stream path was aligned to the modified path to form a primary flow path for the stream. Overland flow paths with an upstream catchment of >= 5ha were predicted to be permanent streams. The Auckland Council stormwater asset network pipes, modified channels and ponds were used to predict the modified flow path of the stream. Assets included in this layer are pipes, culverts, and ponds.	1/01/2000	Auckland Region	ESRI feature class		Creative Commons	Internal versioning

Table B-2: Regional Council survey responses: dataset technical attributes. Responses shown include only key survey fields.

Authority	Name	Spatial resolution	Vertical accuracy	CRS	Data quality / validation status	Data lineage	Known issues or limitations
Tasman District Council	Hydrology Rivers LiDAR Derived Streams	Derived from 1–2m DEM - max zoom scale 1:5,000	?	NZTM2000	QA/QC completed.	See separate Tech memo.	District wide coverage but small area in South of district derived from 8–10m DEM data. Underlying issues with Karstic landscape and Dune lakes.

Authority	Name	Spatial resolution	Vertical accuracy	CRS	Data quality / validation status	Data lineage	Known issues or limitations
Hawke's Bay Regional Council	Hawkes Bay LiDAR Stream Network 2020	1:2,500	NA	NZTM2000	Visual review for general completeness of coverage.	Created by MWLR. Largely automated output from watershed modelling. Some editing of the network was undertaken on the flats, utilising asset drainage features as primary data source.	Data gaps on the coastline where watershed weren't formed by modelling.
Hawke's Bay Regional Council	Hawkes Bay LiDAR Subwatersheds 2020	1:2,500	NA	NZTM2000	Visual review for general completeness of coverage.	Created by MWLR. Automated output from watershed modelling.	Data gaps on the coastline where watershed weren't formed by modelling.
BOPLASS Ltd	Bay of Plenty LiDAR 1m DEM 2019–2022	See LDS	See LDS	NZTM2000	See LDS.	LiDAR was captured for BOPLASS Ltd by Aerial Surveys Ltd between 27 October 2019 – 24 October 2022. These datasets were generated by Aerial Surveys Ltd and their subcontractors.	None known.
Regional Software Holdings Limited (RSHL)	Bay of Plenty LiDAR 1m DEM (2024)	See LDS	See LDS	NZTM2000	See LDS.	LiDAR was captured for Regional Software Holdings Limited (RSHL) by Woolpert Ltd between 1 February 2024 and 30 April 2024. These datasets were generated by Woolpert and their subcontractors.	None known.
Bay of Plenty Regional Council	CatchmentsStreams_DTM_2011/Model 1 inputs/zx_01_burn	Smallest feature is an approximately 2m culvert line	No elevation	NZTM2000	Only shows what was found on the aerial photography so accuracy as per the 2011 aerial photography and not complete.	Just digitised from 2011 aerial photography. No processing history, no transformations, no idea about edits and modifications.	Not complete - just what was needed to fix obvious issues with resulting stream network.
Bay of Plenty Regional Council	CatchmentsStreams_DTM_2011/Model 1 inputs/z2_03_sink, z3_02_sink	smallest sink (not size of sink polygon) approximately 2000sqm	No elevation	NZTM2000	Only by Opus when data created.	Created by Opus, no modifications.	Only sinks known. May not be complete.
Bay of Plenty Regional Council	CatchmentsStreams_DTM_2011/Model context/BOPRC_lakes	Smallest lake - 3ha	no elevation	NZTM2000	No checks apart from when created by Opus.	Created by Opus, 2014. No modifications.	I believe this is relatively complete for the significant lakes.

Authority	Name	Spatial resolution	Vertical accuracy	CRS	Data quality / validation status	Data lineage	Known issues or limitations
Tasman District Council	Stormwater Assets	0.2153	± 0.30m	NZTM2000	QA/QC completed.	Compiled from both historical and modern asbuilt plans, since 2020 asbuilt data captured confirms to the Nelson Tasman Land Development Manual. Source data is generally provided in NZGD 2000 Nelson Circuit, which is then converted to NZTM2000.	Accuracy and completeness can vary.
Hawke's Bay Regional Council	HBRC Asset Lines	1:2,500	NA	NZTM2000	Dataset is activity managed by Engineering team, be it with light resources. Some asset types are better validated that others.	Features have been digitized using heads-up digitizing from aerial imagery, DEM, some ad hoc field validation has been completed. Drains and streams have been digitized with flow directions, other features such as culverts (currently) don't.	Data quality and completes varies from across the dataset. Quality and completeness checks are ad hoc.
Hawke's Bay Regional Council	HBRC Asset Points	1:2,500	NA	NZTM2000	Dataset is activity managed by Engineering team, be it with light resources. Some asset types are better validated that others.	Features have been digitized using heads-up digitizing from aerial imagery.	Data quality and completes varies from across the dataset. Quality and completeness checks are ad hoc.
Hawkes Bay Regional Council	Hawkes Bay Surface Catchments	1:2,500	NA	NZTM2000	Manual review during creation.	Catchments are named. Created using a combination of Hawkes Bay LiDAR Subwatersheds 2020, engineering drainage watersheds, MHWS 2023 and asset lines datasets.	Spatial definitions within vicinity of estuary and flat plains less reliable than through hill country.
Auckland Council	Permanent Streams	Overland flow paths based on 1m resolution DEM raster	+/- 20cm (95%)	NZTM2000 (EPSG:2193); Vertical datum: NZVD2016	Yes.	Contact the Lifecycle Planning Team.	Contact the Lifecycle Planning Team.

Table B-3: Regional Council survey responses: DN generation attributes. Responses shown include only key survey fields.

Authority	Name	Surface flow pathways	Routing data	Other information included	Data type (what the data represents)	Origin
Tasman District Council	Hydrology Rivers LiDAR Derived Streams	Rivers,Streams	Connections_between_segments, Flow_directions	Sinuosity,Strahler_Order, Catchment_Areas,Upstream/downstream_Connectivit	River/stream_lines_(mapped_hydr,Digital_Network	Remote_Survey_(desktop_mapping_
Hawke's Bay Regional Council	Hawkes Bay LiDAR Stream Network 2020	Streams,Ditches	Connections_between_segments, Flow_directions		Digital_Network	Remote_Survey_(desktop_mapping_
Hawke's Bay Regional Council	Hawkes Bay LiDAR Subwatersheds 2020			Watershed_Data	Digital_Network	Remote_Survey_(desktop_mapping_
BOPLASS Ltd	Bay of Plenty LiDARr 1m DEM 2019-2022				DEM	Field_Survey
Regional Software Holdings Limited (RSHL)	Bay of Plenty LiDAR 1m DEM (2024)				DEM	Field_Survey
Bay of Plenty Regional Council	CatchmentsStreams_DTM_2011/Model 1 inputs/zx_01_burn	Rivers,Streams			Obstacle_bypass_lines	Remote_Survey_(desktop_mapping_
Bay of Plenty Regional Council	CatchmentsStreams_DTM_2011/Model 1 inputs/z2_03_sink, z3_02_sink	Sinks			legitimate sinks	Remote_Survey_(desktop_mapping_
Bay of Plenty Regional Council	CatchmentsStreams_DTM_2011/Model context/BOPRC_lakes				Lake_outlines_and_outlets	Remote_Survey_(desktop_mapping_
Tasman District Council	Stormwater Assets	Streams,Ditches	Connections_between_segments, Flow_directions		Artificial_channels_-_drains,Artificial_channels_-_diversion,Artificial_channels_-_realignme,Flood_protection_assets_(stopba	Field_Survey,Asbuilt plans

Authority	Name	Surface flow pathways	Routing data	Other information included	Data type (what the data represents)	Origin
Hawke's Bay Regional Council	HBRC Asset Lines	Streams,Ditches	Flow_directions		River/stream_lines_(mapped_hydr,Artificial_channels_-_drains,Urban_drainage_infrastructure,Flood_protection_assets_(stopba	Remote_Survey_(desktop_mapping_
Hawke's Bay Regional Council	HBRC Asset Points	other			Flood_protection_assets_(stopba	Remote_Survey_(desktop_mapping_
Hawke's Bay Regional Council	Hawkes Bay Surface Catchments			Catchment_Areas	Regulatory main and sub catchments	Remote_Survey_(desktop_mapping_
Auckland Council	Permanent Streams	Rivers,Streams,Ditches,Ephemeral_Channels,Stormwater infrastructure	Flow_directions	Watershed_Data,Catchment_Areas	River/stream_lines_(mapped_hydr,Artificial_channels_-_drains,Artificial_channels_-_diversion,Artificial_channels_-_realignme,Urban_drainage_infrastructure	Field_Survey,Remote_Survey_(desktop_mapping_

Appendix C Instructions provided with pilot DN

Abstract

This pilot national Digital Network version 4.0 provides river line geometries for the main islands of New Zealand. River line geometries have been generated from a DEM derived from LiDAR topography where it is available and digitised contours where LiDAR is not available. The DEM was previously altered to represent continuous channel geometries for flood modelling purposes where appropriate data were available (e.g., removal of the blocking effect of bridges as describe by Pearson et al. (2023). No burning was therefore applied when generating the digital networks river line geometries.

Purpose

This pilot national Digital Network version 4 is draft and provisional. The data is provided solely for the purpose of seeking feedback and comment on segment alignment and connectivity from end users. The data is provided "as is", without warranty as to accuracy or completeness, and is shared in confidence with staff from the project oversight group to provide feedback. It has not been finalised, approved, or quality assured and must not be relied on for operational, legal, planning, or decision-making purposes.

The data may be reviewed internally within the organisations who are given access and used only to provide feedback to the project team. It must not be shared, published, redistributed, or used for any other purpose without prior written permission, except where disclosure is required by law.

No intellectual property licence is granted (including any Creative Commons licence). All rights remain with the data owner. Any future public release or licensing will occur only once the data is final and formally approved.

Please provide feedback about segment alignment and network routing to Earth Sciences NZ by emailing river.network@earthsciences.nz. An example of how to provide feedback is provided in the attachment.

Secondary routing along all South Island canals represented in LINZ Topo 1:50k maps has been incorporated into the dataset for demonstration purposes. Propose additions to secondary routing to Earth Sciences NZ by emailing river.network@earthsciences.nz. Instructions on how to provide suggestions for secondary routing are described in the attachment. We particularly encourage the addition of secondary routing for the North Island.

Data associated with each segment includes the following.

- id = segment id
- Type = blue indicates downstream of a head; purple indicates downstream of a sink; grey indicates secondary routing
- Down = downstream segment id
- Intolake = downstream lake id
- Outlake = upstream lake id
- Order = Strahler stream order

- hid and sid relate to id's for headwaters and sinkholes respectively.

Not applicable values are indicated by -1.

Attachment Word document

Thank you for agreeing to provide feedback about segment alignment and network routing of pilot national Digital Network version 4.

River line geometries (https://data.mfe.govt.nz/layer/124347-digital-river-network-v4-2026-rivers/) are to be used in conjunction with a separately supplied dataset describing lake polygon geometries (https://data.mfe.govt.nz/layer/124346-digital-river-network-v4-2026-lakes/). Lake IDs are a reference to the LID field in the lake data layer.

Information about segment connectivity allows upstream and downstream tracing to be applied. Instructions on how to apply tracing of flow pathways using QGIS are provided below. Utilising the tracing capability in QGIS may be a useful tool for identifying segment alignment or connections that require improvements.

Instructions to enable tracing

Tracing of flow pathways from a selected reach can be useful when checking a digital network. The "id" and "Down" attributes can be used to trace upstream and downstream within the network.

The following instructions allow tracing using QGIS:

1. Download these files to a local directory
   - https://data.mfe.govt.nz/layer/124347-digital-river-network-v4-2026-rivers/
   - https://data.mfe.govt.nz/layer/124346-digital-river-network-v4-2026-lakes/
   - Python script "traceupanddown.py"
2. Open QGIS.
   - Load the data using Layer -> Add Layer -> Add vector layer -> Source -> select the file holding the river geometries. Repeat for the file holding the lake geometries.
   - Go to Plugins -> Python Console, hit Show Editor button (third from left in the Python Console), then Open Script (first button) and load: "traceupanddown.py ".
   - Ensure you have the rivers layer selected and hit the green play button (fifth button). It will take 5–10s to cache the routing.
   - You can then use the Select tool (Edit -> Select -> Select features) to view the upstream and downstream trace emanating from a selected segment.
   - The default Selection colour in QGIS is yellow which might be hard to see. You can set your selection colour in Settings -> Canvas & Legend -> Selection colour (we recommend you make rivers blue and selection red).

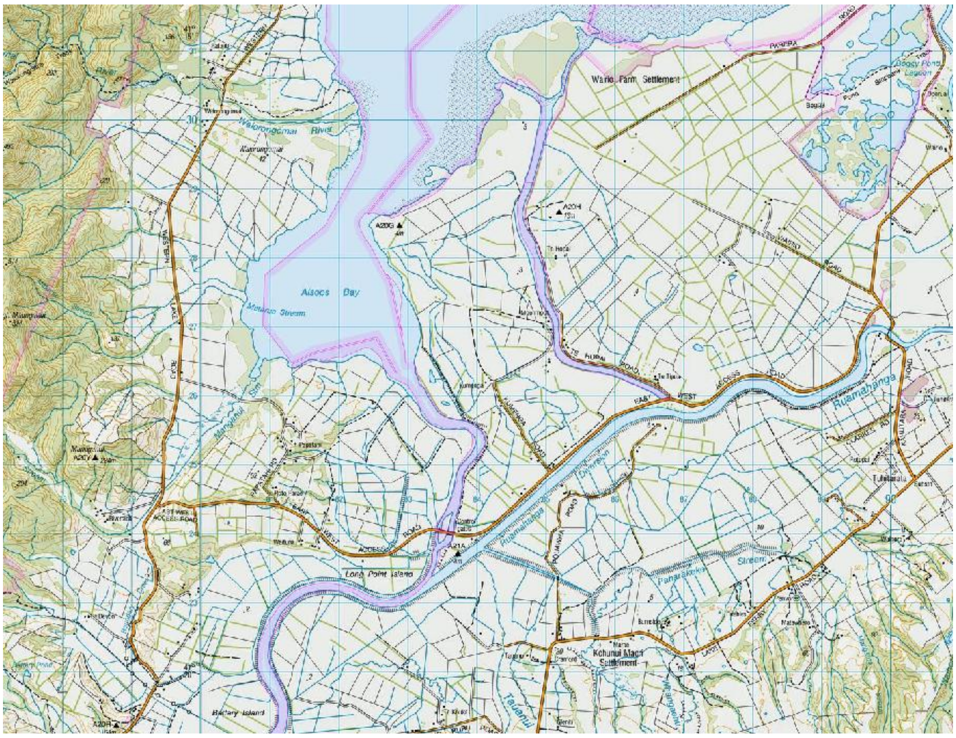
Example of feedback to improve routing

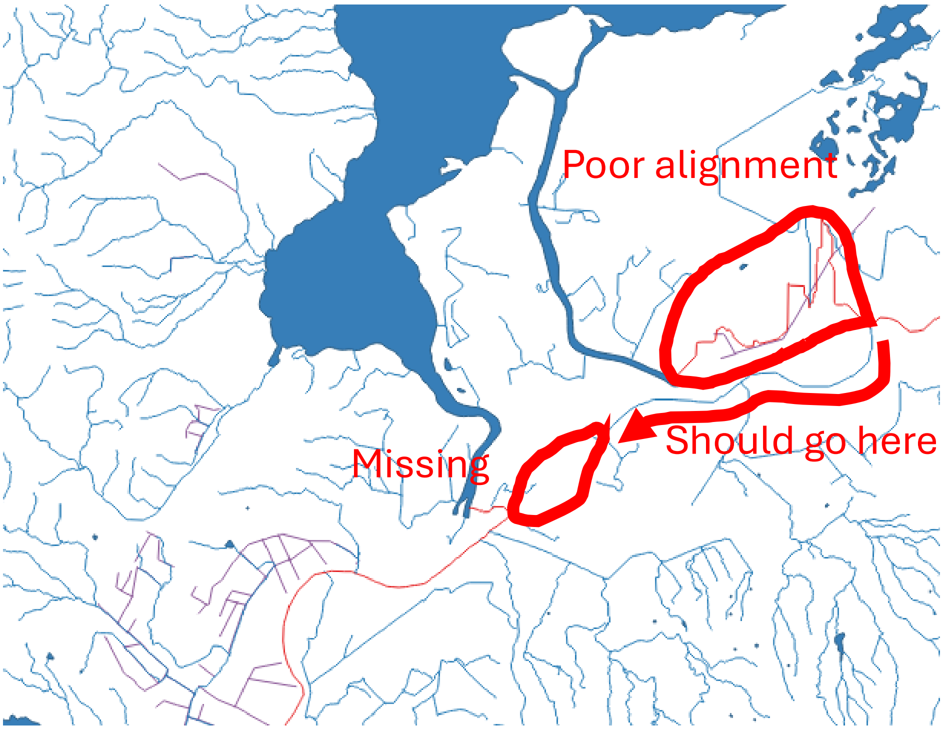
Feedback describing improvements to segment alignment and network routing of the draft river network is encouraged. Below is an example of a request for an improvement that users can follow. When providing feedback, please use place names used by LINZ that can be searched for in MapToaster TopoNZ.

Location: Ruamahanga River southeast of Lake Wairarapa (around WGS-84 -41.28141S, 175.25031E).

Description: Unexpected channel alignment upstream of Ruamahanga Diversion. Ruamahanga Diversion missing from both primary and secondary routing.

Screen shot below highlights the issue. For this example, the top image is taken from MapToaster TopoNZ and the bottom image is taken from QGIS. Blue lines are DN river segments.





Secondary routing

Conceptualisation

Secondary routing allows segments representing artificial flow pathways such as canals, ditches, or drains to be added to the river network. For ease of communication, we apply the following colour labels to distinguish types of segment:

- blue for river channel segments;
- red for overland flow segments upstream of river channels;
- purple for overland flow segments downstream of sinkholes; and
- grey for secondary routing segments.

Red, purple, and blue segments have watersheds which are polygons that drain to a segment within their boundary. Flow pathways along red, purple, and blue segments can only take one downstream route (unless a lake with multiple outlets is encountered).

Grey segments are used to apply secondary routing. They have no watersheds. They only exist to represent a flow pathway from one location to another location in the network. Grey segments often represent flow pathways along artificial channels or pipes such as diversions or tunnels. However, grey segments could also represent natural flow pathways such as braided/anastomosing rivers or flow pathways that emerge under flood conditions.

Unlike the blue or purple segments, grey segments can cross blue or purple segments without intersecting or interacting with them. For example, a grey segment representing a diversion may initiate at a blue segment, cross (and not interact with) several blue segments, and then terminate at a blue segment. This situation might represent a diversion that takes water from a river, is piped under or over some tributaries, before discharging water into a river in another catchment.

Specifying a grey flow pathway

A grey flow pathway may comprise one or more grey segments. A grey flow pathway is defined by its geometry and two "ends". The geometry comprises an ordered list of LineString or Polygon geometries. Ordered means that the first and last elements of the list are at either end of the grey flow pathway, and the elements in between are ordered in the direction of flow. There is no requirement for the points within each individual geometry to be oriented in the direction of flow. For example, individual geometries in the list might have come from the LINZ canal data layer, and the points within these individual LINZ geometries are often not ordered in the direction of flow.

The "ends" of the grey flow pathway are indicated by coordinate locations. Each end of a grey flow pathway may be:

- unconnected to any other segments, representing an artificial headwater or sinkhole (a grey segment that ends near the coastline can be viewed as representing an artificial channel that flows to the ocean);
- connected to a blue or purple segment, representing flow from or to a river; or
- connected to another grey segment, representing flow from or to another artificial channel (thus sets of grey segments can be configured into their own complex networks).

Inputted end locations are indicative. They do not have to coincide exactly with the geometries of the grey segment they pertain to or to the object that the end of the grey object is intended to connect to. For connected ends of grey flow pathways, these locations specify the centre of a search for the closest object being connected to (river, lake, other grey). The closest point on the object being connected to will become the start or end of the grey flow pathway.

There are two acceptable formats for specifying grey segments: a) a single JSON file; or b) two CSV files that cross reference each other. Simplified examples follow.

JSON format

A JSON file is a txt file with a prescribed format, that is editable in programs such as notepad. Here is an example for a single grey segment:

```
{
  "name": "an optional name for you",
  "start": [1497544.0, 5175607.0],
  "end": [-43.6733, 171.8676],
  "snap_start": "river>=5",
  "snap_end": "lake",
  "geoms": ["50250:8308217", "canalfile.gpkg:lines:id=13"]
}
```

The name is an optional identifier which might help the user. Both start and end are coordinates of two points, in NZTM easting/northing or WSG84 lat/lng coordinates (either format is acceptable). Omission of snap_start and/or snap_end will result in the grey flow pathway being unconnected at its start and/or end respectively. If included, snap_start and snap_end specify the type of object the start and end of the grey flow pathway are connected to. Possible values of snap_start and snap_end are:

- “river” for the closest point on a blue segment.
- “river=x” for the closest point on a blue segment of Strahler order x.
- “river>=x” for the closest point on a blue of Strahler order x or more.
- “lake” for the closest point on the closest lake.
- “grey” for the closest point on the closest grey segment that is also specified in this file.

In the above example a search from easting=1497544.0 northing=5175607.0 for a river with Strahler order 5 or more will be made. The closest point will be prepended to the geometry specified in the geoms field. This will result in the grey flow pathway being attached to the nearest blue segment of Strahler order 5 or more. In the above example the lake shoreline closest to the point at latitude=-43.6733, longitude=171.8676 will be selected as the end of the grey flow pathway.

In the above example the grey flow pathway will be made up of two geometries. The first geometry will be the object with a t50_fid of 8308217 from LINZ data layer ID 50250 (NZ Canal Centrelines from Topo, 1:50k). The second geometry is the object with an ID of 13 contained in the inputted file canalfile.gpkg.

Multiple grey segments can be specified by listing them one after another with commas between, e.g.,

```
[
  {
    ... first segment ...
  },
  {
    ... second segment...
  }
]
```

CSV format

An alternative way of specifying grey flow pathways is with two CSV files. CSV files may be familiar to end-users, but extra care is needed because consistent formatting is required to allow correct cross-referencing between the two files. Two files are needed because the geometries and start/end points do not share columns. One file is needed for start and end points. A separate CSV file is needed to specify segment geometries.

Here is an example from the file that specifies the start and end points:

```
Group,  xNZTM,      yNZTM,       IsFrom,  riverorder,  lake,  grey
1,      1457922.2,  5152762.2,   1,       5,           0,     0
1,      1498473,    5174857,     0,       0,           1,     0
2,      1497544,    5175607,     1,       7,           0,     0
2,      1508786.35, 5176364.03,  0,       0,           0,     0
```

The Group column is used to match each point with a corresponding geometry in the geometry CSV file shown below. Two rows must be specified for each group that appears in the geometry CSV file (one row for start; one row for end). The second and third column specify the locations of the points. IsFrom is a boolean to indicate if the row is a starting or ending point (0 for end; 1

for start). The final three columns indicate how the points are to be used. Care must be taken to ensure that at most only one of these final three columns is set. Riverorder indicates that the closest river of the given order is used, lake and brown are boolean columns, and if 1 (True) then the closest lake or grey river is used as a starting or ending point.

Here is an example from the file that specifies segment geometries:

```
t50_fid,   DataLayer, Group,   Sequence
100096611, 50251,     1,       1
2428311,   50309,     1,       2
100123913, 50251,     1,       3
2428432,   50309,     1,       4
6768169,   50250,     1,       5
```

The Group column is used to associated geometries with start/end points specified in the first CSV file. The Sequence column is used to specify the flow direction of the geometries. The first two columns indicate the source of the geometry using a similar format to the JSON format above.