NIWA
Taihoro Nukurangi

Digital Networks:

challenges, solutions, and case studies to inform nationwide integrated freshwater-land mapping

Prepared for Ministry for the Environment

May 2024

Prepared by:
Doug Booker, Clare Wilkinson, Matt Wilkins

For any information regarding this report please contact:

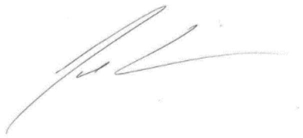
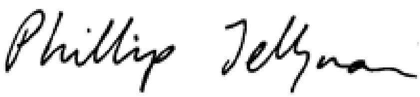
Doug Booker
Hydro-ecological modeller
Freshwater Modelling Group
+64 3 348 8987
doug.booker@niwa.co.nz

National Institute of Water & Atmospheric Research Ltd
PO Box 8602
Riccarton
Christchurch 8440

Phone +64 3 348 8987

NIWA CLIENT REPORT No: 2024138CH
Report date: May 2024
NIWA Project: MFE24502

Revision	Description	Date
Version 1.0	Final version sent to client	31 May 2024

Quality Assurance Statement		
	Reviewed by:	Jason Alexander
	Formatting checked by:	Rachel Wright
	Approved for release by:	Phil Jellyman

Contents

Tables

Figures

Executive summary

Digital Networks (DNs) are virtual representations of spatially explicit connections across coupled freshwater-land systems. DNs are an important tool for many aspects of freshwater policy, planning, reporting, management, and research because they are a fundamental input to classifications, typologies, models, simulations, and quantitative analysis. A nationally consistent DN is needed to ensure consistency between regions for nationwide applications such as environmental reporting, policy analysis, and flood forecasting. The generation of DNs is challenged by technical constraints, data availability, and because requirements for DN characteristics differ between uses.

A DN is characterised by its level of spatial detail (resolution), accuracy of object positioning (alignment), and spatial extent (coverage), which are determined by a combination of input data, technical methods, and developer decisions. Large rivers are usually represented within DNs, but fine resolution DNs can include small streams, drainage ditches, or ephemeral channels along which surface water may flow for short-lived periods. This is important because biophysical modelling may benefit from detailed spatial representation inside and outside of river channels whereas river management and policy development purposes may not require representation of surface flow pathways outside of river channels.

High resolution Digital Elevation Models (DEMs) derived by applying algorithms to process Light Detection and Ranging (LiDAR) data are currently available for many, but not all, land areas. High resolution DEMs can be advantageous because they can resolve individual channels, however, they increase computational demands. They also have the disadvantage that they include false or inconsequential topographic artifacts that need to be removed or ignored during DN generation.

A sequence of technical steps for DN generation, regardless of technical details about data or algorithms, is outlined. Technical advancements designed to increase DN utility are then described. A multi-coloured labelling system for DN objects is devised and demonstrated. The multi-coloured labelling system can 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 rivers, purple objects can be added to represent ephemeral flow pathways, green objects can be added to represent engineered flow pathways, and red objects can be used to represent surface flow pathways at the highest possible level of detail. A negative Strahler stream order scheme is devised and demonstrated. The negative Strahler stream order scheme can be used to facilitate consistent comparisons between DNs regardless of their resolution.

Challenges for DN generation associated with various landscape phenomena are outlined. Lakes, wetlands, estuaries, braided rivers, springs, and sink holes are not explicitly represented in DEMs, and flow directions within these landscape phenomena are not easily extracted from DEMs or represented in DNs. Existing nationwide data and potential solutions to improve representation of these phenomena are outlined. Multi-channel rivers and artificial channels require bifurcations to be included in DN routing. Methods were developed to improve DN functionality by representing bifurcating channels so that islands, braided rivers, and artificial channels can be represented in DNs. Various solutions are set out to meet challenges of routing through lakes.

Methods for testing DN functionality, in terms of correctness of routing, are outlined. Approaches for testing DN alignment, including subjective visual comparisons against remotely sensed images and objective measures of difference with digitised maps, are discussed. Both approaches for testing alignment are challenged by a lack of independent ground-truthed data for DN validation.

Two case studies are used to demonstrate that a variety of automated procedures are available to generate DNs. In some situations, DN alignment and routing is more sensitive to DEM data than to DN generation algorithms. Automated DN generation procedures generally worked well in steep locations but less well in flatter areas where calculated flow directions are more sensitive to DEM randomness or interference from bridges, culverts, and vegetation. Careful use of sink-filling (removing areas of the DEM with no outflow), daylighting (removing humps in the DEM along known flow pathways) and burning (lowering DEMs along known flow pathways) can improve representation of flow directions in flat areas. Due to the complexity of DNs, some level of manual/bespoke checking and alteration would be required following automated DN generation procedure, especially within lakes and where engineered structures influence flow directions. The tension between the need for spatial consistency versus using bespoke procedures and data to best represent particular landscape phenomena in DNs is discussed.

The following recommendations are made to guide DN 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 updateable**, 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 updateable**, 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 fill in missing input data, and 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 muti-coloured labelling system should be applied so that a DN can be sub-sampled to be viewed and 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, including version history,** DN users should be able to obtain DNs, including previous versions, from a stable source.

1 Introduction

1.1 The importance of fresh water

Fresh water is a vital supporting element to the wellbeing of land and people because water is essential for life (Dudgeon et al. 2006). Freshwater environments and their states are intrinsically linked with indicators of wellbeing such as ecological integrity[1] and human health[2]. For example, flow of water within and between rivers, wetlands, lakes, and estuaries is vital for maintaining ecological integrity associated with in-stream values such as healthy ecosystems (Baron et al., 2002), basic human health (Gleick 1998), and local customary practices (Stewart-Harawira 2020). River flow is an important driver of various physical, chemical, and ecological states that are in turn linked to ecosystem health and human health including cultural wellbeing, landscape character, recreation, and water supply for human use (Poff et al. 1997; Sofi et al. 2020). Surface water flowing through these environments is a fundamental determinant of their size, structure, dynamics, and chemistry (Zeiringer et al. 2018).

Freshwater environments are influenced by local conditions (e.g., slope, riparian vegetation, bank material) as well as upstream catchment conditions (e.g., rainfall, temperature, geology, catchment vegetation). Freshwater flows, aquatic habitats, and water quality are impacted by various human activities in addition to being influenced by natural processes. Some human activities influencing fresh water are applied locally within freshwater environments (e.g., consumptive abstraction of surface water, manipulation of flow for hydroelectricity production, gravel extraction). Other activities influencing fresh water are applied distally on land within catchments draining towards freshwater environments (e.g., landcover alteration, flood protection works). Regardless of whether they are applied locally or distally, the influence of human activities on freshwater environments tend to propagate and accumulate downstream (e.g., Snelder et al. 2023). The nested hierarchical nature of freshwater systems and the spatially distributed nature of human activities across the landscape combine to create a need for mapped information on freshwater environments describing the feature positions, flow directions, and links with land.

1.2 The need for functional integrated maps of rivers and land

Mapped information about freshwater-land systems is vital to support management, environmental reporting, and policy development relating to freshwater environments. Maps showing the position and interconnectedness of land, rivers, lakes, wetlands, and estuaries are important to support the work of the Ministry for the Environment (MfE), regional councils, iwi, Crown Research Institutes, universities, and others involved in management and research of connected freshwater-land systems.

Simple maps showing the positions of freshwater environments such as centrelines of rivers or outlines of lakes are useful because they show the extent and position of freshwater habitats and resources. Functional maps indicating flow directions within freshwater environments as upstream-downstream connections are an extension to simple maps because they show transport routes within freshwater environments. Integrated maps indicating connections between freshwater environments and their upstream catchments are an extension to functional maps because they

[1] The Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) defines ecological integrity as the ability of freshwater ecosystems to support and maintain ecological processes and a diverse community of organisms (IPBES).
[2] The World Health Organization (WHO) defines human health as a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity (WHO).

show which freshwater environments are likely to be influenced by conditions on which parcels of land.

Nationwide mapped information about freshwater environments is particularly useful for MfE because they are responsible for environmental reporting and resource management to ensure future wellbeing of land and people across Aotearoa-New Zealand (NZ). The national remit of MfE requires mapped information describing freshwater environments to have broad coverage so that maps cover the whole county, and consistent spatial utility so that maps work consistently regardless of location.

1.3 A description of digital networks

From a conceptual perspective, Digital Networks (DNs; also sometimes referred to as digital river networks) are virtual representations of explicit spatial connections across real-world, coupled freshwater-land systems. From a technical perspective, DNs comprise segment data describing surface flow pathways, routing data describing connections between surface flow pathways, and watershed data describing areas contributing to each surface flow pathway.

Segment data can represent different types of surface flow pathways including large rivers, small streams, drainage ditches, or ephemeral channels along which surface water may flow for short-lived periods. Each segment is essentially a line that provides information about the length, sinuosity, aspect, and position within the landscape of a surface flow pathway. When combined with topographic data, segment data also provides information about the altitude and slope of each segment.

Routing data represents the connections between a local segment, its downstream neighbours, and its upstream neighbours. When viewed over several segments, routing data provides information about the possibilities for transport between segments in either the upstream or downstream direction. When combined with watershed and other spatial data, routing data provides information about upstream conditions such as upstream catchment area, upstream average rainfall, or upstream dominant geology. Within a standard DN, routing within each segment can only be in one direction, and bifurcations (where flow from an upstream segment is received by two downstream segments) cannot be represented.

Watershed data represents the area of land whose surface flow directions point towards each segment. Each watershed is essentially a polygon that provides information about what parts of land are associated with a particular segment. When combined with other spatial data (e.g., landcover, rainfall, geology), watershed data also provides information about the local land conditions associated with each segment.

When segment and routing data are analysed together, they provide information on pathways of travel throughout whole catchments. When segment, watershed, and routing data are analysed together, they provide information on connections between surface flow pathways and the surrounding land across entire landscapes. DNs can therefore be used to identify catchments (e.g., which land areas flow to which outlet to the sea), upstream pathways of travel (e.g., distance to headwaters), upstream conditions (e.g., catchment area), and downstream conditions (e.g., distance to the sea).

1.4 The utility of digital networks

DNs are useful analytical tools because they are a fundamental input to classifications, typologies, models, simulations, and quantitative analyses that are subsequently used to inform many aspects of freshwater policy, planning, management, and research (Brown and Pasternack 2019). DNs are key components for many analytical needs, such as:

- Mapping drainage networks (e.g., river catchments) and their defining characteristics such as area, wetness, slope, or Strahler order (a metric of river size and form).
- Characterising pathways taken by water, sediment, contaminants, or other substances that travel in the downstream direction through soil, aquifers, and waterbodies such as rivers, lakes, and wetlands.
- Characterising movements of biota that inhabit freshwater systems (e.g., fish, riverine birds, algae, aquatic vegetation, riparian vegetation).
- Representing the fate and flux of various substances (e.g., water, nutrients, sediment, animals, plants) as they are transported across the landscape within freshwater systems.
- Calculating and displaying river classifications, such as catchment planning units (e.g., Leathwick et al. 2012), ecoregions (e.g., Harding and Winterbourn 1997) or data-driven multiscale geographically-independent classes (e.g., Snelder and Biggs 2002).

1.5 Applications of digital networks

The utility provided by DNs means that they are routinely applied within studies across hydrology, hydraulics, geomorphology, water quality, and ecology despite differences in approaches applied across these freshwater sub-disciplines. Examples of applications that have depended on DNs and that have impacted freshwater policy, planning, or management in NZ include the following.

- Biosecurity modelling and planning, including simulating the spread of invasive species through freshwater environments (Kilroy et al. 2008).
- Climate change investigations, including assessment of the effects of climate change on river habitats, water resources, and river flows (Booker and Snelder 2023).
- Conservation planning, including habitat mapping, species distribution modelling, and identification of high value sites/catchments (Leathwick et al. 2012).
- Environmental reporting, including mapping of water allocation (Booker 2018) and water quality status (Larned et al. 2020) for national environmental reporting.
- Economic assessments and development of policy options, including estimation of the cost of stock exclusion regulations (Ministry for the Environment 2022).
- Geomorphological and sediment modelling, including mapping of surface substrate cover (Haddadchi et al. 2018) and estimation of policy-relevant reference conditions (Stoffels et al. 2021).
- Hydraulic modelling, including hydraulic geometry prediction (Morel et al. 2020) and flood forecasting (Cattoën et al. 2016).

- Hydrological modelling, including assessments of catchment storage (Yang et al. 2020) water resource assessment and delineation of groundwater recharge zones (Singh et al. 2019).
- Local planning, including delineation of water management zones and river habitats (Whitehead et al. 2022) for legislative purposes.
- Water accounting, including assessment of water availability and river flow depletion effects from water abstraction (Booker 2018; Bright et al. 2022).
- Water quality modelling; including calculation of nutrient accumulation and attenuation (Elliott et al. 2016), and assessment of land-use suitability (Snelder et al. 2023).
- Underpinning of statistical analysis to inform on the effects of agricultural and urban land cover on estuarine water quality (Dudley et al. 2020).

Booker (2023) simplified discussion of DN applications by describing two broad categories of purposes for DNs application. The first category was biophysical modelling purposes that can often benefit from detailed spatial representation of surface flow pathways both inside and outside of river channels (e.g., hydrological or water quality modelling). The second category was river management and policy development purposes that do not require information on surface flow pathways outside of river channels (e.g., establishing limits for water resource use, indicating where stock exclusion regulations would apply). Booker (2023) did recognise that biophysical modelling often feeds information into river management. The distinction between the two categories is useful because it demonstrates why different applications for DNs would have different requirements for DN characteristics relating to alignment, coverage, and resolution (see Section 2.1 for more details).

1.6 Existing digital networks

DNs are used by many organisations across NZ. Due to differing needs for differing purposes, numerous local DNs for specific areas exist at different resolutions and extents.

Many regional councils have produced bespoke DNs or altered national DNs for their regions (see report of Booker (2023) for more details). In fact, some councils have produced more than one DN for different purposes. For example, at the time of writing, at least two DNs are publicly available for download from Environment Canterbury (ECan).

There are also several versions of a national DN that have been previously generated by the National Institute of Water and Atmospheric Research (NIWA). The first version of a national DN was produced in association with a national river classification system commonly referred to as the REC (River Environment Classification). The REC is a hierarchical classification that classifies river segments based, in order, on climate, topography, geology, and land-cover factors that control spatial variability in river ecosystems (Snelder and Biggs 2002; Booker 2023). The REC requires a DN to project classes (e.g., warm-dry lowland or cool-wet mountain) onto river segments. The original REC classes (Snelder and Biggs 2002) were calculated onto version 1.0 of the national DN in the early 2000's. Although the national DN has seen several iterations since 2000, the REC classification system has remained the same. The REC and the national DN are widely accessed by users at universities, Crown Research Institutes, unitarity authorities, and in policy generation. Several versions (and sub-versions) of a national DN have been developed and released by NIWA over the past two decades

(Table 1-1). Differences between national DN versions have occurred due to a combination of differing user needs, changes in available input data and changes to technical methods.

Table 1-1: Major version and sub-version history of NZ's national DN (note that various minor sub-versions of each version have also been released).

DN Version	Major Release Dates
1.0	2004
2.3	2008
2.4	2012
2.5	2015
3.0	Most recent

Detailed information about existing national DNs and the REC — including the relationship between the DN and the New Zealand Water Model (NZ-WaM) — can be found in various journal articles and NIWA reports (e.g., Snelder and Biggs 2002; Whitehead and Booker 2019; Shankar et al. 2022; Booker 2023).

1.7 Aim and objectives of this report

Despite the importance of DNs in support of coupled freshwater-land policy, planning, and management, there is no set procedure or agreed methodology for generating a DN. Variability in DN generation is likely to stem from a combination of factors, including: a) differing purposes for which DNs are used, which drives differences in spatial detail and coverage; b) continuing improvements to input data available to generate DNs, such as remotely sensed topography data; c) variety of technical methods, software packages, and user choices that can be used to generate DNs; and d) a variety of institutions who may have need to generate or apply DNs. The interaction of these factors has meant that, despite the analytical advancements afforded by DNs and their widespread use, a commonly agreed best practice for DN generation has not been devised or applied.

The general purpose of this work is to provide information to help make informed decisions about future DN development from a nationwide perspective. The overall aim is to advance methods, technical solutions, and operational delivery of DNs and derived products such as river classifications with a view to future nationwide production and maintenance. There are several objectives which combine to meet the overall aim.

1. Propose principles to guide DN generation in view of:
   - a wide breadth of user needs that stem from DNs being used for multiple purposes;
   - likely ongoing improvements to input data; and
   - the need for spatial consistency associated with national environmental reporting and inter-regional comparisons.
2. Propose technical steps that would constitute best practice for DN generation in alignment with the principles.

3. Investigate options for generating a singular network that can be used consistently across multiple purposes despite varying needs for DN characteristics.
4. Outline options for improved DN functionality such as representation of artificial channels and/or bifurcating channels.
5. Devise a procedure for generating a network that can be intermittently updated when new input data becomes available and discuss the advantages and disadvantages of fully automated DN generation versus manual alteration of DNs.
6. Outline options for testing and/or validating and receiving user feedback on DNs.
7. Assess the utility of the proposed framework for DN generation and maintenance described by Booker (2023) considering technical developments and application to two pilot locations.

DNs were generated for two case study catchments. The purpose of these case studies was not to generate the best possible DNs. Case study catchments were used to demonstrate potential challenges and solutions for DN generation.

1.8 Report structure

The remainder of this report contains several sections. Section 2 outlines overarching methodological considerations for DN generation prior to considering technical methods. Section 2.5 describes a sequence of procedural steps typically applied to generate a DN. Section 3 describes several challenges for generating a DN, including issues relating to representation of real-world phenomena (Section 3.1), input data (Section 3.2), mathematical and algorithmic challenges (Section 3.3). Sections 4 and 5 describe methods and results for two case study catchments. Section 6 provides discussion, including how challenges and solutions for DN generation fit within a previously proposed framework for DN generation. Conclusions and recommendations are provided in Section 7.

2 Overarching methodology

2.1 Defining characteristics

The various applications for DNs listed in Section 1.2 above have varying requirements for DN alignment, coverage, and resolution, which are the main characteristics of DNs as described in Table 2-1. It should be noted that, in theory, there is no trade-off between DN alignment, coverage, and resolution since it is feasible to generate a well-aligned, high-resolution network with national coverage. However, in practice, there is an upper limit to the resolution at which features can be resolved due to computational capacity and data storage/data transfer limitations, regardless of whether a very high-resolution and accurate nationwide digital elevation model (DEM) is available.

Table 2-1: Characteristics that distinguish between DNs.

Characteristic	Description	Example of when the characteristic is important	Example of when the characteristic is less important
Alignment	The positioning of individual river segments and watershed boundaries at the local scale (i.e., when those objects are viewed individually).	Correct alignment (positioning) of river segments is important when calculating streamflow depleting effects of groundwater takes, or for mapping areas where gravel extraction from riverbeds may be occurring.	Correct local alignment or a particular river segment would not be important for large-scale conservation prioritisation purposes.
Coverage	The total land extent covered by the DN, including outlying islands.	Nationwide coverage is important for national environmental reporting or large-scale conservation prioritisation purposes.	Nationwide coverage is irrelevant for local planning purposes where there is only one region of interest.
Resolution	The smallest entity (in terms of segment or watershed) that is represented by the DN.	Information on flow pathways outside of river channels is important for flood flow modelling where overland flow may be influential.	Information on flow pathways outside of river channels is irrelevant for river habitat modelling or stock exclusion assessment.

Strahler stream order is often used to characterise river size. The Strahler system labels stream segments with no other stream segments flowing into them as first order. When two stream segments with different orders join, their downstream segment is labelled with the same order as the highest order of the two upstream segments. When two stream segments with the same order join their downstream segment is labelled with the next highest order than the two upstream streams. The Strahler stream ordering system can be applied in the field, to mapped river lines, or to segments of a functional DN.

2.2 Guiding principles

It is important to state intended end-goals when creating any product or service. A DN, as well as the data and procedures used to develop it, should meet the principles set out in Table 2-2.

Table 2-2: Proposed principles and success measures for generation and operationalisation of a national digital network. Informed by workshops with end-users as described in Booker (2023).

Label	Importance to users	Success measures
Reproduceable	Users want to know how a DN has been generated to assess suitability for their needs.	The methods and data used to generate the DN are clearly described such that the generation process could be replicated independently.
Spatially consistent	Spatial consistency of methods is a requirement by Statistics NZ for national environmental reporting and other users conducting nationwide analyses.	Consistent methods and procedures were applied to generate all parts of the DN regardless of location within the country.
Automatically updateable	Users expect a DN to reflect available data, and data available to assist in generating a DN are expected to be improved periodically (e.g., future LiDAR surveys, improved technologies for surveying, improved mapping of culverts and artificial channels).	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.
Version controlled	Users want to know why changes between versions have arisen.	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.
Bespoke updateable	A broad user community can provide more testing and suggested improvements than can be applied at the time of DN generation.	Users can provide feedback on DN behaviour so that future versions can incorporate potential improvements.
Spatially complete	Many analyses require national coverage (e.g., national environmental reporting, national conservation planning, nationwide cost of fencing, etc.).	The DN covers the whole country.
Functionally correct	Users expect a DN will act as intended in terms of routing.	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.
Functionally informative	Users need a DN to contain a basic set of information expected by applications using the DN.	The DN is accompanied with basic information required for subsequent analysis.
Appropriate	Users want to know whether a DN has been generated with their purpose in mind.	There is a list describing the purposes that the DN was intended to be used for at the time of production.
Available, including version history	Users need to be able to obtain a DN to use it. Users may want to replicate the results they have previously obtained using past versions.	The DN can be obtained by its intended users. Users can obtain previous versions.

A basic set of information required by applications using the DN is mentioned under the principle of ‘functionally informative’ in Table 2-2. We suggest that a DN would typically be expected to be accompanied by at least the following attributes for each segment/watershed.

- Segment length (e.g., for calculating distance to sea, travel time through segment, etc.).
- Segment upstream elevation and downstream elevation (e.g., for calculating slope and average altitude).
- Watershed area (e.g., for calculating upstream area).
- Routing information such that the next upstream segment(s) and downstream segment(s) can be identified (e.g., for calculating upstream area, fate and flux of substances including concentrations and loads, Strahler order, etc.).

2.3 Methodological strategy

Our strategy for devising a network that could fulfil multiple purposes is centred around generating a single network that could be viewed and utilised for different applications. We therefore propose a methodology to generate a high-resolution network comprising objects (segments, watersheds, and routing data) generated using consistent methods and represented using consisted formatting. Our proposed methodology would then label each object using a classification system to indicate sub-sets of objects that are suitable for different purposes because they collectively represent features of the landscape at different resolutions.

2.3.1 A multi-coloured labelling system

Our proposed classification system uses a multi-coloured labelling system to distinguish the use of objects within a DN that would be suitable for use in different applications that require different DN resolutions (as outlined in Table 2-3 and exemplified in Figure 2-1 and Figure 2-2).

Table 2-3: Proposed multi-colour classification system for different resolutions of landscape-scale features to be consistently represented within a single digital network.

Label	Phenomena being represented	Practical description	Example applications
Red	The highest resolution at which surface flow pathways and near-sub-surface pathways can be represented that is compatible with the DN.	Resembles a surface flow direction grid with the same resolution as the input DEM used to generate the DN. Does not require segment line or watershed polygon data because it can be represented by a direction for each grid cell. Contains flow direction information used to derive blue and purple objects.	Fine-resolution grid-based modelling (hydrological, nutrients, sediment etc).

Label	Phenomena being represented	Practical description	Example applications
Purple	Surface pathways that are considered ephemeral because they may experience surface flow intermittently but are not considered river channels. Local watersheds and associated segments are distinguished from blue because different nutrient generation and sub-surface water transport processes are dominant.	When joined with blue, resembles resolution of existing DNv3. Represents possible surface flow pathways during high rainfall events. Purple segments cannot exist downstream of blue segments. Have zero and negative stream orders after they have been coloured.	Flood routing and high-flow modelling. Fine-scale nutrient modelling.
Blue	River channels, lakes, and estuaries. Includes intermittently flowing sections of river channel downstream of permanently flowing sections of river channel.	Resembles resolution of existing TopoMap blue lines. May be less dense than DNv1 and DNv2 in flatter areas. Have upstream nodes that can also be assigned as downstream nodes of purple segments. Blue segments cannot exist upstream of purple segments. Have positive stream order after they have been coloured.	River conservation planning. Estimation of length of fencing for stock exclusion. Analysis of river monitoring site representativeness.
Indigo	Artificial channels engineered for the purposes of altering existing surface flow pathways (would usually not break unidirectional routing requirements of traditional DNs).	Represents amendments to catchment routing associated with engineered movements of surface water. Indigo segments could replace blue segments.	Representation of when a channel has been straightened or a lake/wetland has been bypassed.
Green	Artificial channels engineered for the purposes of water transfer away from or between existing surface flow pathways (including those that would break unidirectional routing requirements of traditional DNs).	Represents augmentation to amendments to catchment routing associated with engineered movements of surface water. Augments network routing between other coloured segments but does not replace any other segments. Would usually intersect with blue segments. Watersheds have no area.	Representation of flow between catchments or segments via canals or raceways for water accounting.
Violet	Artificial channels engineered for the purposes of improved drainage. Could include segments whose alignment can be represented, but whose routing cannot be represented within a DN because it is unknown.	Represents land drainage alteration. Could intersect with purple or blue segments or could be independent of purple and blue segments.	Representation of drainage channels to inform hydrological or water quality modelling.

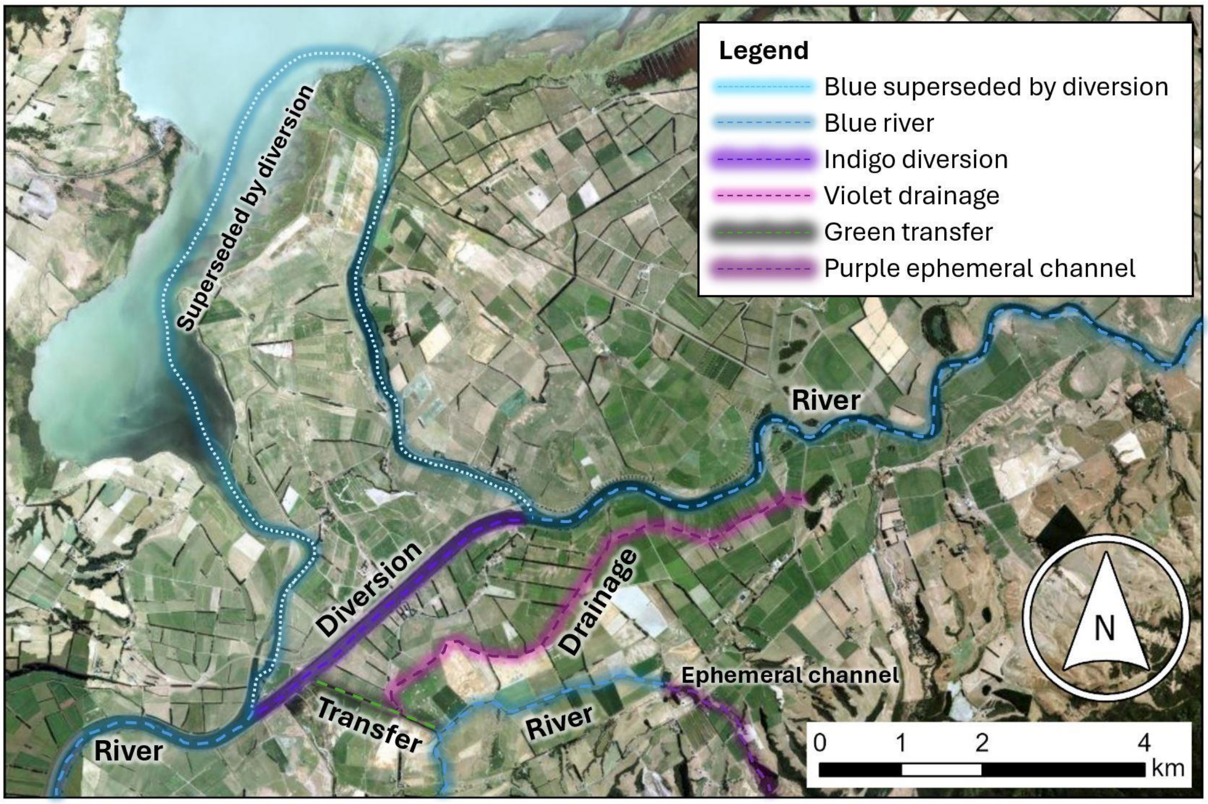


Figure 2-1: Examples of multi-coloured labelled segments. Background map is aerial imagery of the area around Lake Wairarapa and the Ruamahanga River which is bypassed by the Ruamahanga Diversion. Flow direction is generally from top-right to bottom-left.

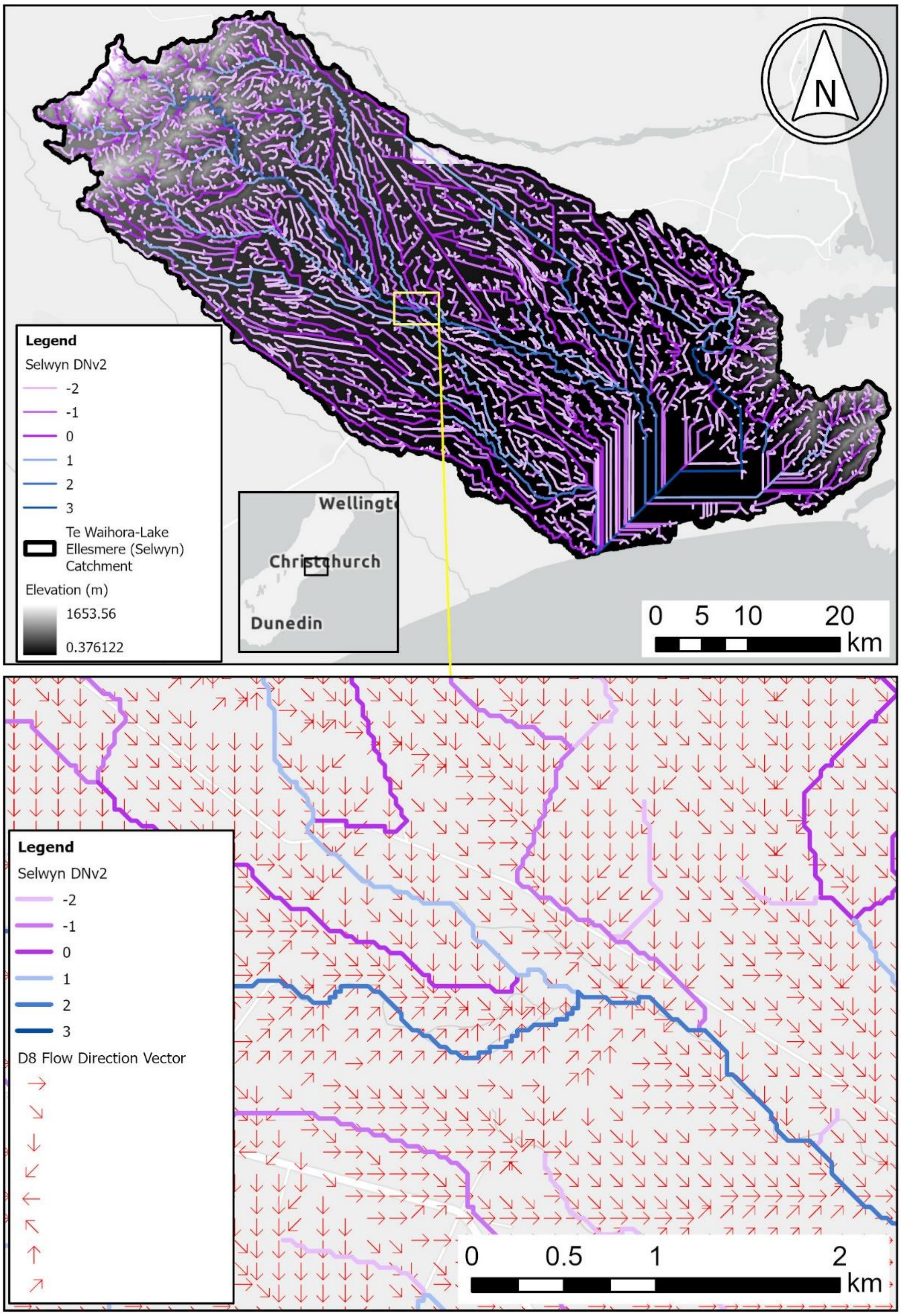


Figure 2-2: Illustration of proposed methodology. The multi-coloured labelling system and negative values for purple segments have been applied to DNv2 for the Te Waihora-Lake Ellesmere (Selwyn) catchment. Map credits: Eagle Technology, LINZ, StatsNZ, NIWA, Natural Earth, OpenStreetMap contributors, Esri, TomTom, Garmin, Foursquare, FAO, METI/NASA, USGS.

The proposed multi-coloured labelling system would match well with the need to service both of the two broad categories of DN application proposed by Booker (2023). Blue labels would indicate objects that should be considered for river management and policy development purposes that do not require information on surface flow pathways outside of river channels. Purple labels would indicate objects that, when combined with blue labelled objects, constitute a DN suitable for biophysical modelling purposes that can benefit from detailed spatial representation of flow pathways both inside and outside of river channels (e.g., nutrient modelling or hydrological modelling). Red labels would indicate objects that contain raw flow direction information used to derive blue and purple objects. Red indicates the finest possible resolution of flow pathway information. Red objects could take the format of a segment and a watershed for each cell of the DEM. However, this format would contain a large amount of redundant data that is more efficiently represented as flow directions for each cell of the DEM.

Indigo objects would replace some blue objects and therefore amend blue routing. Thus, implementation of indigo objects is more complicated than the other colours described in Table 2-3 because indigo objects need to indicate which objects they replace, whereas purple, green, and violet objects just augment other coloured objects.

Green objects would indicate engineered flow pathways away from, or between, natural surface flow pathways. Green objects would supplement, but not replace, the routing represented by blue (and possibly purple). It may be useful to incorporate green segments into water routing (e.g., hydrology, nutrient, or fish movement) models.

The proposed multi-coloured labelling system is easily defined for labelling segments (Table 2-3), but it has some complications with respect to watersheds. When only blue objects of a DN are being considered, blue watersheds would have to be amended to merge with any purple watersheds that they intersect. Merging of purple watersheds with blue watersheds is necessary so that the entire domain is covered by watersheds when just blue segments are being considered.

Green objects would differ to other coloured objects in several ways: a) they may have zero watershed area; b) their flow direction would not necessarily correspond to their slope, especially in flatter areas; and, c) their routing may break with conventions for traditional DNs about unidirectional flow and no bifurcations. It is straightforward to represent a watershed as a sliver with zero area, but the presence of zero or not-applicable (NA) watershed areas for green objects may cause numerical aberrations. For example, an infinity will be produced if a value is divided by a zero-watershed area. Further consideration of how green watersheds are applied in models may therefore be required. Green objects would also necessitate special technical methods with respect to specifying flow direction and treatment of routing.

2.3.2 A negative Strahler stream ordering scheme

The traditional Strahler stream ordering system described in Section 2.1 has often been applied to segments of a DN, and then used when referring to DN resolution. For example, a user might say “we coarsened DNv3 to third order segments” or “we included first order segments of DNv2 to analyse the network at its finest resolution”. However, it is important to note that, if applied to all segments in a DN, results from the traditional Strahler stream ordering system are dependent on DN resolution. Resolution dependency of the traditional Strahler system is demonstrated by the fact that stream orders on national DNv3 are not equivalent in size, position, or number compared to the same stream orders for national DNv2 because DNv3 contains many more segments than DNv2.

We propose a negative Strahler stream ordering scheme that combines with the proposed multi-coloured labelling system to produce comparable stream orders regardless of DN resolution. The standard Strahler stream ordering system would be applied only to blue segments to give the best possible representation of stream order with respect to river channels. The system would then label purple segments flowing into blue segments (or flowing into the sea) as zero or negative, with more negative stream orders belonging to more upstream segments.

Green segments would have no Strahler order. Treatment of green segments with analysis that use Strahler order may require further consideration.

2.4 Types of input data

Spatially continuous elevation data are the minimal data requirement to generate a DN, but several sources of data can be used in DN generation (Table 2-4). Digital Elevation Models (DEMs) are the main input to DN generation. When DEMs have been derived from remotely sensed data, it is important to distinguish DEMs from Digital Surface Models (DSMs). A DSM includes influences of natural and built features (e.g., the tops of buildings, tree canopies, and powerlines) within observations of earth surface elevation. A DEM is the best possible representation of the bare ground after having removed the influence of buildings and vegetation. See Section 3.2 for more details about the technical challenges associated with some of these input data types.

Table 2-4: Types of input data that could be used to generate digital networks.

Data type	Description and use	Data sources	Main traits
Survey data	Point observations of elevation at particular coordinates. Can supplement a DEM.	Field surveys e.g., GPS observations.	Presumed to be accurate. Spatially discrete.
Contour maps	Linear representations of locations on land with the same elevation. Can be converted to a DEM.	Derived from field observations.	Can be inaccurate due to interpolation or extrapolation. Spatially continuous. Resolution dictated by contour height intervals.
DEM	Regular grid representation of elevations. The main input to DN generation.	Post-processing of raw LiDAR data to remove the effects of vegetation and buildings, and they project onto a regular grid. Possibly post-processing of contour maps onto a regular grid. Can be supplemented with survey data.	Often presumed to be accurate but accuracy is influenced by resolution and can be influenced by presence of vegetation and buildings. Defines finest resolution of analysis.

Data type	Description and use	Data sources	Main traits
Mapped river lines	Linear representations of river locations. Can be used to enforce flow pathways within a DEM (through a process known as burning). Can be used to remove the influence of bridges within a DEM (through a process known as daylighting). Can be used to assess alignment of DN segments.	Digitized from mapped lines or survey information (e.g., TopoMaps).	Can be inaccurate. Can become outdated if rivers move. Can be discontinuous, and include bifurcations, braiding etc. Does not explicitly indicate direction of flow. Does not necessarily align with flow directions derived from DEM.
River headwater locations, including springs	The upstream starting locations of rivers. Can be used to inform and/or assess positions of headwater segments.	Can be post-processed from river lines when flow direction is known. Field surveys.	Can be inaccurate, possibly in relation to being subjectively defined. Does not necessarily align with flow directions derived from DEM.
Lake outlines	Polygon representation of extent of lakes. Can be used to improve representation of routing through lakes.	Field surveys. Inspection of maps.	Does not necessarily align with elevations from DEM.
Lake outlets	Point observations of outflows of lakes.	Field surveys. Inspection of maps.	Can be informative for DN routing. Can be insightful if more than one lake outlet is associated with a single lake.
Mapped artificial channels	Linear representations of artificial channel locations. Can be used to amend or supplement DN routing. Can be used to remove the influence of culverts within a DEM (through a process known as daylighting).	Some information in maps (e.g., Open Street Map or TopoMaps).	Difficult to define. Possibly includes a wide variety of features ranging from large canals to small ditches, or underground drainage. Can be inaccurate, possibly in relation to being subjectively defined.

Data type	Description and use	Data sources	Main traits
Ocean bathymetry	Linear representations of locations in the ocean with the same elevation. Can be incorporated into a DEM to avoid missing values around the coast, represent estuaries, and improve routing around outlets.	Ocean bathymetry surveying.	Small-scale variations in ocean bathymetry are likely to be irrelevant for DN generation. Ocean bathymetry just has to be accurate enough to fill gaps around coast locations.

2.5 Procedural steps

In 1997, staff at the Environmental Systems Research Institute, Inc. (Esri)[3] published a methodology for watershed delineation (which involves generation of a DN; Djokic et al. 1997, as cited in Djokic and Ye 2000). The methodology, termed the Fast Watershed Delineation method (FWD), provided a set of standard steps described at that time as "DEM preprocessing that provides the basis for fast (under a minute) and consistent watershed delineation on DEMs of any resolution and size using desktop GIS technology" (Djokic and Ye 2000, p. 66). This methodology built upon earlier work by others (e.g., Mark 1983, 1988; O'Callaghan and Mark 1984; Band 1986; Morris and Heerdegen 1988) and has since been used across many studies that require watershed or DN delineation.

The standard steps established by Esri for watershed delineation are included in Table 2-5. Their methods assume that a DEM has already been acquired for analysis. We supplement the Esri-standard methods with steps outlined by Booker (2023) to include DEM preprocessing steps that influence the accuracy and utility of DNs. We also supplement the Esri-standard methods with steps relating to DN assessment.

[3] Now known as "Esri", Esri is reportedly the global market leader in GIS software. Esri hosts one of the most widely used GIS platforms (ArcGIS and its related platforms) and provides geospatial technology to many organisations worldwide. Esri provides detailed GIS instructional material and support services for GIS users. Because Esri is the GIS industry standard, we draw upon their methodology for generating a DN throughout this review.

Table 2-5: Outline of standard steps for network generation. Adapted from Djokic et al. (1997) and Booker (2023).

No.	Name	Purpose	Stage
1	Collect raw topography data	Gain raw information about topography. For example: survey, contour, or LiDAR data.	DEM preprocessing
2	Create or acquire DEM	Project topography information onto a regular grid. May include removal of topographic effects of bridges or culverts that would otherwise result in falsely blocking of flow along true flow pathways.	DEM preprocessing
3	Remove sinks	Remove influence of sinks (sometimes small and/or false, but sometimes large and true) in topography on flow pathways. Removal of all sinks will ensure that all downstream flow pathways eventually lead to the sea.	Hydrological conditioning
4	Burn river lines	Increase the likelihood that alignment of river segments will match with mapped river channels.	Hydrological conditioning
5	Determine flow direction grid	Intermediatory step to producing segments and watersheds.	DN generation
6	Determine flow accumulation grid	Intermediatory step to producing segments and watersheds.	DN generation
7	Characterise river channels on the flow accumulation grid	Intermediatory step to producing segments and watersheds. Often applies a threshold for the minimum accumulated area that defines channel initiation.	DN generation
8	Convert river grid to river segments	Identifies cells that are characterised as being associated with river segments. Identifies locations of headwater nodes and junction nodes.	DN generation
9	Determine sub-watersheds	Generate a polygon area draining to the downstream end of each segment.	DN generation
10	Vectorise segments and watershed	Generate polyline segments and polygon watersheds.	DN generation
11	Refine network/vector processing	Apply manual interventions to improve network representation.	DN generation
12	Verify network behaviour	Apply mathematical checks of network routing functionality.	DN assessment
13	Initial assessment	Apply internal checks to DN representation prior to release to users.	DN assessment
14	User assessment and feedback	Receive feedback from users and consider possible improvements for future versions.	DN assessment

As Geographic Information System (GIS) technology has advanced, so too have the available input data and digital tools to generate DNs, although the underlying process remains much the same. The naming convention of tools across different GIS interfaces or versions of GIS applications (e.g., ArcGIS, ArcPRO, QGIS, GRASS GIS) has also evolved through time, but, again, the underlying mathematical process remains largely unchanged.

The following sub-sections provide a more detailed description of each step outline in Table 2-5.

2.5.1 Step 1: Collect raw topography data

Topography data is a fundamental input required for DN generation. For the purposes of generating a DN, raw topography data could include bare-earth elevation data collected with LiDAR, structure-from-motion methods, or topographic maps. Collecting topographic data introduces the first source of uncertainty or error: the resolution of topographic data is user-defined, based on factors such as cost, method of data collection, extent of data collection, etc. For LiDAR-derived data, which is the modern industry standard, data precision is influenced by target surface characteristics (reflectance, roughness), acquisition geometry (range, angle), instrument effects (e.g., transmitted energy, aperture size), and environmental effects (atmospheric transmittance, wetness) (Kashani et al. 2015). User decisions regarding the sensor or instrument used, the height from which data is collected, data overlaps, and speed at which the sensor collects data may also be influential. In many cases, post-processed LiDAR is publicly available, so Step 1 can be skipped.

2.5.2 Step 2: Create (or acquire) DEM

Following collection of raw topographic data, a digital surface such as a DEM, DSM, digital terrain model (DTM), or triangular irregular network (TIN) can be built. This step involves processing the raw data (often held as a point cloud) into a grid dataset and includes assigning a coordinate system, cleaning the data to remove gross errors, and choosing a method of interpolation between points to build a continuous surface. Grid resolution (also known as pixel size) is a set characteristic of a DEM. Raw data collection methods and raw point density do not directly determine grid resolution, although choice of grid resolution should be informed by raw data collection methods and selected grid resolution.

It is important to note that raw LiDAR data is routinely post-processed to produce a DEM that represents ground elevations after having removed the effects of LiDAR returns from nonground objects such as vegetation (or buildings). It is ground elevation DEMs that are routinely used to derive DNs. However, information about the location of vegetation (which is often removed from the DEM) could be useful in identifying river corridors because vegetation often grows alongside or within river channels.

DEMs can supplement land elevation information with lake and/or ocean bathymetry. For example, near-shore coastal bathymetry information can be represented within a DEM. Inclusion of coastal bathymetry is desirable for flood inundation modelling (because waves and river flooding can combine to cause flooding), but not desirable for river network generation (because it is unnecessary to know flow pathways that project into the ocean).

2.5.3 Step 3: Remove sinks

“Sink”, “depression” and “pit” are terms often used interchangeably when discussing digital surface grids, but each may have a distinct meaning. Lindsay (2016a) provides the following hierarchical descriptions for these terms (Figure 2-3):

- Sink: “an area of undefined lateral flow or internal drainage” (Lindsay 2016a, p. 847)
- Depression: “bowl-like features, denoted by an area of internal drainage, and completely surrounded by grid cells of higher elevations. The extent of a depression is defined by the elevation of its outlet, also called a spill.” (Lindsay 2016a, p. 847)
- Pit: “a single DEM grid cell that is surrounded by eight neighbouring cells of higher elevation.” (Lindsay 2016a, p. 847)

In this report, unless specifically required, we use the word “sink” when discussing areas of the DEM with no defined onward drainage path.

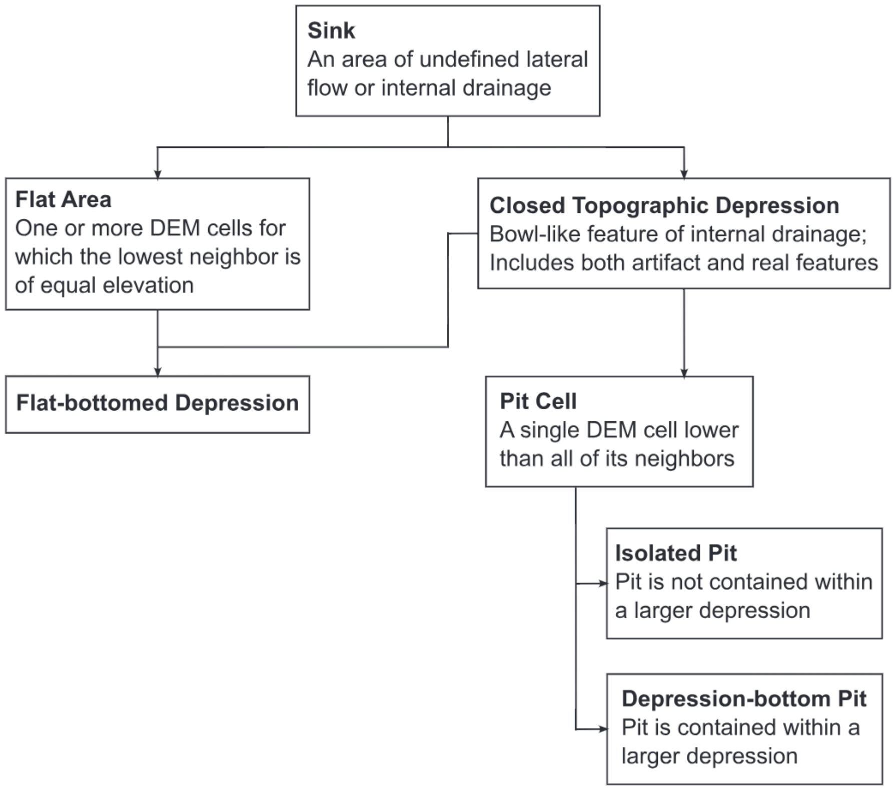


Figure 2-3: Descriptions for terminologies related to areas of undefined drainage in DEMs. From Lindsay (2016a).

Functional DNs cannot be generated from DEMs with sinks because segments that enter sinks cannot be further routed downslope (Lindsay 2016a). Sinks have often been considered erroneous data in DEMs (Tarboton et al. 1991) that need to be filled prior to generating DNs, but they can also represent real-world landscape features where water would naturally drain to groundwater, or where surface water may spill over a saddle in a ridge. Regardless of whether a sink is a data artefact or a real-world feature, the algorithms for developing a DN require the removal of sinks. Filling of sinks often results in a near-to-flat surface across the filled sink (Figure 2-4).

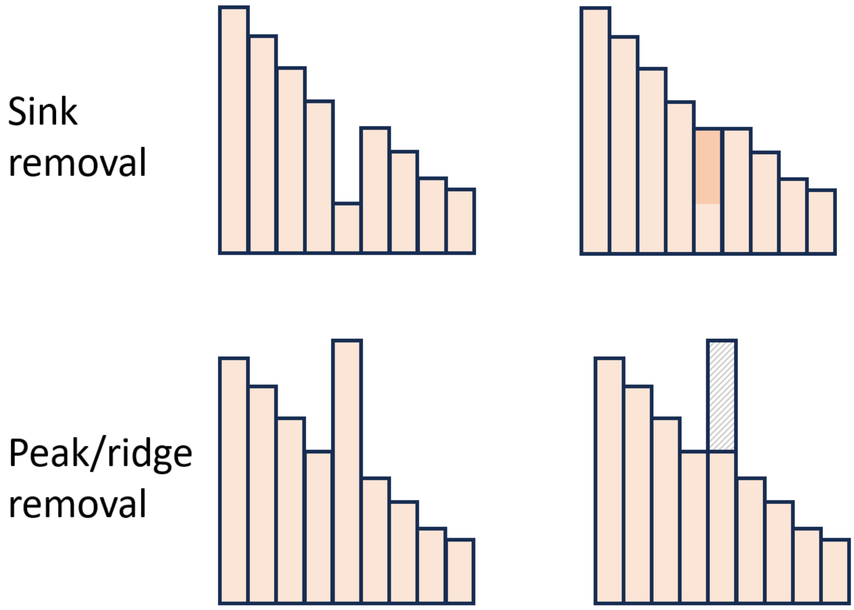


Figure 2-4: Visual representation of removing sinks and peaks/ridges. Modified from Ayad (2021); Esri (2024).

Sink-filling is routinely applied prior to network generation. However, alternative treatments to sink-filling include digging and spilling. Digging involves lowering cells of a DEM so that digital water can escape the sink toward lower elevations located outside the sink. Spilling involves calculating the path that water would take to travel between the centre of the sink and the lowest saddle of the ridge defining the boundary of the sink.

2.5.4 Step 4: Burn river lines

Sometimes, users elect to "burn" river lines into their digital surface. This is an optional step that may be omitted if the digital surface is very accurate and has a high horizontal resolution and precise vertical information. Burning river lines has often been undertaken when DEM grid resolution is relatively coarse compared to river channel width, or where DEM vertical accuracy is insufficient to allow detection of river channels.

Burning river lines involves systematically lowering the height of the DEM in cells that intersect river lines obtained separately from the DEM. Intentionally lowering the height of the digital surface is intended to ensure that subsequent analytical steps (Step 5, Step 6) result in flow pathways whose alignment matches mapped river lines. External data sources representing mapped river lines may come from digitised topographic maps, Open Street Map, or manually digitised vector lines from aerial images. To be effective, burned river lines should flow from source to sea, as burned river lines that do not flow to the sea would cause unwanted sinks in the DEM.

2.5.5 Step 5: Determine flow direction grid

A flow direction grid is generated using an automated algorithm that is often applied via GIS software. A flow direction grid is a representation of the direction of flow for each cell in the DEM grid. There are multiple methods for determining flow direction (e.g., Jenson and Domingue 1988; Qin et al. 2007) but all methods produce an output grid that represents surface flow direction across the grid (Figure 2-5). The flow direction is based on the elevation value of each cell and its relationship to the elevation value of neighbouring cells. For standard flow direction algorithms, cells with higher elevation values will flow towards cells with lower elevation values. Therefore, cells at the bottom of sinks will still have a flow direction, but the cell that they are flowing towards will have a flow direction that directs flow back towards the sink. This behaviour allows cells defining the low points of sinks to be easily identified.

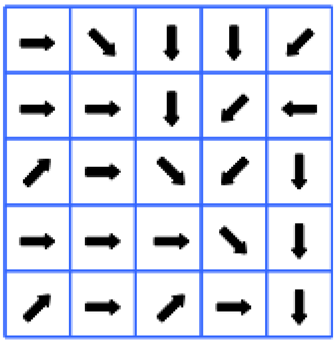


Figure 2-5: Representation of a flow direction grid. Sourced directly from Coggin (2008).

The eight-directional (D8), multi-flow direction (MFD), and infinite flow direction (DINF) algorithms are examples of standard flow direction algorithms. D8 results in flow directions towards one of the eight neighbouring cells. Various conventions can be applied to obtain a flow direction in the event of equal elevations of neighbouring cells (Greenlee 1987). MFD partitions flow away from a cell to all downslope neighbours by creating a flow-partition exponent based on local terrain conditions to determine the fraction of flow draining to each downslope neighbour (Qin et al. 2007). The DINF flow method determines flow direction as the steepest downward slope on eight triangular facets formed in a 3-by-3 cell window centred on the cell of interest (Tarboton 1997). The DINF flow direction output is a floating-point value for each cell representing an angle in degrees counterclockwise from 0 (due east) to 360 (again due east).

2.5.6 Step 6: Determine flow accumulation grid

A flow accumulation grid is determined from the flow direction grid (Figure 2-6). The flow accumulation grid represents the amalgamation of flow from one cell to the next, based on previously determined flow directions. The flow accumulation grid identifies cells of concentrated flow, which can be used in further steps to delineate surface flow pathways and, therefore, a DN.

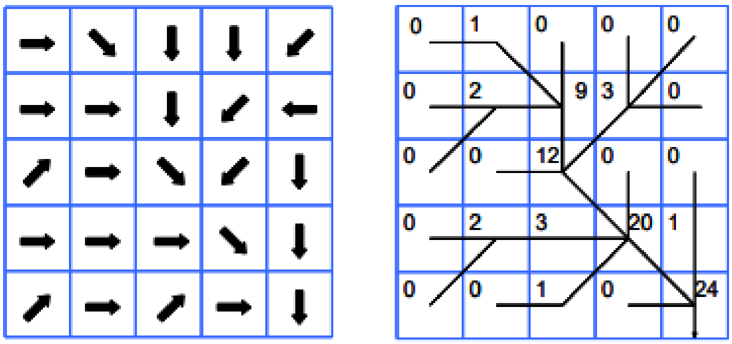


Figure 2-6: Representation of how the flow direction grid is used to generate a flow accumulation grid. From Coggin (2008), with provisional network segments shown.

2.5.7 Step 7: Characterise river channels

A grid of river channels is generated by specifying an accumulation threshold above which cells are characterised as being part of the channel network. Two common approaches to determining channel initiation include the constant area-threshold method (Jenson and Dominique 1988) and the slope-dependent critical support area method (Dietrich et al. 1993). Both methods require specifying a minimum drainage area for a channel. The minimum drainage area should be based on empirically determined relationships for the transition from hillslope processes to channel initiation (e.g., Montgomery and Dietrich 1989). The specified threshold will determine the density of the resulting DN (Figure 2-7).

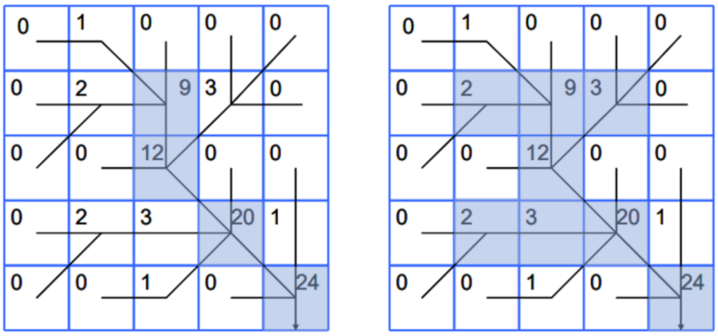


Figure 2-7: Visual representation of DN density as a function of user-specified stream thresholds. Modified from Coggin (2008). A higher threshold returns a less dense DN, while a lower threshold returns a denser DN.

Specifying an accumulation threshold above which cells are designated as channels is a user-defined method that is a common cause of inconsistency between generated DNs.

It is possible to override results produced by applying an accumulation threshold by identifying bespoke locations where the headwaters of river channels have been mapped. For example, it is possible to force segment headwaters to appear at cells that coincide with mapped springs.

2.5.8 Step 8: Identify segments and junctions within river grid

In this step, sequences of cells that collectively represent separate segments of rivers are identified and assigned unique identifiers (Figure 2-8A). A segment is a downstream-pointing sequence of stream cells that (when the sequence is followed in the downstream direction) does not meet with any adjoining river cells with a different headwater cell. Cells where rivers begin and join are called nodes, and the sections of river lines in between nodes represent the segments. In some GIS terminology, river segments are referred to as links; nodes at the downstream end of segments are referred to as junctions; and nodes at the upstream end of segments are referred to as headwaters (Figure 2-8B). Segments and nodes identified in this step remain in grid format. Identification of segments and their associated upstream and downstream nodes is important because this information is required for many aspects of network routing and accumulation.

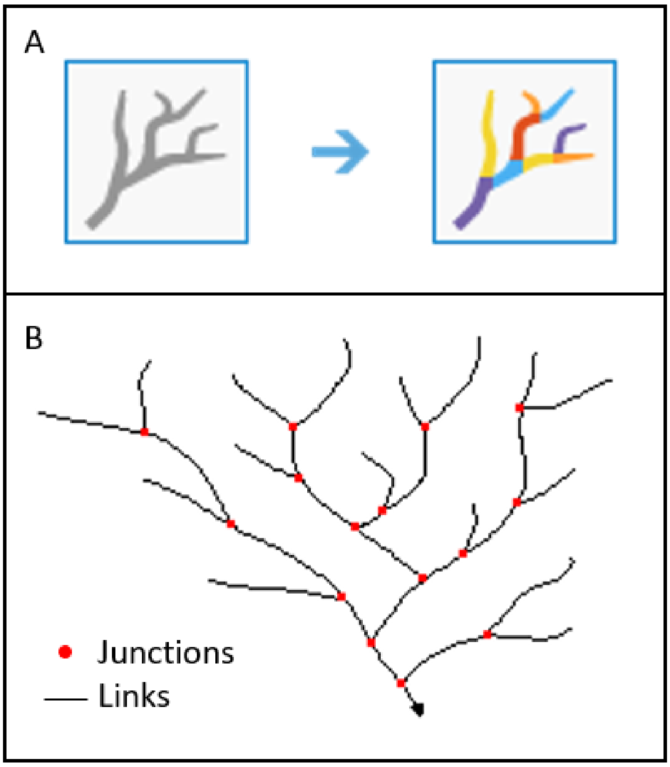


Figure 2-8: Creation of segments and nodes from stream grids. A) identification of individual segments. B) identification of nodes defining interactions between segments.

Several methods are available for identifying segments and nodes. Examples of methods applied in previous studies (e.g., Tarboton and Mohammed 2023) include:

- Channel definition by a global threshold for upstream accumulated area.
- Drop analysis (TauDEM stream definition with drop analysis).
- Slope-area channel definition calculated as slope (S) raised to an exponent (m), then multiplied by the specific catchment area (a) raised to an exponent (n) (i.e., (Eq. 1):

$$S^m a^n >= T \quad \text{(Eq. 1)}$$

- Peuker Douglas stream definition

It should be noted that the outcome of some methods is dependent on the resolution of the input grid. For example, the calculated slope input to the slope-area method will be dependent on the resolution of the input DEM grid.

A brief description of the information required for network routing is as follows.

- Nodes are identified at the upstream end of all headwater segments and the locations defining junctions between all segments.
- Each node is given a unique identifier (e.g., 1 to n).
- For each segment, flow direction and/or flow accumulation information is used to identify an upstream node and a downstream node.

2.5.9 Step 9: Determine subwatersheds

Points represented by junctions (Step 8) can be used to generate a grid of subwatersheds that surround each segment (Figure 2-9). Subwatersheds represent the surface area that drains to each segment as determined by the flow direction of each cell. Subwatersheds determined in this step therefore remain in grid format.

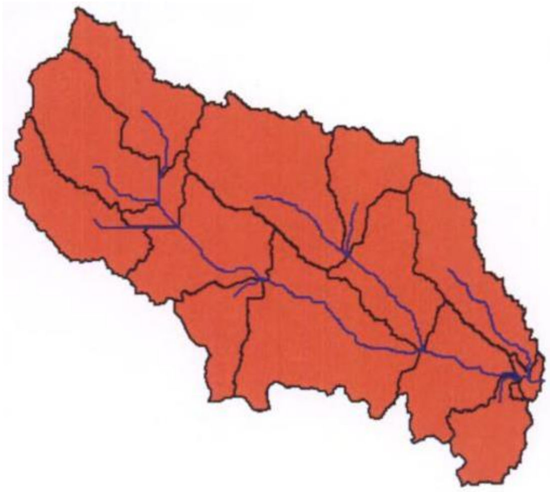


Figure 2-9: Subwatersheds associated with segments of the stream network. Blue lines represent the stream network. Red polygons (with black outlines) represent the subwatersheds that drain to each segment. Directly sourced from Djokic and Ye (2000).

2.5.10 Step 10: Vectorise segments and watersheds

At this stage in DN creation, segments and subwatersheds still exist in a grid form. Sometimes, it is beneficial for the DN components to be converted to vector data (e.g., points for junctions; polylines for segments; polygons for subwatersheds or catchments; Figure 2-10). Vectorised DNs are more easily modified than raster data (e.g., if a segment is known to poorly represent the real-world channel network, it can be manually edited to improve alignment more easily in vector format). Vectorised watersheds can generate more detailed polygon outlines (Lehner and Grill 2013) and vectorised streamlines can provide a more accurate representation of river channels than grid systems, provided that the same number of spatial components (grid boxes or catchments) is used (Mizukami et al. 2021).

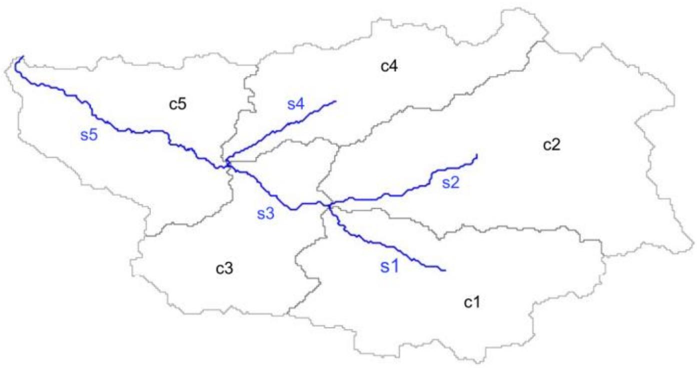


Figure 2-10: A representation of a DN with segments (blue lines labelled "s_") and subwatersheds (black polygon outlines labelled "c_"). Sourced directly from Mizukami et al. (2021).

A benefit of vectorising data is that each segment and each subwatershed will be identified as an individual feature, rather than as a grouping of multiple grid cells with the same values. When vectorised, segments of a river network do not have an associated channel width unless these widths are explicitly specified.

There may be less incentive to conduct this step when a finer resolution DEM has been used because a sufficient resolution of segment alignment is represented within a raster (grid) representation. However, storing segments and watersheds in polygon format rather than raster format may produce smaller output file sizes (depending on resolution of grid and polygons).

2.5.11 Step 11: Refine network/vector processing

With the DN in vector form, manual changes can be made if required. This step is optional, but may be beneficial if additional sources of information could improve the representation of the DN. An example of when manual changes might be made to a DN include when aerial imagery can be used to visualise the location of river channels and river segments of the DN could be re-aligned to better reflect the river channel.

2.5.12 Step 12: Verify network behaviour

After having produced a provisional network, checks can be applied to verify that the network is behaving correctly. If it is assumed that the network is not intended to represent bifurcations, then checks include the following:

- The sum of upstream areas of all terminal segments (segments with no downstream segments) should equate to the sum of all local subwatershed areas in the network domain.
- Each node should only appear once in the column that identifies the upstream node from which each segment is flowing (often referred to as From_node).
- Upstream elevation should be greater than or equal to downstream elevation.
- When calculated using segment length and routing information, distance to sea increases incrementally.

2.5.13 Step 13: Initial developer assessment

After having generated a DN and verified its behaviour mathematically, the makers of the DN will often perform an initial assessment of the DN. Because DNs are used for a variety of purposes, assessments can take a variety of forms and be applied with different levels of rigour. Methods for assessment can therefore range from superficial and subjective (e.g., visual inspection of a part of the network next to mapped river lines or aerial photos) to extensive and objective (e.g., calculation of distances between mapped river lines and DN segments across the entire network).

2.5.14 Step 14: User assessment

DNs can be applied for a variety of applications after they have been released. Users of DNs may apply application-specific assessments of the DN. For example, a hydrological modelling application of a DN may check the magnitude of predicted flows at particular locations or whether mass has been conserved across the model domain (e.g., change in water storage equals rainfall minus evaporation minus outflows). Application-specific assessments may provide useful information that could be fed back to the DN developer.

2.6 Existing framework

A framework for generating DNs and river classifications was outlined by Booker (2023). We mapped the procedural steps outlined in Section 2.5 onto the previously proposed general framework for DN generation (Figure 2-11). Strong correspondence between the procedural steps and the previously proposed framework is indicated by the sequential alignment of the steps within the framework. This suggest that the framework could be applied to generate a DN with a given set of characteristics (see Section 2.1) and therefore meet user needs (see Table 2-2) by conforming to the proposed principles (see Section 2.2).

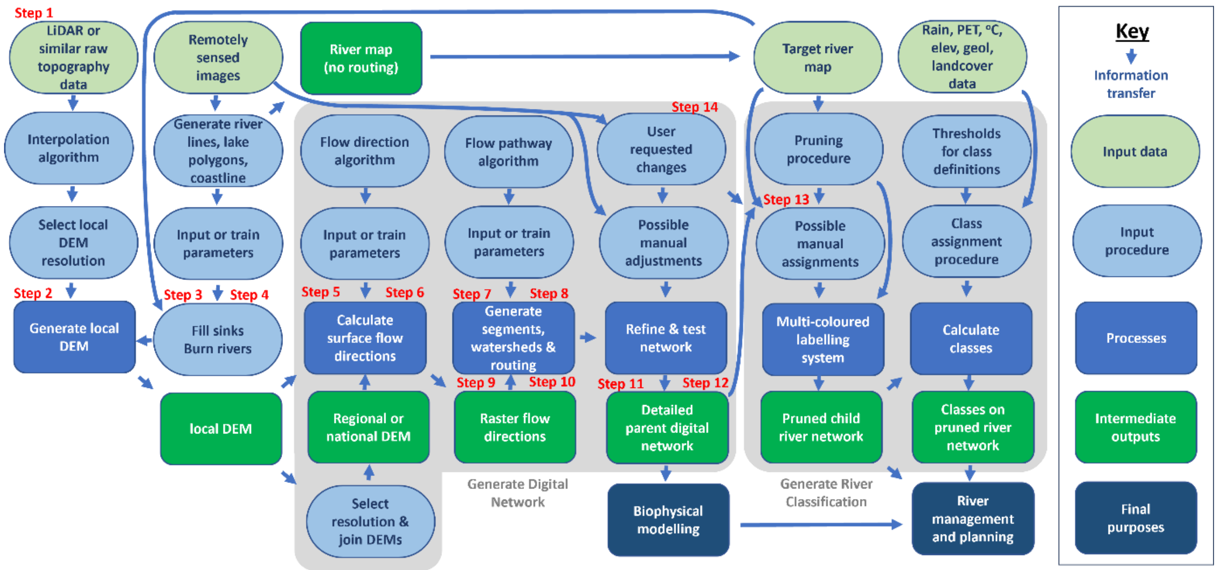


Figure 2-11: Procedural steps orientated in relation to previously devised framework for DN generation. Original framework is from Booker (2023). Procedural steps shown in red text.

2.7 Software tools for implementation/operationalisation

Methods to generate digital networks and consequently advance digital terrain analyses date back to the 1980's (e.g., Mark 1983, 1988; O'Callaghan and Mark 1984; Band 1986; Morris and Heerdegen 1988). Historically, generating DNs was a way to advance cartographical (e.g., field-based) representations of channel networks and watersheds, such as those on topographic maps (Colson 2006; Ehsani et al. 2010; Datta et al. 2022). Today, the availability of detailed digital topographic data (such as DEMs) spanning local, regional, national, and global scales has enabled generation of DNs that often trade off level of spatial detail against breadth of spatial coverage. DNs with global coverage often contain relatively coarse local information (e.g., Giachetta and Willett 2018). DNs with local coverage can contain a very high level of fine-scale local information (e.g., NZ DNv3).

DNs can be generated using a wide variety of GIS or scripting applications (e.g., ArcGIS, ArcPRO, QGIS, GRASS GIS, R, Python). Different GIS applications offer some differences in functionality, but many of the tools and methodological approaches available within these applications are similar.

Table 2-6: Software with DN generation capability.

Software	Description	Availability
ArcGIS Pro	Desktop GIS interface hosted by Esri, the global market leader in GIS software for mapping and spatial analysis.	Subscription
QGIS	Desktop GIS interface supported by Open Source Geospatial Foundation.	Free, open source
GRASS GIS	Geographic Resources Analysis Support System (GRASS), originally developed by the United States Army Corps of Engineers. Used for "vector and raster geospatial data management, geoprocessing, spatial modelling and visualization."	Free, open source
R Programming	Programming language and environment with a wide range of applications, including statistical computing and data visualisation.	Free, open source
Python	Programming language with a wide range of applications, including data analysis and visualisation.	Free, open source
MATLAB	Programming language and environment with a wide range of applications, including simulation modelling and data visualisation. See TecDEM toolbox for network preprocessing (Shahzad and Gloaguen 2011). See GeoNet software (Sangireddy et al. 2016)	Subscription

Some functionality of different GIS applications is the same because the applications act as interfaces to call common base algorithms. TauDEM (Terrain Analysis Using Digital Elevation Models) and SAGA (System for Automated Geoscientific Analyses) are two examples of base tools that can be interfaced from different GIS applications (e.g., ArcGIS, QGIS) to generate DNs. TauDEM (Tarboton 2003) is a collection of tools for extracting and analysing hydrologic information from topographic data. It is a freely available software that includes standalone command line executable programs and a graphical user interface (GUI) toolbox compatible with ArcGIS. SAGA is a GIS software programmed in the C++ programming language that also has a GUI for data management, analysis, and visualisation (Conrad et al. 2015).

It is important to note that combinations of software tools could be used to implement and/or operationalise DN generation. Use of multiple software tools may be beneficial when generating a single DN, however, interoperability between software tools would require conversion of file formatting and naming conventions.

3 Technical challenges and solutions

Procedural steps outlined in Section 2.5 identified some instrument-based, analytical, and user-defined variables that influence the outcome of DN generation. In this section, we unpack the intersection between landscape phenomena, input data, and user-defined algorithms/parameters that can cause variability in generated DNs. Given that DEMs are the typical digital elevation surface used for DN generation (Djokic and Ye 2000), we focus on the input data characteristics and user-defined manipulation of DEMs that can result in variability to DN generation.

3.1 Landscape phenomena

DEMs are often hydrologically conditioned prior to network generation (as discussed in Section 2.5). Global application of DN generation algorithms is theoretically the best way to achieve consistent hydrologic conditioning across an entire DEM domain. However, many different types of naturally occurring geomorphological phenomena (e.g., mountains, valleys, floodplains) are present across landscapes, and various freshwater environments (e.g., rivers, lakes, wetlands) can be represented within a DN. It is thus conceivable that different geomorphological phenomena may be best represented within a DN using different algorithms or parameters for network generation (e.g., sink-filling, burning), and therefore optimal treatment of these phenomena does not always align with globally applied algorithms for DN generation. Globally applied hydrological conditioning may cause some features of the landscape to be lost from, or diminished in, the digital representation of the landscape. Landscape phenomena relevant to DN generation include rivers, lakes, wetlands, springs, and multi-threaded channels/channel bifurcations. The challenges presented by each of these features for DN generation are discussed below. For each phenomenon, we discuss: its definition; why it is relevant to DN generation; why it presents a challenge for DN generation; how it has been handled previously; and options for improved DN generation.

3.1.1 Rivers, including single and multi-channel

There are several contrasting definitions of “river”. One example dictionary definition is a natural stream of water of fairly large size flowing in a definite course or channel or series of diverging and converging channels (dictionary river definition). Rivers can flow on the land surface or inside caves. Rivers can be permanently or temporarily flowing. Rivers can have single or multiple channels. Anastomosing rivers have multiple, interconnected, coexisting channel belts on alluvial plains that enclose floodbasins (Makaske 2001). Braided rivers are characterised by multiple, unstable channel and ephemeral bars formed by intense bed-load transport and a set of very active channel processes (Ashmore 2013). Carling et al. (2014) provide detailed discussion about definitions and semantics of multi-channel rivers. Regardless of their planform, rivers are a fundamental part of freshwater systems.

The main challenges that single channel rivers present for DN generation are precision of segment alignment and location of river headwaters. Precision of segment alignment is important because it influences calculation of segment length, sinuosity, slope. Segments need to be precisely positioned for some DN applications (e.g., calculating distance from river channels to groundwater abstraction points or precisely locating a field site within a DN). Precision of segment alignment also influences the aesthetics of mapped DNs. Location and density of river headwaters is important because they drive DN resolution and therefore river lengths, density of segments, Strahler orders, and watershed areas within a DN. Costabile and Costanzo (2021) provide the following summary of the importance of headwater detection:

"In this context, one of the main relevant aspects is represented by the precise detection of channel heads, which plays a fundamental role for drainage network extraction for a couple of reasons (Li et al., 2020). First, the differentiation between overland runoff on hillslopes and water flow in channels is essential for physically-based models at the catchments scale and for the associated parameters for model simulations. Second, the morphometric and scaling properties of a river network (drainage density, length of drainage pathways, etc.) are influenced by the location of channel heads."

Single channel rivers are the typical landscape phenomenon that DN generation procedures have been developed to represent. Several DN generation steps outlined in Section 2.5 are geared towards incorporating single channel rivers into DNs. In general, DN generation of single channel rivers should be less challenging in hilly areas where segments should be located in valley bottoms, which are well represented in DEMs. Challenges for representing single channel rivers are more likely in flatter areas where calculated flow directions are more sensitive to random noise and the DEM does not represent true flow pathways because of interference from bridges, culverts, vegetation, or other alterations to the land surface.

Multi-channel rivers present a specific challenge for DN generation because standard DNs cannot represent bifurcations (Kleinhans et al. 2019). As a result, braided rivers with multiple channels often appear as single segments in standard DNs because standard methods for DN generation produce networks with unidirectional routing. Therefore, confluences (segments that join as they flow downstream) are represented in standard DNs but difluences (also known as bifurcations, which are segments that split as they flow downstream) cannot be represented in standard DNs. Unidirectional routing causes an inability to represent some types of river planform that would otherwise be included in DN routing. For example, braided channels and islands cannot be represented within a unidirectional DN. See Booker (2023) for more examples of phenomena that cannot be represented within a standard DN due to the unidirectional routing constraint.

Connor-Streich et al. (2018) and Kleinhans et al. (2019) developed a graph theory-based approach and a computational approach, respectively, to determine networks for braided rivers, but neither of their discussions include speculation of how to incorporate flow routing information of braided river segments into catchment-wide DNs. One of the main challenges with representing multiple segments in a braided river relates to routing and calculations that are dependent on routing. Standard DN generation algorithms are not typically able to indicate that two bifurcating segments should have the same upstream catchment area.

We developed a method to demonstrate that it is technically feasible to supplement a standard DN with additional routing-information to represent bifurcations and multidirectional routing. We applied the method using bespoke code written in the R programming language. The method involves creating a set of primary routing pathways derived using standard DN functionality, supplemented with multiple hidden routing pathways. The decision to apply a hidden routing pathway is determined by a function that must pass a condition (e.g., if flow at the upstream segment > x, then apply secondary routing). Primary and secondary routing can be applied simultaneously to represent bifurcations. Multidirectional routing methods would require user input information for each multidirectional junction to be applied. One disadvantage of the multidirectional routing method is that information describing hidden routing pathways is stored as unstructured lists that are difficult to translate to a simpler format (e.g., a single csv spreadsheet).

3.1.2 Lakes

Lakes can be defined as relatively large and permanent bodies of water characterised by slow moving or standing water that is surrounded by land. Richardson et al. (2022) provide additional detailed discussion of definitions and semantics relating to differences between ponds, lakes, and wetlands. Lakes are unique environments that provide habitat and food resources for a diverse set of fish, aquatic life, and wildlife. Lakes may need to be represented explicitly in DNs because they affect transport, mixing, travel times, routing, etc. but behave differently other freshwater environments within DNs, such as rivers.

DN generation is challenged by lakes for the following reasons.

- Lakes do not contain discrete flow directions between inlets and outlets because they route water through polygon areas, which clashes with the requirements for DNs to comprise segments that are lines with no width, fixed position, precise length, and a specific flow direction.
- Mixing of mass/energy/constituents occurs within lakes, which (if the lake contains more than one DN watershed) clashes with the functionality of DNs because mixing does not occur across multiple DN watershed boundaries.
- Lakes are represented as flat areas in DEMs, and DEMs do not have integrated data to identify lake features, which makes it difficult for an automated algorithm to determine coherent flow directions within a lake feature. Near-to-flat areas also cause problems for the requirement for segments to slope downward in the downstream direction.
- Lakes may have more than one outlet, which clashes with the requirement for unidirectional routing within a DN.
- Lakes behind engineered dams (also known as reservoirs) are a special case because lake outflow locations may be altered by human operations (e.g., operation of gates and spillways), which clashes with the requirement for non-transient unidirectional routing within a DN.
- Lake inflows and outflows may be of interest to DN users, but points where lake tributaries intersect with lake boundaries do not coincide with segment confluences, which clashes with the standard practice for DN segments to define the length of river between two confluences.
- Methodological decisions about routing within lakes will influence Strahler stream orders, which clashes with the need for objective calculation of Strahler stream orders.

Segments inside lakes must have valid co-ordinates that can be mapped so that they have valid parameters (length, sinuosity, slope, etc). There are no consequences for routing if segments inside lakes cross each other or the lake boundary. It is visually appealing and mathematically convenient for calculations of watersheds if these lines to not cross each other or the lake boundary. Watersheds inside lakes must have valid co-ordinates that can be mapped so that they have valid areas.

Flow routing through lakes has been represented in previous national DNs by creating segments that can run through, and join within, a lake's boundaries. The watershed boundaries within lakes have been demarcated by the half-way point between neighbouring segments. This method often produces a series of parallel lines running through lakes when segments within lakes were mapped. The advantage of this method is that it does not disrupt models which use the DN because all segments can be treated as surface flow pathways since they have co-ordinates, length, sinuosity, slope, etc. A disadvantage of this method is that segment co-ordinates, length, sinuosity, slope, etc. do not correspond well with expectations about lake hydrodynamics. Segments may be partially inside and partially outside of a lake because they cross the lake boundary. Another disadvantage is that routing through lakes can be expressed in several equally valid, but contrasting, ways that will impact subsequent calculations (e.g., stream order and distance to sea). The DN may also have lacked information needed for models to apply different treatments/methods to segments inside lakes to those outside lakes.

Turcotte et al. (2001) proposed a method for integrating a digital river and lake network (DRNL) alongside a DEM to generate a digital drainage structure that represents lakes in the DN. Their proposed approach was to determine flow directions within lake boundaries using digital river and lake network (DRLN) connections only.

Assuming that lake polygons are available, alternative valid strategies for incorporating lakes into DNs include the following.

- Create a To_node at the lake boundary identifying the downstream end of every lake tributary segment, regardless of segment confluences.
- Label all segments as being either inside or outside of lakes. It is recognised that this labelling incorrectly assumes that lake boundaries are fixed.
- Create a “lake outlet node” away from which a single segment flows downstream of the lake.
- Either:
  - Connect each lake boundary node directly to the lake outlet node. This method has the advantage of simplicity but can reduce downstream Strahler stream order compared to other methods.
  - Create a series of “within lake nodes” towards which each lake tributary segment flows such that they sequentially lead towards the lake outlet node. This method is harder to apply but is preferable for calculation of Strahler stream order.
  - Force flow directions towards the lake centre and then onwards to the lake outlet by artificially adjusting elevations within the domain of each lake. We applied this method using bespoke code written in the R programming language to calculate the distances from each lake cell to the lake boundary and the lake outlet. However, it has been difficult to integrate our bespoke routine into predefined DN generating routines such as available in GRASS.

3.1.3 Wetlands

Wetlands can be defined as environments where the water table is at or near the surface of the land, or where the land is permanently or temporarily (as with the tides) covered by water (Department of Conservation wetland definition). Wetlands can provide valuable ecosystem services because they can filter out nutrients and remove contaminants, reduce flooding, maintain water tables, and return nitrogen to the atmosphere. Wetlands – both natural and constructed – are important landscape phenomenon because they provide habitat for many unique species (e.g., Jenkins and McCauley 2006), they help improve water quality (e.g., Moshiri 1993; Shutes 2001), and they can reduce the impacts of flooding (e.g., Mitsch and Gosselink 2000; Acreman and Holden 2018). Wetlands have implications for network generation from both an environmental and a flood protection or flood modelling perspective.

Wetlands can exhibit similar characteristics as rivers and/or lakes depending hydrological conditions. Challenges for DN generation that are associated with rivers and lakes also apply to wetlands. Wetlands are a particular challenge for DN generation for the following reasons.

- Wetlands can appear as sinks within DEMs that are then removed by a sink-filling algorithms (Jenkins and McCauley 2006).
- By definition, wetlands can transition between dry land, river-type behaviour, and lake-type behaviour.
- Like lakes, wetlands may not have definitive flow directions within their boundaries.
- Unlike lakes, wetlands may not have definitive boundaries.

In previous national DNs, wetlands were not explicitly recognized, and flow pathways were represented as single channels; smaller isolated wetlands may have been ignored due to the application of sink-filling algorithms.

Understanding how wetlands influence real-world surface flow routing and integrating that into DNs may be a crucial step to increasing accuracy of DNs in areas of the landscape with abundant wetlands. It may be useful to identify existing and relict wetland areas (e.g., wetlands that have been drained) by creating a map of grid cells that were "filled" during DN generation. It is possible to associate information about filled sinks with subsequentially generated segments and watersheds so that a DN user can determine whether segments or watersheds may have been associated with wetlands. It is also theoretically possible to create a hidden routing pathway (as described above in Section 3.1.1) that specifies that a segment can terminate under one set of conditions (i.e., not flow downstream as would be the case for water entering a wetland in low flow conditions), and can also flow downstream under another set of conditions (i.e., flow downstream as would be the case for water entering a wetland in high flow conditions).

3.1.4 Springs, aquifers, and karst landscapes

Springs have been defined as areas where a natural discharge of groundwater emerges to the surface (White 2005). Scarsbrook et al. (2007) pointed out that one significant area of confusion in the literature is caused by misuse of the term "spring' when referring to "springbrooks", which are the surface flows leading away from springs. Springs have distinct physico-chemical and biological characteristics, which become modified as groundwater mixes within a downstream surface water body (Barquín 2004). Springs are often associated with aquifers, which have been defined as a unit of rock or sediment that can yield a usable quantity of water (StatsNZ aquifer defintion).

Karst environments are those characterised by extensive underground water systems, comprising a network of caves, sinkholes, springs, underground streams (Ford and Williams 2007). Karst landscapes are typically composed of easily soluble rock types (e.g., carbonate rocks, marble, or gypsum) with well-developed fractures (Ford and Williams 2007). The most well-known karst landscapes in NZ are located around Waitomo in the North Island and the marble country of Mt Owen, Mt Arthur and Tākaka Hill in the South Island.

Springs are a challenge to DN generation because they represent situations where the upstream water source area is not well represented by land surface area and elevations. The upstream area contributing flow to a spring is likely to be more than the local accumulated upstream area calculated from a DEM. Karst landscapes have similar issues, but with the added challenge that surface flow pathways can disappear underground into sinkholes.

Springs are often missing from DNs, including previous national DNs, because accumulated upstream area is a standard input to calculate stream initiation points. Likewise, legitimate sinkholes are often missing from DNs because sink-filling algorithms have been applied and because there is often the assumption and expectation that all flow pathways lead to the sea.

Possible solutions to improving representation of springs and sinkholes in DN generation include the following.

- Data describing point locations of springs could be used to initiate DN segments.
- Data describing point locations of sinkholes could be used to terminate DN segments.
- Spatially varying area thresholds for channel initiation that incorporated DEM and geology data might be used to better detect springs.
- If the majority of flow entering a sinkhole is known to emerge at a spring, then a segment could be drawn through the sinkhole and spring.

3.1.5 Estuaries

Estuaries have been defined as partially enclosed, coastal water bodies where freshwater from rivers mixes with salt water from the ocean (NIWA estuary definition). Tagliapietra et al. (2009) provide detailed discussion about definitions and semantics of estuaries, lagoons, and similar other freshwater-saltwater environments. Estuaries are important because they provide habitat and feeding grounds and for many fish, birds, shellfish, mammals, and other wildlife. Estuaries are also receiving environments for sediment, nutrients, and various pollutants (Kennish 2002). It is useful for estuaries to be incorporated into DNs if users wish to calculate flows or loads into an estuary, or summarise upstream catchment conditions (e.g., Dudley et al. 2020).

Estuaries present a challenge to DN generation because they can have relatively flat slopes and exhibit multidirectional flows. Land-based DEMs may also lack inclusion of topography data for estuaries, which is a challenge for automated calculation of flow directions. One of the most challenging aspects of representing estuaries in a nationwide DN relates to their definition and delineation. For example, some definitions include sounds as a type of estuary (e.g., Dusky Sound, Fiordland), implying that the DN should extend into ocean environments whose topography is well below sea level. Solutions to improving representation of estuaries in DNs therefore centred around tight definitions and decisions about how far DNs should extend into estuary/ocean environments.

3.1.6 Artificial channels

Artificial channels have been defined as waterways or engineered channels built for drainage management, water conveyance, or transport. Terms such as “ditch”, “drain”, “water race”, “raceway”, and “canal” can each have a diverse set of definitions, but they can all be considered as types of artificial channel. Artificial channels are important for some DN applications because they can influence the transport of water, chemical substances, and animals. Here we distinguish between three types of artificial channel:

- Drains are artificial channels engineered to increase drainage towards natural surface flow pathways. Tile drains and storm overflows are types of underground drains. Open ditches are a type of surface drain. Drains are a challenge to DN generation because they are often small enough to not be represented in DEMs. Drains can also have shallow enough slopes to make it difficult to determine their flow direction. Drains may be of interest for some DN applications (e.g., nutrient modelling) but not others (e.g., calculating length of river).

- Diversions are artificial channels engineered for the purposes of transferring water away from, or between, natural surface flow pathways. Raceways that transport abstracted water for irrigation are a type of diversion. Canals that transport water between river reaches or between catchments are types of diversion. Diversions present a challenge to DN generation because they are often associated with bifurcations, which clashes with the requirement for unidirectional routing within a standard DN.

- Realignments are artificial channels engineered for the purposes of replacing natural surface flow pathways. Straightening of a natural river channel is a type of realignment. Realignments would typically not present a challenge to DN generation if they are resolved in the DEM. Realignments may be difficulty to identify due to lack of definitive data that can be used to distinguish natural alignment from realigned channels.

Previous national DNs applied global area-thresholds for segment initiation. A lower area-threshold was applied to generate DNv3 than to generate DNv2. Consequentially, more drains are represented in DNv3 than DNv2. Diversions have largely been ignored in previous national DNs. Realigned channels have largely been incorporated into previous national DNs through burning of river lines.

The method that we developed to supplement a standard DN with additional routing-information to represent bifurcations and bidirectional routing (see Section 3.1.1) was able to represent drains, diversions, and realignments. We developed the method such that additional segments could: a) flow away from the primary network; b) flow towards the primary network; c) flow between segments of the same catchment within the primary network; d) flow between segments of different catchments within the primary network.

3.1.7 Bridges and culverts

Bridges can be defined as structures built to allow transportation over an obstacle such as a river or lake. Culverts can be defined as a structure built to allow water to pass under the surface of something that would otherwise block the flow of water such as a road, railway, footpath, etc. Bridges and culverts differ only in their relative elevations and their relative widths compared to the water they are crossing.

Bridges and culverts are of interest to DN generation because they disrupt the ability to calculate flow directions from DEMs. Relatively high topography in DEMs cause by bridges and culverts may cause sink-filling algorithms to fill in legitimate upstream channels. In previous national DNs, the effect of bridges (and some culverts) on segment alignment was removed from the DN generation process by burning river lines into the DEM.

3.2 Data sources

Sources of data that can cause challenges – but also possible solutions – for DN generation are discussed below. Sub-section 3.2.1 provides some general discussion about generation of DEMs because DEMs are the fundamental input to DNs. Sub-section 3.2.2 lists various non-DEM datasets and describes how they might be used to improve DN generation.

3.2.1 DEM topography

DEMs come in a wide variety of horizontal resolutions and vertical accuracies (Wu et al. 2008). The resolution and quality of DEM data depends on original topographic data (e.g., LiDAR) collection circumstances, including but not limited to, the scan pattern, internal scanner variability, data collection overlap (swath overlap), aircraft altitude during data collection, aircraft instability during data collection, interference due to urban structures, or interference due to vegetation (Petras et al. 2023). The initial data collection returns an irregular point cloud of data (Guo et al. 2010). Differences or inconsistencies in collection circumstances can lead to variations in the resulting point cloud data (Figure 3-1), which may result in artifacts or distortions in products derived from the point cloud (e.g., the DEM). Petras et al. (2023) provide an extensive review of the effects that can result from inconsistencies or variations in original point cloud datasets.

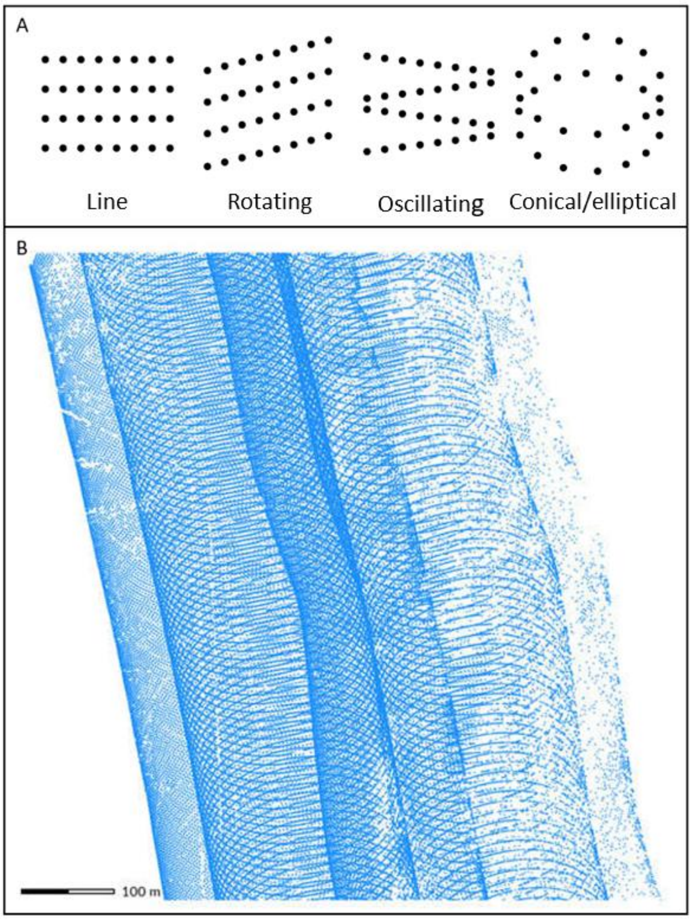


Figure 3-1: Differences or inconsistencies in collection circumstances of LiDAR data. Exemplified by different scan patterns (A) and overlaps (B). Sourced from Petras et al. (2023).

After a LiDAR point cloud has been acquired, a DEM can be built from the data. This step introduces numerous sources of variability due to the variety in methods by which point cloud data can be used to generate a DEM (Caruso and Quarta 1998; Guo et al. 2010; Ajvazi and Czimber 2019). A summary of some of the main interpolation methods are included in Table 3-1. Different interpolation methods produce different results (Heritage et al. 2009; Polidori and Hage 2020). Selection of interpolation method should be based on the characteristics of the initial data points and the intended use of the output (Caruso and Quarta 1998).

Table 3-1: Interpolation methods for creating DEMs.

Interpolation Method	Description	Sources
Natural neighbour (NN)	"Finds the closest subset of input samples to a query point and applies weights to them based on proportionate areas to interpolate a value." (Esri)	Guo et al. 2010 Esri 2024
Inverse distance weighted (IDW)	"Determines cell values using a linearly weighted combination of a set of sample points. The weight is a function of inverse distance. The surface being interpolated should be that of a locationally dependent variable."	Esri 2024
Triangulated irregular network (TIN)	"Partitions a surface into a set of contiguous, non-overlapping triangles…interpolated elevation is the weighted sum of elevations of its surrounding triangle verticals, and the weights are defined as the areal proportions of the sub-triangles to the original triangle."	Guo et al. 2010, p. 703
Spline	"Estimates values using a mathematical function that minimises overall surface curvature, resulting in a smooth surface that passes exactly through the input points."	Esri 2024
Kriging	"Generates an estimated surface from a scattered set of points with z-values."	Esri 2024

Much of the time, users constructing DNs start with a pre-generated DEM. The spacing of data points in the original point cloud can influence decisions about setting the grid resolution of the resultant DEM (Zhang and Montgomery 1994; Li and Wong 2010). DEM resolution can impact calculations of attributes including slope, aspect, curvature and upslope contributing area (Wu et al. 2007). Notably, the effect of DEM resolution on slope and upslope contributing area are particularly relevant to DN generation, as the constant threshold area method (Jenson and Dominique 1988) and the slope-dependent critical support area method (Dietrich et al. 1993) are the two most common approaches for determining digital channel networks from DEMs (Montgomery and Foufoula-Georgiou 1993; Wu et al. 2007). Either of these methods – or others, e.g., curvature-based thresholds (Tarboton and Ames 2001) – can be applied to work out the flow accumulation threshold (2.5.7: Step 7) that a GIS deems as a river. As such, finer resolution DEMs (smaller grid sizes) can provide more detailed topographic surfaces that result in a more detailed DN, while coarser resolution (larger grid sizes) provides coarser topographic surfaces that can reduce the detail of a DN[4]. Wu et al. (2007) and

[4] A 1994 study by Garbrecht and Martz, as cited in Wu et al. (2007) indicated that "a DEM should have a grid area less than 5% of the network reference area to reproduce drainage features with a 10% accuracy."

Polidori and El Hage (2020) provide detailed accounts of the flow-on effects of DEM interpolation and quality assessment methods used in water resource sciences and geoscience.

In Aotearoa-New Zealand, there is a national 8m DEM that was primarily derived from the 2012 LINZ Topo50 20m contour maps (LINZ 2023). This DEM is distributed with a warning that it is “not suitable for analytical purposes.” LiDAR data is becoming increasingly available (Figure 3-2), although with varying degrees of point density (which affects DEM resolution Figure 3-1). When users start their DN generation with a pre-generated DEM, they have no control over the way the DEM was generated. It is important that the user understands how the DEM was generated so that any variability in the DEM can be accounted for in DN generation methods.

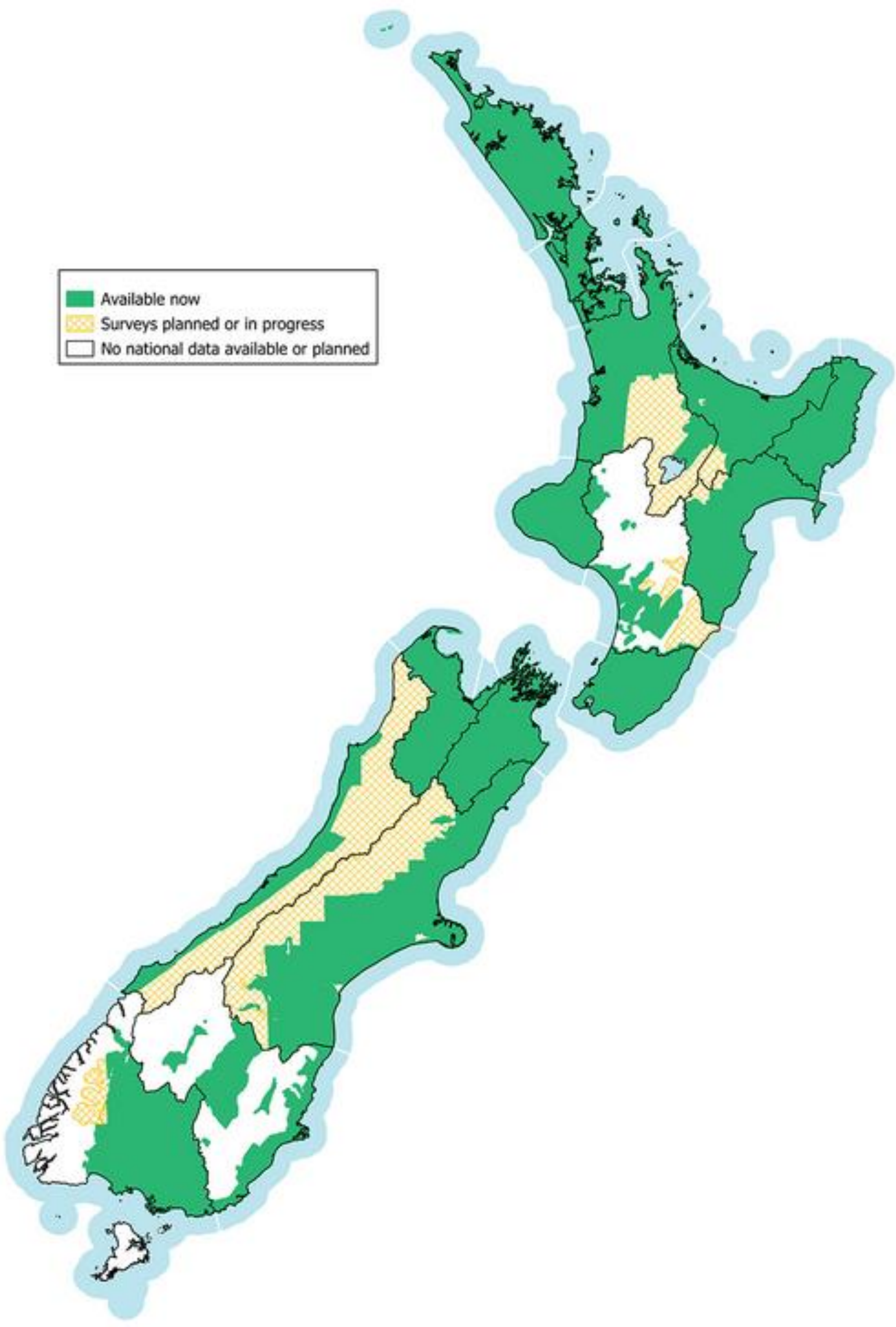


Figure 3-2: Extent of LiDAR coverage across NZ. Obtained from LINZ website May 2024.

At the time of writing, LiDAR-derived DEMs for discrete areas are available via the LINZ data portal (e.g., Canterbury Banks Peninsula lidar 1m). LiDAR-derived DEMs for different area may have been obtained using different procedures. A LiDAR base specification provides a foundation for public sector LiDAR procurements to set minimum standards to ensure LiDAR point cloud data and digital elevation models are suitable for inclusion in the National Elevation Programme.

At the time of writing, "GeoFabrics DEM" data are available from NIWA's Mā te haumaru ō te wai/flood hazard Endeavour programme. The GeoFabrics DEMs are hydrologically conditioned DEMs that combine raw elevation data (LiDAR where available and contours derived elsewhere), natural feature data (e.g., mapped river lines), and engineered infrastructure data (e.g., bridges) as described in Pearson et al. (2023) and in Section 4.2.

3.2.2 Mapping of rivers, springs, lakes, wetlands, and coastline

There are several available datasets with nationwide coverage that could be useful to supplement DEM data when generating or assessing DNs. Table 3-2 describes how various datasets have potential to help solve challenges posed by various landscape phenomena. For example, maps of river/streams/waterways could aid precision of segment alignment and location of segment headwaters in DN generation. The GeoFabrics DEM described in Pearson et al. (2023) and in Section 4.2 was produced by supplementing LiDAR data with some independently mapped data such as waterway alignment and positions of bridges. The workflow for producing the GeoFabrics DEM is an example of how to automatically produce a hydrological conditioned DEM by combining LiDAR and other mapped data. The main challenges for incorporating mapped data into DN generation and assessment are: a) selecting a dataset to apply from several with similar content but different characteristics; b) understanding how errors or uncertainties in mapped data will filter through to the DN generation or assessment; and c) bring together mapping of different freshwater environments such as rivers, wetlands, and lakes; d) maintaining independence between DN generation and DN assessment.

Table 3-2: Nationwide datasets that may supplement DEMs to inform DN generation. LINZ = Land Information New Zealand. MWLR = Manaaki Whenua Landcare Research. GNS = Geological and Nuclear Sciences. LCDB = Land Cover Database. WONI = Waters of National Importance. FENZ = Freshwater Environments of New Zealand.

Dataset and web link	Description	Potential use in DN generation	Notes
LINZ NZ Coastlines and Islands Polygons (Topo 1:50k) LINZ NZ Coastlines	Polygon coastline and islands layer which is based on the Topo50 products.	Define extent of DN segments and watersheds so that inland areas are filled to the coastline.	Somewhat arbitrary inclusion/exclusion of estuary environments, e.g., ToeToes Harbour is ocean whereas Waituna Lagoon is inland.

Dataset and web link	Description	Potential use in DN generation	Notes
LINZ NZ Coastline – Mean High Water NZ Coastline – Mean High Water	Mean High Water coastline including offshore islands at a scale of 1:50,000, and describes the type of coast along the coastline, for example, steep coast, mangrove, or stony shore.	Define extent of DN segments and watersheds so that inland areas are filled to the coastline.	Not necessarily consistent with coastline polygons.
MWLR LCDB version 5.0, Mainland, New Zealand Land Cover Database (iris) or Land Cover Database (Landcare)	Thematic classification of land cover, identifying 33 mainland land cover classes.	“Lake or pond” class to delineate lake polygons. “Built-up Area” class for special treatment of routing expected due to urban drainage (e.g., sewers).	Not necessarily consistent with LINZ NZ Coastlines, e.g., ToeToes Harbour and Waituna Lagoon both listed as “Estuarine Open Water”. “River” class only represents very large rivers. “Gravel and rock” class somewhat coincides with some river braid plains. “Herbaceous Freshwater Vegetation” and “Herbaceous Saline Vegetation” possibly associated with wetland extent.
MWLR NZ LRI Land Use Capability 2021 Land Use Capability	Information about physical factors (rock type, soil, slope, present type and severity of erosion, and vegetation) and ability to sustain agricultural production.	Physical factors to define spatially distributed area-thresholds. “lake” class to delineate lake polygons. “town” class for special treatment of routing expected due to urban drainage (e.g., sewers).	“River” class only represents very large rivers. Lake extent does not correspond with Landcover Database (LCDB) lake extent.
GNS Geology Map of NZ Geology Map	Current mapping of the geology and faults at a scale of 1:250k and 1:1M.	Geological Units 250k or similar to inform/set area-thresholds.	Various types of geological information available. See White et al. (2019) hydrogeological unit type map and similar.

Dataset and web link	Description	Potential use in DN generation	Notes
LINZ River lines NZ River Centrelines (1:50k, 1:250k, 1:500k) Topo50 map series Topo250 map series Topo500 data	River lines as shown on Topo maps at three different level of detail.	River lines to assess alignment of segments. River lines to burn into DEM. Density of river headwaters to inform/set area-thresholds.	Lines do not have directions, watersheds, or routing. Lines do not extend across lakes or wetlands.
LINZ NZ River Name Lines (Pilot) NZ River Name Lines	River name lines for mainland NZ that can be used for searching for a named river.	River lines to assess alignment of segments. River lines to burn into DEM. Density of river headwaters to inform/set area-thresholds.	River lines much denser than LINZ River lines of any resolution. Some drains present but not distinguished from rivers. Names missing for many features. Lines do not have directions, watersheds, or routing. Lines do not extend across lakes or wetlands.
LINZ NZ Lake Polygons NZ Lake Polygons	Polygons delineating lakes, a component of the Topo50 map series.	"Lake or pond" class to delineate lake polygons.	More detailed and numerous lakes than in Land Resource Inventory (LRI) or LCDB. Includes very small ponds that may not have definitive tributaries or outlets.
LINZ NZ Spring Points Spring Points	Points where "water issues from the ground naturally", component of the Topo50 map series.	Segment headwaters bespoke initiated at spring locations.	Spring locations are very sparse.
GNS Location and extent of NZ's aquifers, 2015 Aquifers 2015	Polygon areas known to be above one or more aquifers.	Density of river headwaters to inform/set area-thresholds.	See White (2001). Some boundaries were updated by Moreau and Bekele (2015).

Dataset and web link	Description	Potential use in DN generation	Notes
LINZ NZ Swamp Polygons (Topo, 1:50k) Swamp Polygons	Polygons defined as being 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. A component of the Topo50 map series.	Sink-filling algorithm would be expected to falsely fill in these areas.	Validation unknown.
MWLR WONI/FENZ current wetland layer Current wetlands	Polygons identifying nationally important palustrine and inland saline wetlands.	Improvement or checking of sink-filling. Labelling of wetland objects within DN.	See discussion in Ausseil et al. (2011).
MWLR Wetland area, 1996–2018 Wetland area (MfE & StatsNZ)	Spatial data product which nominally combine FENZ/WONI wetlands with LCDB.	Improvement or checking of sink-filling. Labelling of wetland objects within DN.	See discussion in Dymond et al. (2021). No more "up to date" than LCDB v5.0 despite the later date.
Bespoke spring point surveys (held by regional councils).	Point locations of known springs.	Segment headwaters bespoke initiated at spring locations.	Data not freely available. Anecdotal evidence of piecemeal data through personal communications.
OpenStreetMap	Global coverage maps with various features. Includes community contributed information.	Waterway features to assess alignment of segments. River lines to burn into DEM.	Waterway features do not have directions, watersheds, or routing.

3.3 Mathematical and algorithmic issues

Challenges and possible solutions to generic mathematical and algorithmic issues are discussed below.

3.3.1 DEM extent

DN generation for a catchment requires a DEM that extends beyond the entire catchment of interest, but the extent of a catchment is not known before DN generation. This means that the DN developer (or automated procedure) must clip (delineate and isolate) a DEM that is likely to extend over all parts of the catchment of interest. DN generation algorithms are then applied to all parts of the clipped DEM. If all watersheds of the catchment of interest are contained within the clipped DEM, then the whole catchment has been generated. If any parts of the catchment of interest coincide with the edge of the clipped DEM, then the DEM must be expanded. If the catchment of interest is small, DEM resolution is coarse, or computational resources are plentiful, then the DEM can extend well beyond the boundary of the catchment of interest. If the catchment of interest is large, DEM resolution is fine, or computational resources are scarce, then the DEM must be clipped near to the boundary of the catchment of interest. DEMs are often clipped in regular or rectangular shapes, but this does not have to be the case.

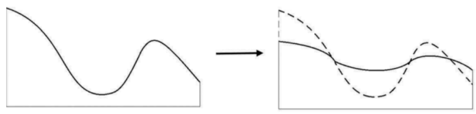
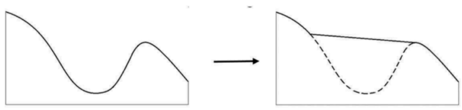
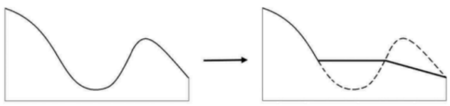
The consequence of needing a DEM to extend beyond the catchment of interest is that DNs are needlessly generated for parts of catchments that may not of interest. Unwanted parts of the DN may then be discarded. Issues about DEM extent relative to catchment extent are relatively straightforward when dealing with single catchments, but are computationally and memory burdensome when dealing with nationwide coverage due to redundancy in DN generation.

3.3.2 Sink removal

Sink removal is a necessary step in DN generation (see Section 2.5.3) but presents mathematical and algorithmic challenges. There are numerous algorithms that have been developed to remove sinks in DEMs, but no single algorithm can appropriately handle different types of sinks (e.g., spurious data versus real topographic depressions; Wang et al. 2019). Lindsay (2016a) and Wang et al. (2019) provide detailed reviews of various algorithms that researchers have developed for sink removal.

One of the original sink removal algorithms was a smoothing filter (O'Callaghan and Mark 1984), which removes shallow and small sinks. Smoothing filters sometimes need to be applied iteratively if the DEM covers a large area (Wang et al. 2019). Jenson and Domingue (1988) proposed sink "filling" to reduce the chance of over-smoothing a DEM (Wang et al. 2019). There are several variants of sink filling algorithms, including the flood-water shedding method of Planchon and Darboux (2002) and the priority-flood method proposed by Wang and Liu (2006). Sink removal algorithms involving the lowering of elevations that acts as dam-points in DEMs are known as breaching algorithms (e.g., Rieger 1998). Soille (2004a) describes a modified sink removal approach to provide a hybrid solution that attempts to minimize the impact of flow enforcement through an optimal combination of depression breaching and filling. Wang et al. (2019) provide a concise representation of digital surface manipulation via smoothing, filling, and breaching (or carving; Table 3-3).

Table 3-3: Sink removal algorithm categories.

Sink removal algorithm category	Cartoon (from Wang et al. 2019, p.87)	General process	Primary challenge
Smoothing filter		Removes shallow and small depressions.	Can remove topographic details by over-smoothing data, especially if multiple iterations are required.
Filling		Raises cells within the sink to produce an area with gentle slopes and reasonable flow directions.	Removes trace of potentially real-world features such as wetlands or ponds.
Breaching (or carving)		Creates descending path from the bottom of the sink to the outlet, then carve the terrain to enable breaching.	Can generate false relationships between carved cells and unchanged neighbouring cells, such as indicating steeper slopes than in reality.

Sometimes, a combination of multiple algorithms with different methods for sink removal can be applied to achieve the best outcomes for DEM conditioning (Wang et al. 2019). As discussed in Section 2.5.3, there are different types of sink characteristics (e.g., sinks with undefined lateral flow or internal drainage, depressions with internal drainage or surrounded by cells of higher elevation, or single-cell pits) in DEM data. Sinks can also be nested within each other. These different types of sinks need to be treated with different algorithms to result in the most suitably conditioned DEM (Wang et al. 2019). Lindsay and Creed (2005) developed an approach called the impact reduction approach (IRA) that applies different sink removal algorithms to different sinks that will have the least impact on a given sink based on its characteristics (Wang et al. 2019). The impact of the algorithms is assessed using the number of DEM grid cells that must be altered under a given algorithm and the mean elevation of the sink before and after removal.

Despite there being a comprehensive literature on different algorithms for conditioning DEMs (Wang et al. 2019), there is no known, formalised documentation covering all algorithms and their most appropriate usages and applications (Qin et al. 2016). Qin et al. (2016) propose a methodology for formalising knowledge related to digital terrain analyses, but the authors recognise that additional research and effort is required. Until a database of the different algorithms and their best usages is documented, it may remain challenging to select the best tools to remove sinks and appropriately condition DEMs.

3.3.3 Segment initiation

As briefly discussed in Section 2.5.7, generating a river grid is a key step in generating a DN. Generating a river grid requires implementing a method that tells the DN-generating algorithm what parts of the digital surface represents a segment, versus what parts do not. This step often involves determining a flow accumulation area-threshold to identify the headwater location of each segment.

Flow accumulation thresholds are important in DN generation because the threshold determines the value above which grid cells are identified as being part of the river network. Different accumulation thresholds effectively determine the resolution of the resulting network because a smaller threshold will create a finer resolution DN. Various methods have been devised for determining river channel initiation values for generating DNs, and all come with some advantages and disadvantages. Thresholds are based on field-determined relationships of channel initiation, and channel initiation in the real landscape varies due to a variety of controlling factors.

Wohl (2018, and references therein) provides a comprehensive review of factors controlling the upstream extent of a river network. In general, the accepted location of where a river begins is called the “channel headwater”, which is “the upstream-most point of concentrated water flow and sediment transport between definable banks” (Wohl 2018, p. 3). Channel headwaters may not always correspond with “stream headwaters”, which is “the upstream-most extent of perennial flow within a river” (Wohl 2018, p. 3). The distinguishing feature of a channel headwater is that it depicts the transition from diffusive hillslope-dominated processes that may generate temporary channels to a persistent, defined channel.

Persistent, defined channels can be dominated by either colluvial or fluvial processes. Colluvial processes are those that transport sediment in the form of weathered material by gravity action. Fluvial processes are those that transport sediment in the form of detrital material by surface water. In colluvial channels, sediment transport and channel formation are dominated by nonfluvial erosion, such as debris flows. Colluvial channels may not always contain persistently flowing surface water, but they may become fluvially active during periods of rainfall. If persistently flowing surface water is present within a colluvial channel, the channel may be altered by fluvial processes, but during less frequent, higher-intensity rainfalls that trigger debris flows, the channel may be subject to substantial modification. It can be difficult to discern the transition between colluvial and fluvial channels in the field, based on the time since the last debris flow and/or the presence or absence of persistently flowing surface water. Making this distinction, however, is important for DNs because some networks may include colluvial channels (either intentionally or unintentionally) while others may exclude them. If the network excludes colluvial channels, then the river network might start significantly downstream of the determined channel headwater.

3.3.4 Burning

Burning involves adjusting the elevations of DEM grid cells that are coincident with the features of mapped river lines or similar hydrography layer (see Section 2.5.4). However, burning of mapped river lines into a DEM prior to DN generation may not be beneficial in some cases. One particular concern is topological errors resulting from the mismatched scales between the river line and DEM data sets (Lindsay 2016b). Burning of incorrectly mapped rivers will overwrite correct information about river location contained in a DEM.

In situations where DEMs (or topographic information used to make DEMs) are relatively coarse compared to river channel size, burning may be required because river channels are not represented within the DEM. In situations where DEMs (or topographic information used to make DEMS) are relatively fine compared to river channel size, burning may be required to “daylight” false blockages to river channels caused by bridges, culverts, or erroneous elevation data.

Burning of river lines that include bifurcations (e.g., islands or braided rivers) may cause mathematical complications for automated generation of standard unidirectional routed DNs.

3.3.5 Multi-directional routing

Graph theory-based approaches have been developed to determine networks for braided rivers (e.g., Connor-Streich et al. 2018; Kleinhans et al. 2019) but the consequences of these approaches for catchment-wide routing are not well developed. When bifurcations are not represented in DNs, there is one definitive routing behaviour that occurs, which we refer to as the standard routing. There are at least two distinctive routing behaviours that may occur downstream of bifurcations, which we labelled as “preserve routing” and “split routing”:

- Preserve routing: both downstream segments inherit the condition in the upstream segment as would be suitable for calculating catchment area or modelling nutrient concentration. This behaviour preserves upstream values but breaks continuity.
- Split routing: downstream segments inherit a proportion of the condition in the upstream segment as would be suitable for calculating flow or modelling nutrient loads. This behaviour splits upstream values and therefore maintains continuity.

It should be noted that that application of preserve versus split routing behaviour would have ramifications for the application of algorithms devised for standard DNs (e.g., calculating upstream area, distance to sea, Strahler order). For example, the standard algorithm for calculating upstream area for a standard DN is applied incrementally in the downstream direction by totalling the local area and the upstream areas of all (usually two) immediately upstream segments. The standard calculation of upstream area would not produce correct results under either preserve routing (correct either side of an island but double counts downstream of an island) or split routing (incorrect either side of an island but correct downstream of an island). The logical solution to calculate upstream area with bidirectional routing is to either: a) for each segment individually, find and sum all upstream local watershed areas; or b) apply the Standard method to the entire DN first, and then apply the Preserve method by summing local area with the immediately upstream areas found by the Standard methods.

4 Methods

In this section, we describe a suite of methods that we applied to demonstrate potential challenges and solutions for DN generation.

4.1 Case study catchments

Catchments referred to as Te Waihora-Lake Ellesmere and Te Awa o Mokotūāraro were selected to demonstrate DN generation methods. These case studies were chosen in consultation with MfE and the relevant regional councils (ECan and Hawke's Bay Regional Council; HBRC). Together these two catchments exhibit several characteristics that are challenging for DN generation:

- Both catchments contained reasonable diversity of landscape setting, including steeper mountain/hills, flatter lowland areas.
- The catchments contained interesting lake and/or estuary configurations.
- The catchments contained some areas with natural landscapes and some areas where human engineering has altered surface water drainage.
- The catchments contained sufficiently large catchment areas to test computational time and resources associated with DN generation.

Te Waihora-Lake Ellesmere is located in Waitaha-Canterbury, Te Waipounamu-South Island. Its upstream catchment contains the Selwyn River, Harts Creek, and several other tributaries to Te Waihora-Lake Ellesmere.

Te Awa o Mokotūāraro is located in Te Matau-a-Māui-Hawke's Bay, Te Ika-a-Māui-North Island. It flows across the Heretaunga Plains, contains the Ngaruroro River and Karamu River, and several other tributaries. Te Awa o Mokotūāraro was formerly known as the Clive River.

4.2 Input data

We sourced input "GeoFabrics DEM" data from NIWA's Mā te haumaru ō te wai/flood hazard Endeavour programme as input to the DN generation process. GeoFabrics DEMs are hydrologically conditioned DEMs that combine elevation, natural feature and infrastructure data (Pearson et al. 2023). GeoFabrics is a processing package that has five processing stages, including: "unconditioned DEM generation, riverbed interpolation, riverbed estimation, open and closed waterway inclusion, and, finally, hydrological conditioning" (Pearson et al. 2023, p. 3). Detailed information about each of these processing stages can be found in Pearson et al. (2023). The GeoFabrics DEM resolution was predetermined to be 8 m. We understand that 8m was chosen to represent the best trade-off between required detail/processing speed for hydrodynamic modelling versus false detail and high computational requirements.

Table 4-1 summarises the key components of each stage used to create the GeoFabrics DEM. Daylighting mentioned in Table 4-1 is an algorithmic procedure that attempts to remove the unwanted effect on flow direction of bridges and culverts on a DEM. Daylighting is a process that maintains hydraulic connectedness of mapped waterways within the DEM by lowering cells that are crossed by waterways but are higher than the ends of the mapped waterway. Waterways defined in Open Street Maps with labels of "rivers", "streams" and "drains/ditches" were daylighted within the GeoFabrics DEM. If daylighting was not applied, then bridges and culverts would block flow paths along flow pathways derived from the DEM. Quite severe daylighting had been applied to the DEM we used. Because the daylighting was applied for the purposes of hydraulic modelling of floods. In the previously applied process, drains/ditches were daylighted to 8 m, streams to 16 m, and rivers to 32m. The source of information for each DEM cell was identifiable from labels shown in Table 4-2.

Table 4-1: Processing stages in the GeoFabrics workflow (summarised from Pearson et al. 2023).

Processing Stage (Pearson et al. 2023)	Summarised Explanation
Unconditioned DEM generation	Converts one or more classified point clouds (LiDAR or sonar) into a grid of elevations through data filtering and averaging. Each grid cell is square, with its width equal to that of the DEM resolution. The NZ 8 m Digital Elevation Model (2012), a nation-wide coarse DEM available on the LINZ data service, is used where there is no LiDAR data.
Riverbed interpolation (interchangeable with riverbed estimation, below)	Incorporates a dataset of surveyed cross sections of riverbed elevation and interpolates riverbed elevations between those cross sections. Requires information on the location of riverbanks.
Riverbed estimation (interchangeable with riverbed interpolation, above)	For use in river areas where no elevation data (e.g., point cloud bathymetry or surveyed cross sections) is available. Estimates riverbed elevation based on hydraulic equations. Requires information about bank-full flow, bed roughness, and river centrelines.
Open and closed waterway inclusion	Analogous with daylighting, this stage removes obstructions along flow paths in areas where flow, roughness or channel geometry is unknown. Utilises waterway features from Open Street Maps (a global, community driven open source dataset).
Hydrological conditioning	Combines the unconditioned DEM with any available riverbed and waterbed elevations, and ocean bathymetry data (from the other processing stages or from externally sourced data). Generates layers as riverbed, waterways, ocean and unconditioned DEM for amalgamation into a single, hydrologically conditioned DEM.

Table 4-2: Labels associated with each cell of the GeoFabrics DEM. After Pearson et al. (2023).

Label	Description	Comments
-1	‘no data’	Indicates missing data as indicated by NA in elevation data.
0	‘interpolated’	On land or in ocean (where ocean bathymetry has been interpolated).
1	‘LiDAR’	LiDAR data were available to derive a height for the cell.
2	‘ocean bathymetry’	Derived from ocean contour data
3	‘rivers and fans’	Derives from OpenStreetMap labelling.
4	‘waterways’	Derives from OpenStreetMap labelling. Include various types of “waterways”, including “rivers”, “streams” and “drains/ditches”
5	‘coarse DEM’	Digital Elevation Model (2012) sourced from LINZ.

Inspection of the GeoFabrics DEM revealed some interesting characteristics within the catchment upstream of Te Waihora-Lake Ellesmere. Some parts of the DEM contained more detailed representations of elevation than others because they were sourced from different data sources (Figure 4-1). For example, relict channels, roads and possibly some empty water storage ponds can be seen in the bottom-left of Figure 4-1 whereas fewer details can be seen towards the top-right of that figure. Elevations within Te Waihora-Lake Ellesmere itself were uniform (exactly 0.5 m implying they were filled manually) but were bordered in some locations by lower elevations at the lake margins (Figure 4-2). Elevations lower than 0.5 m at lake margins may have occurred because LiDAR data were collected during a time of relatively low lake levels or because the lake water level is actually lower than specified in the DEM.

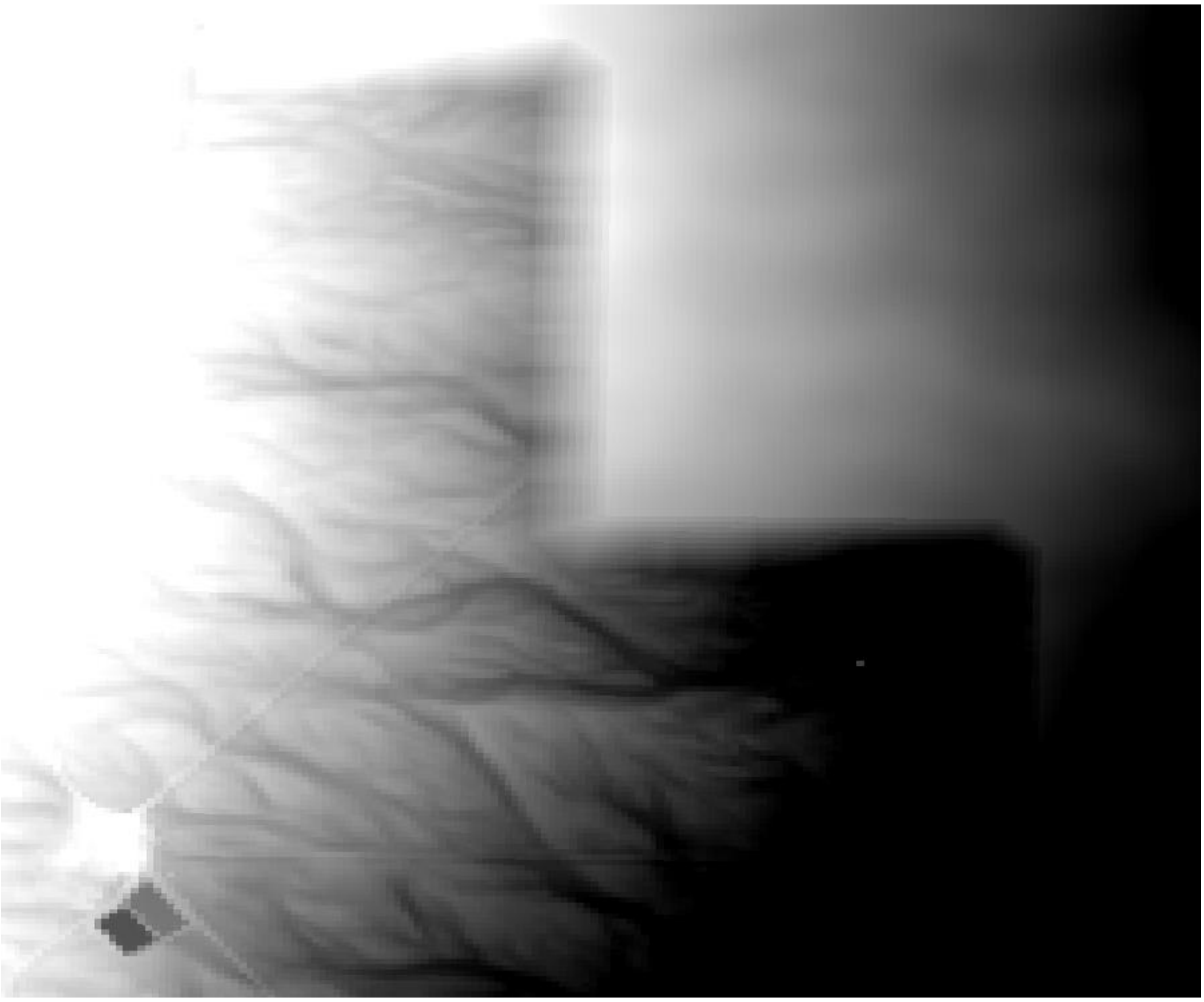


Figure 4-1: GeoFabrics DEM illustrating difference in representation of topographic detail within Te Waihora-Lake Ellesmere. Elevations clipped for illustrative purposes. Darker shading are lower elevations.

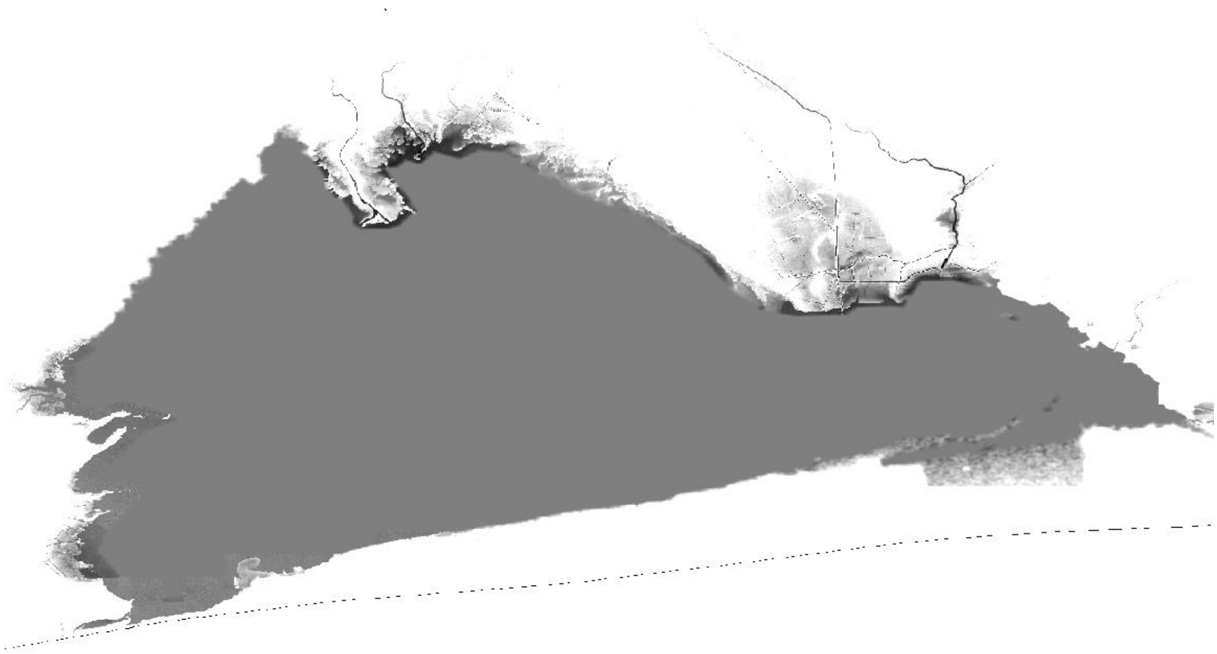


Figure 4-2: GeoFabrics DEM illustrating topographic details around within Te Waihora-Lake Ellesmere. Elevations clipped for illustrative purposes. Darker shading are lower elevations.

GeoFabrics DEM includes elevation data that extends into the ocean. Inclusion of ocean bathymetry in a DEM is beneficial for hydrodynamic modelling around estuarine and coastal because floods are often influenced by tides and river flows. Inclusion of ocean bathymetry in a DEM is also beneficial for DN generation because the occurrence of missing data is reduced. However, inclusion of ocean bathymetry creates difficulties for automated algorithms when locating catchment terminal segments representing outflows to the sea. DNs would extend across the ocean floor if ocean bathymetry were included in a DEM. We clipped the DEM data using data from LINZ NZ Coastlines and Islands Polygons (Topo 1:50k) with a 10 m buffer.

4.3 Generating a network from river lines

We downloaded publicly available set of river lines from Environment Canterbury on 9 Nov 2023 (ECan River Network). This ECan River Network is described as a representation of the location of the main rivers, streams, and drains in Canterbury. We read the downloaded shapefile, which covered all of Waitaha-Canterbury, into the R programming language, inspected the data, and converted the lines into a functioning routed set of segments. In this case routing functionality meant that all segments must take one route downstream to the sea, and bifurcations and diversions could not be represented. The downloaded network contained 5151 segments before manipulation. We took the following steps to investigate processes required to convert river lines into a routed network. The numbers listed below refer to segment identification numbers within the 5151 starting segments.

1. Inspect direction of segments and confirm that the order of segment co-ordinates is starting at upstream and ending at downstream (this was mostly the case).

2. Identify segments that represent loops because the segment start point is in the same place as its end point. Remove these. Segment numbers 4778 and 586 removed automatically.

3. Identify segments those geometry duplicates and other segment's geometry. Remove these. Segment 4692 removed automatically.

4. Identify segments that can be removed from the network based on visual inspection (e.g., because they are canals or isolated segments that join elsewhere to the network). The following segments were removed:

   A. 3108 Ōhau canal.

   B. 1040 tiny stream running into Ōhau canal.

   C. 743 tiny stream running nowhere.

   D. 2973 flowing north from Waimakariri, not a river?

5. 2394 flowing south from Eyre, not a river?

6. Identify segments whose direction needs to be reversed because starting node is downstream and ending node is upstream. Reverse order of these. 4968, 4170, 2919, 5045, 2618, 2059, 4601, 4864, 1915 directions reversed.

7. Identify segments which have another segment joining somewhere within its length. Split these segments based on where they coincide with the interceding segment. 3855, 3957, 3389 were split at their intersection with 2300, 1757, 242 respectively.

8. Identify pairs of segments that only have one upstream and one downstream segment such that they could be joined to be one continuous segment. 5060 and 4159 joined with 4601, 2590 respectively.

9. Use start point as fnode (from node) and end point as tnode (to node) for routing.

10. Identify all segments whose tnode does not match with any fnodes. Label these segments as terminal nodes. Terminal segment should be located around the coast, unless there is a legitimate inland terminal segment (e.g., inflow to Tarn with no outflow). Use output to feedback to inform Step 4 (removed segments), Step 5 (segment direction reversing), and Step 6 (splitting of segment by interceding segment).

11. Identify segments whose fnode matches with more than one tnode. Label these segments as bifurcations.

12. Identify dominant pathway through a bifurcation. Amend fnode of sub-dominant segment(s) flowing away from bifurcations to be unique to all existing tnodes and fnodes. Essentially, this method breaks the routing that caused the bifurcation. The current default method is to assign the shortest path to the sea to be the dominant flow pathway.

13. Identify any bifurcations for which dominant and sub-dominant pathways should be switched. The following bifurcation dominance was switched to demonstrate functionality.

    A. 2233-153 legitimate Rakaia bifurcation. 153 (south branch) should be the dominant pathway.

    B. 3232-1705 small tributary should join Ashburton 1705 should be the dominant pathway.

    C. 1336-878 latter looks more river-like than straighter drain.

    D. 566-316 latter looks more river-like than straighter drain.

We developed an interactive R Shiny app to map the network, locate bifurcations, and interactively inspect routes to the sea (Figure 4-3). We found that it was relatively straightforward to create set of routed segments from the ECan river lines. Artifacts found within the ECan river lines included the following.

- The majority of segment data represented the upstream-to-downstream order, which greatly aided conversion of mapped lines to routed segments.
- Some duplicated segments mapped on top of each other.
- Several bifurcations associated with either illegitimate routing or legitimate routing because of islands, diversions, or splits.
- Noticeable differences in segment density in different parts of the landscape, with relatively few segments running across some, but not all, parts of the Canterbury Plains.
- Relatively few segments appear to represent “drains” despite the provided description of the data, although we acknowledge difficulty in defining rivers, streams and drains.

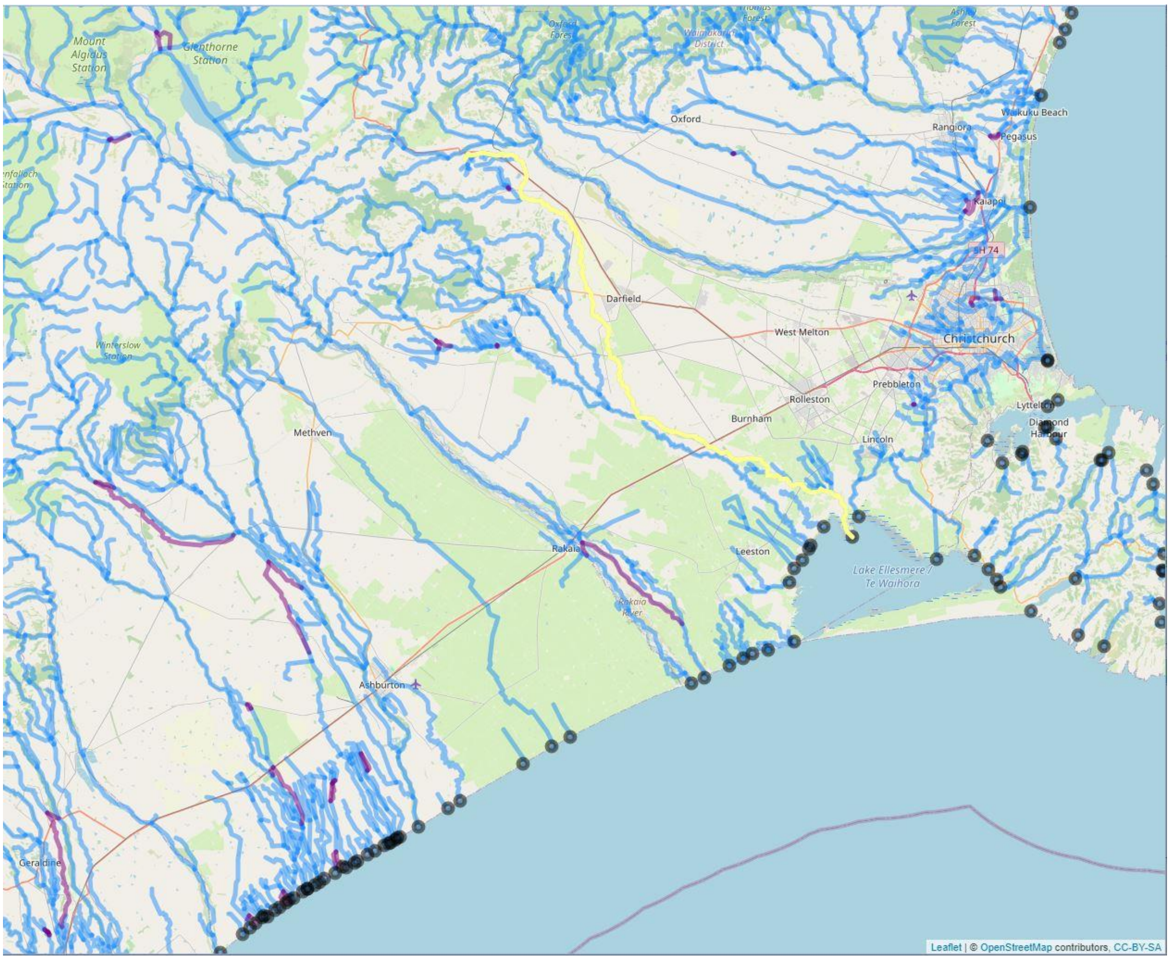


Figure 4-3: River lines that have been converted to a routed network. Blue lines are routed network. Yellow line is a trace to the sea. Purple segments are bifurcations. Black circles are terminal nodes.

4.4 Generating a network from a DEM

It was impractical for us to investigate the effects of all possible combinations of input data (e.g., DEM resolutions), methodological choices (e.g., choice to burn river lines or not burn river lines), and parameter choices (e.g., are-threshold for segment initiation of segment headwaters) on DN generation. We therefore devised a default combination of input data, choices, and parameters to apply Steps 3 to 10 outlined in Section 2.5. Four DNs were newly generated for each of the two case study catchments. Three DNs were generated using GRASS and one DN using TauDEM using the following routines.

- GRASS networks. Two routines, r.watershed and r.stream.order were used to generate three networks, only varying in the upstream area-threshold for segment initiation values (100, 1000, and 10,000). The exact command to make GRASS (10000) were:

```
r.watershed -b elevation=DEM.tif \
 accumulation=accum.tif threshold=10000 \
 drainage=direction.tif \
 stream=stream.tif basin=wshed.tif
 r.stream.order stream_rast=stream.tif \
 direction=direction.tif elevation=DEM.tif \
 accumulation=accum.tif stream_vect=river
```

The three networks are labelled GRASS (100), GRASS (1000), and GRASS (10000).

- TauDEM network. Five commands, PitRemove, D8FlowDir, AreaD8, Threshold (with 1000 for upstream area segment initiation), and StreamNet, were used to generate a network. The exact commands were:

```
PitRemove.exe -z DEM.tif -depmask depmask.tif -fel
DEMfel.tif
 D8FlowDir.exe -fel DEMfel.tif -p pfile.tif
 AreaD8.exe -nc -p pfile.tif -ad8 ad8out.tif
 Threshold.exe -thresh 1000 -ssa ad8out.tif \
 -src stream_raster.tif
 StreamNet.exe -fel DEMfel.tif -p pfile.tif -ad8
ad8out.tif \
 -src stream_raster.tif -ord sorder.tif -tree tree.txt
\
 -coord coords.txt -net reach.shp -w watershed.tif
```

We also accessed previous national DN version 2.5 (DN2) for both case study catchments.

We accessed previous national DN version 3.4 for the catchment upstream of Te Waihora-Lake Ellesmere and version 3.6 for the catchment of Te Awa o Mokotūāraro. These versions of version 3 (DN3) represent the best available DNs for the catchments of interest at the time of our investigation. DN2 and DN3 were previously generated using similar procedures but different input data and a different area-threshold. DN3 contains many more segments than DN2 because the area-threshold used to generate DN3 was less than the area-threshold for DN2. DN2 and DN3 were clipped to the same DEM area that were supplied to the GRASS and TauDEM DN generation processes. This is problematic because it results in far more terminal reaches than one might expect or want since segments are generated outside of the catchments of interest. However, for an illustration of the process of assigning attributes this will suffice.

4.5 Characterising a network using metrics and attributes

Due to the multistep generation process, with different methodologies and input parameter choices within those steps, numerous wildly varying river networks with the same coverage can be produced. It is useful to assign attributes to a network to allow a quick assessment of DN characteristics (as set out in Table 2-1), and also allow comparisons between DNs.

4.5.1 Blue Line Initial Stream (BLIS)

The multi-coloured labelling system (Section 2.3.1) and the negative Strahler stream ordering scheme (Section 2.3.2) both depend on identification of headwater nodes for blue segments. Blue headwater nodes represent the transition from purple to blue and from zero-order to first-order. We applied the multi-coloured labelling system and the negative Strahler stream ordering scheme to our networks using the newly developed concept of Blue Line Initial Stream (BLIS).

Application of the multi-coloured labelling system would require an area-threshold to identify headwaters of purple segments and a BLIS area-threshold to initiate blue segments. The area-threshold for purple segments should be less than the BLIS area-threshold because purple segments can only appear upstream of blue segments. When inspecting blue objects only, the area-threshold for purple segment initiation is irrelevant as long as it is low enough to produce a high-resolution DN.

We apply a uniform BLIS area-threshold. All segments with accumulated upstream areas greater than the BLIS threshold were labelled blue, whilst segments with smaller accumulated upstream areas are labelled purple. We used the average watershed area of first order segments from DN2 as our BLIS threshold for illustrative purposes. Future applications could apply non-uniform BLIS area-thresholds if they any of the following methods could be applied.

- Create a classification of BLIS area-thresholds to apply different thresholds in different types of location such as lowland versus mountain areas.
- Create a continuously varying map of BLIS area-thresholds because the most appropriate thresholds may vary across locations such as would be the case if lower thresholds were more appropriate at lower elevations.
- Identify headwater location points independently from the DN (e.g., from a non-networked map or rivers). Match each headwater location to its nearest DN segment and then label all downstream segments to be blue. There is no need for a BLIS area-threshold in this case.

The algorithm that we applied for colouring the network was as follows.

1. Set the blue line initiation of stream (BLIS) threshold. For this exercise the mean of the watershed area of the Strahler one DN2 reaches in the clipped area containing catchment upstream of Te Waihora-Lake Ellesmere was used.
2. Segments with upstream area greater than BLIS were set to blue segments.
3. Non-blue segments have their Strahler order decremented by the BLIS Shift. The BLIS Shift is equal to the Strahler order of the first downstream blue segment below the non-blue segment if there is one, else the Strahler order of the terminal segment below the non-blue segment.
4. Blue segments have their Strahler orders recalculated starting from the blue headwaters (segments with no blue upstream segments) having a Strahler order of one.

4.5.2 Attributes representing DN resolution

Possible attributes of a network that might be useful to represent DN resolution include the following.

- Number of segments.
- Number of headwater segments.
- Number of terminal segments.
- Distribution of stream order before applying the negative Strahler stream ordering scheme.
- Distribution of stream order after applying the negative Strahler stream ordering scheme.
- Distribution of segment lengths.

- Distribution of watershed areas.

We calculated these attributes for each DN.

4.5.3 Metric representing DN alignment

A metric that produces a "distance" between two DNs could be defined using a combination of the above attributes, with the possible addition of the following more elaborate calculation to assess correspondence of segment alignment between two DNs. For each segment in the first network, consider N points equally spaced along the segment, then calculate the minimum distance from each point to a segment in the second network. So, for each segment a set of N distances is calculated, the distribution of these distances could be considered as a measure of distance between the two networks. Alternatively, one could take the mean of these values to produce a distance for each segment, and the mean of these means could produce a single number for the distance between the networks.

We did not calculate correspondence of segment alignment between two DNs because there are several potential problems with this approach. Some issues are as follows.

- Our metric is not symmetric, in other words the distance from network A to B is not necessarily the distance from network B to A. As an example, consider the two segments shown in Figure 4-4-a. Most of the distances from points on the red line to the blue line are small, however if one measures from the blue line to the red line there are a number of larger distances when considering points near the right-hand end.
- The above point illustrates a second problem, the metric does not necessarily agree with visual inspection by eye. A typical user of these river networks could possibly consider these two segments very close when looking by eye, the slight extension upstream might not be relevant. However, our measure from blue to red has many large distances.
- It is possible to measure the distances to the 'wrong' segment. Consider measuring the distance between B and R in the situation shown in Figure 4-4-b. For many points along B the shortest distance will be to segment M, but this is unlikely to be what is intended. A possible solution to this problem is to only consider distances to segments with identical Strahler order.
- If one restricts measurements between segments of identical Strahler order, problems can arise due to variability in stream initiation. Consider the situation shown in Figure 4-4-c. By eye these networks are similar, however if measurements are only made between segments with identical Strahler order the distance will be large.

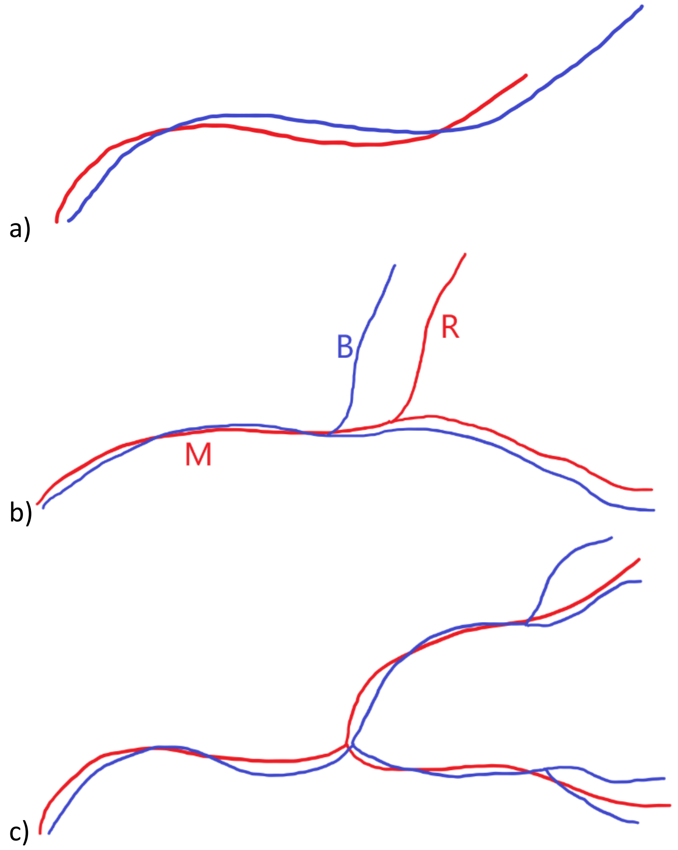


Figure 4-4: Hypothetical comparisons between networks.

Due to these issues, and in general the difficult nature of defining a metric that measures distance between networks, the six attributes previously presented were used to summarise each network.

5 Results

In this section, we show example results from two case studies to demonstrate potential challenges and solutions for DN generation. Section 5.1 shows example results for the entire domain analysed when the DN for the catchment upstream of Te Waihora-Lake Ellesmere was generated.

5.1 Quantified catchment characteristics

As the upstream area-threshold for segment initiation is increased the total number of segments (including headwaters and terminals) in the GRASS networks decreases (Table 5-1). The TauDEM network was initiated with an area-threshold of 1000 cells, so we would be expected it to be similar to the GRASS (1000) network, however, the TauDEM network had fewer segments. One explanation is that the TauDEM software and GRASS process apply different algorithms. DN3 is known to have many more segments than DN2 resulting from the different area-thresholds applied to generate those previously generated networks. The number of headwaters was approximately half the total number of segments across all networks.

Table 5-1: Scalar (single number) attributes of six DNs applied to a rectangular domain around the catchment upstream of Te Waihora-Lake Ellesmere.

Network	DEM resolution (m)	Stream initiation (Number of cells)	Stream initiation (m^2)	Number of segments	Number of headwaters	Number of terminals
GRASS (100)	8	100	6,400	537,481	270,583	1,614
GRASS (1000)	8	1000	64,000	91,364	46,037	569
GRASS (10000)	8	10000	640,000	14,248	7,231	201
TauDEM	8	1000	64,000	75,457	38,026	595
DN2	30	222	200,000	10,833	5,558	268
DN3	8	156	10,000	35,095	17,704	391

The distribution of, segment length (Figure 5-1) , watershed area (Figure 5-2), Strahler order (Figure 5-3), and Strahler order after colouring of the networks (Figure 5-4) for each network are shown below. A common y-axis has been used to make comparisons easier. Also note that the y-axis is not the raw variable of interest, but rather it is proportional to the total number of segments in that network. This is to make it easier to compare between networks, since the total number of segments varies markedly between networks. Colouring of the TauDEM network was not performed due to time constraints and differences in the routing conventions between GRASS and TauDEM. Results showed that all networks had many more short segments than longer segments (Figure 5-1). There were very few watersheds with large areas across all networks (Figure 5-2). TauDEM and GRASS networks, contained many watersheds with very small areas, likely because they were generated from a DEM containing more detailed information, which created small watersheds around the border of the analysed domain. DN2 and DN3 contained very few watersheds with very small areas, possibly because they were generated from a DEM containing less detailed information.

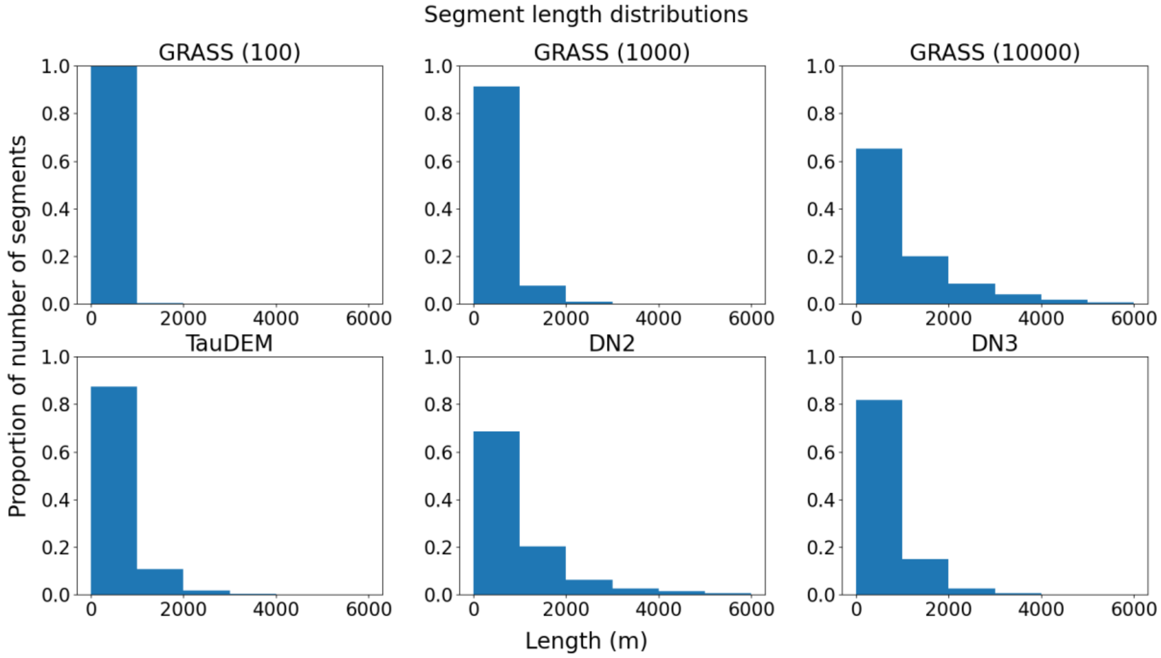


Figure 5-1: **Distribution of blue segment lengths for six DNs.**

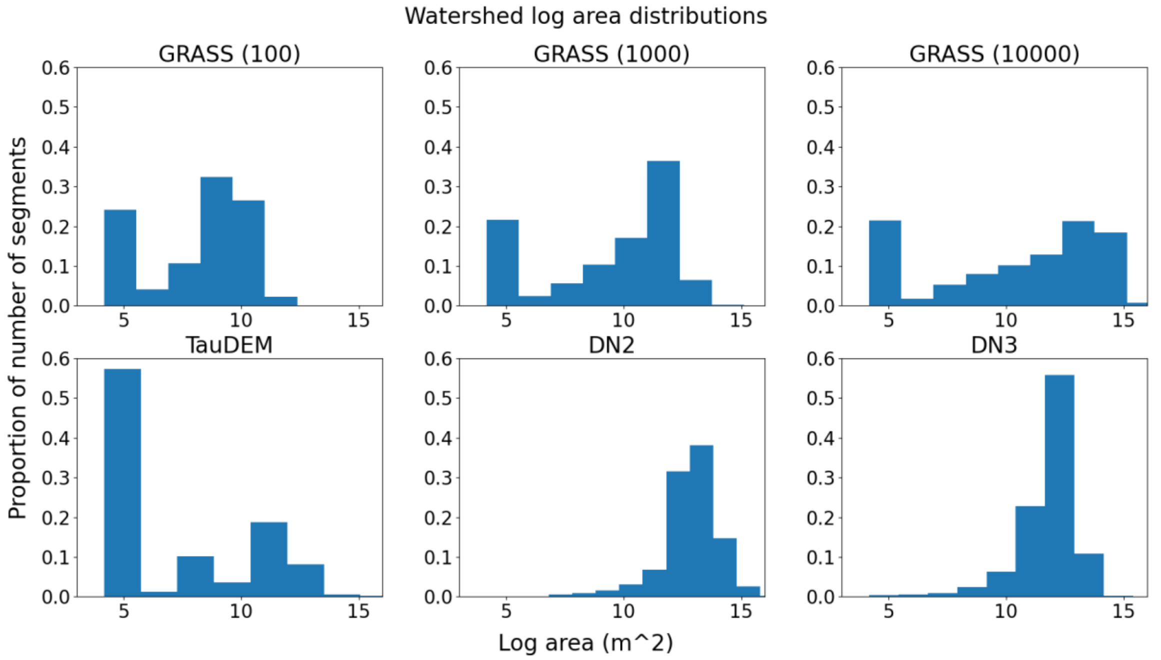


Figure 5-2: **Distribution of watershed areas for six DNs.**

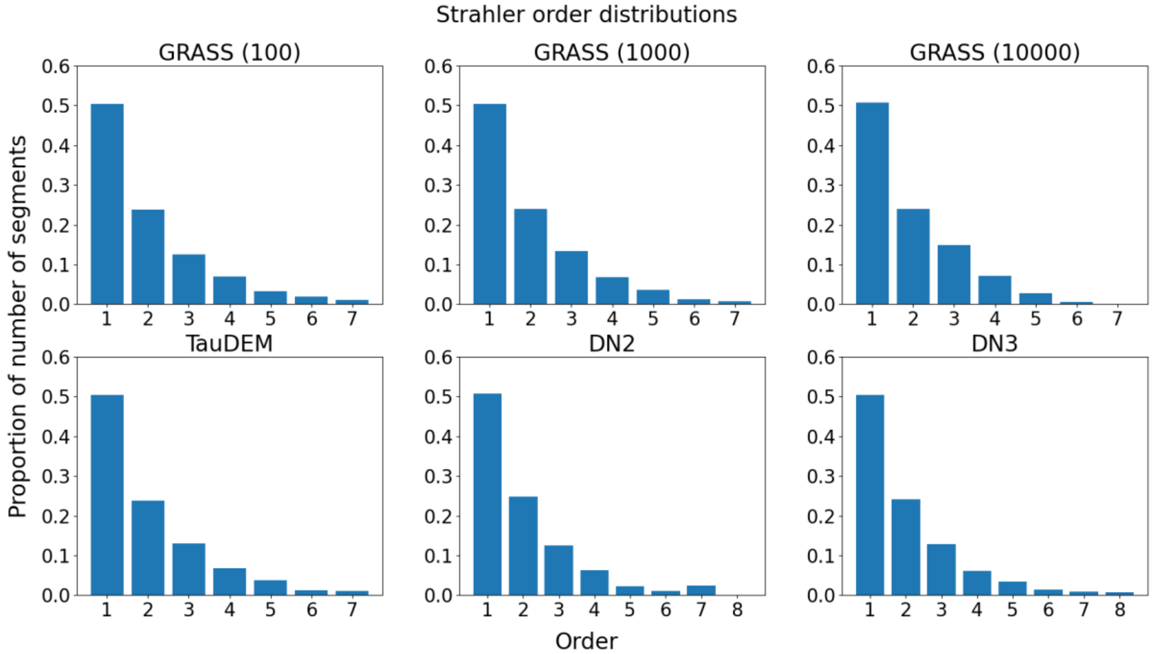


Figure 5-3: Distribution of Strahler stream orders for six DNs before the negative Strahler scheme was applied. Note differences in x-axis limits. Full domain used in DN generation.

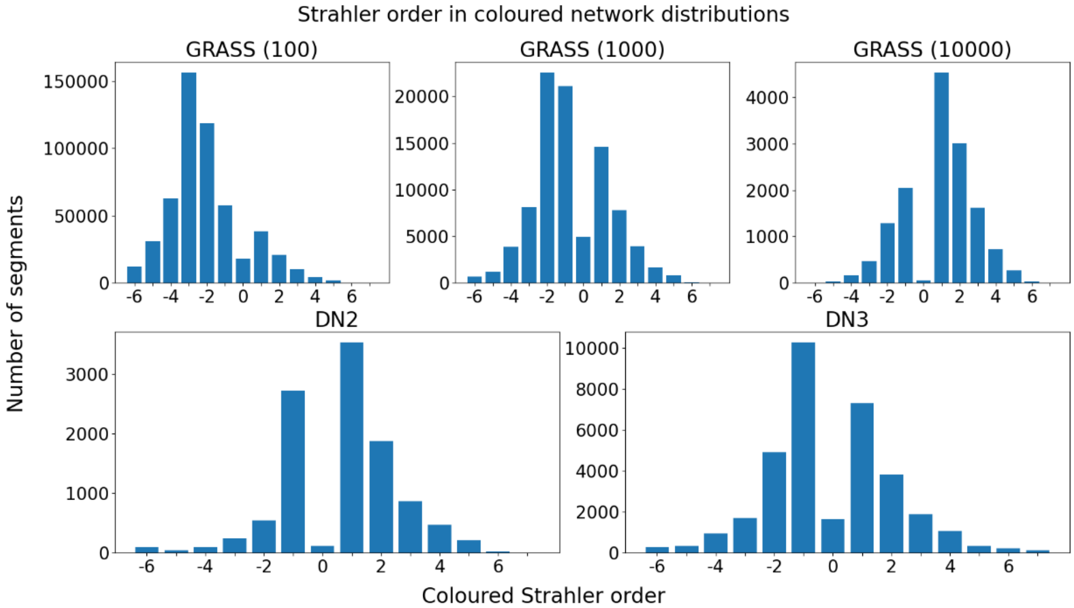


Figure 5-4: Distribution of Strahler stream orders for six DNs after the negative Strahler scheme was applied.

Strahler order distributions were not sensitive to the area-threshold for segment initiation for GRASS (100) and GRASS (1000) networks. The larger area-threshold applied for GRASS (10000) resulted in very few fourth or greater order segments. We note that the strange DN2 distribution (more seventh order than sixth order segment) was because of the truncation of the network to the analysed DEM extent.

We applied the negative Strahler stream order scheme after having applied the multi-coloured labelling system to obtain the negative Strahler orders for purple segments and new positive Strahler orders for blue segments (Figure 5-4). The scheme worked somewhat as anticipated to equalise the distribution of positive stream orders between DNs and illustrate differences in the distribution of negative stream orders between DNs. This result indicates that the multi-coloured labelling scheme was able to make comparable blue networks from DNs that originally comprised different resolutions of segments. The lack of zero Strahler orders segments occurs because they only occur when we have a first order segment and a higher order segment meeting, and the one downstream is the first blue reach. Zero order segments only occur in situations that equate to a minor tributary joining a main river stem.

5.2 Maps

Number of segments within the case study catchment boundaries (Table 5-2) were less than those within the analysed domains (Table 5-1). This demonstrated that an area much larger than the catchment of interest must be analysed to capture all the catchment of interest within a network. For example, the catchment upstream of Te Waihora-Lake Ellesmere contained only one quarter of the area analysed to generate that network. The GRASS (1000) network generated a total of 23,207 segments within the catchment upstream of Te Waihora-Lake Ellesmere (containing the Selwyn River) and a total of 28,774 segments inside the Te Awa o Mokotūāraro catchment (containing the Ngaruroro River) (Table 5-2). In this section we show maps of DNs generated for both case study catchments using the GRASS (1000) network and GeoFabrics DEMs. The maps illustrate network alignment (or misalignment) with aerial imagery and LINZ topographic maps.

Table 5-2: Segment details for the two case study catchments using GRASS (1000) and GeoFabrics DEMs.

Metric	Te Waihora-Lake Ellesmere (Selwyn)	Te Awa o Mokotūāraro (Ngaruroro)
Total number of segments	23,207	28,774
Number of negative order (purple) segments	15,506	18,936
Number of zero and positive order (blue) segments	7,701	9,838
Percent negative order (purple) segments	66.8	65.8
Percent zero and positive order (blue) segments	33.2	34.2
Catchment area (km^2)	2,600	3,369

5.2.1 Te Waihora-Lake Ellesmere

We plotted the DN2 against the GRASS (1000) network generated the GeoFabrics DEM data, with a focus on Te Waihora-Lake Ellesmere, to demonstrate issues associated with routing through lakes (Figure 5-5). The most noticeable difference between the two networks is routing around and through Te Waihora-Lake Ellesmere. Except for a few segments that route straight through the lake, the GRASS (1000) network appears to route segments around the perimeter of the lake before being routed through the lake. DN2, on the other hand, routes segments through the lake that converge on a lake centre-line and then to the lake outlet. Other differences include the number of segments included in each network, illustrated to the west and north of Figure 5-5). These differences – the different number of segments, the different alignment of segments, and the routing of segments around the lake in the GRASS (1000) network – are linked to underlying DEM data (Figure 4-1, Figure 4-2), as well as different values for stream initiation thresholds.

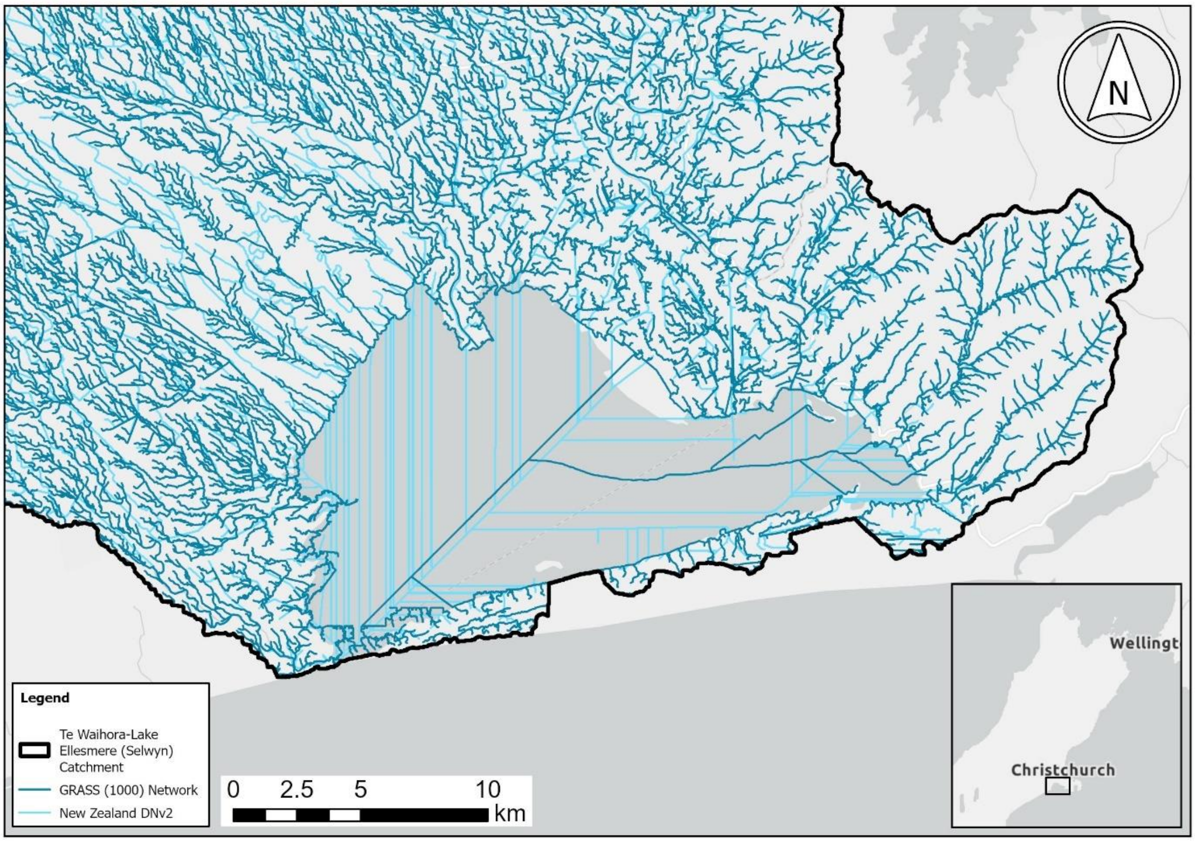


Figure 5-5: Comparison of network alignment and routing near Te Waihora-Lake Ellesmere. Networks shown here are the GRASS (1000) network generated for this report and the national DNv2. Map credits: Stats NZ, Esri, TomTom, Garmin, Foursquare, METI/NASA, USGS, Eagle Technology, LINZ, NIWA, Natural Earth, OpenStreetMap contributors.

The network upstream of Te Waihora-Lake Ellesmere shows, relative to the total number of segments, a proportion of 67% and 33% of purple and blue segments respectively. Visually, the catchment appears to be dominated by purple and low-order blue segments (Figure 5-6). The large proportion of purple segments demonstrates how network resolution and labelling might look for analysis such as calculating the length of river in a network.

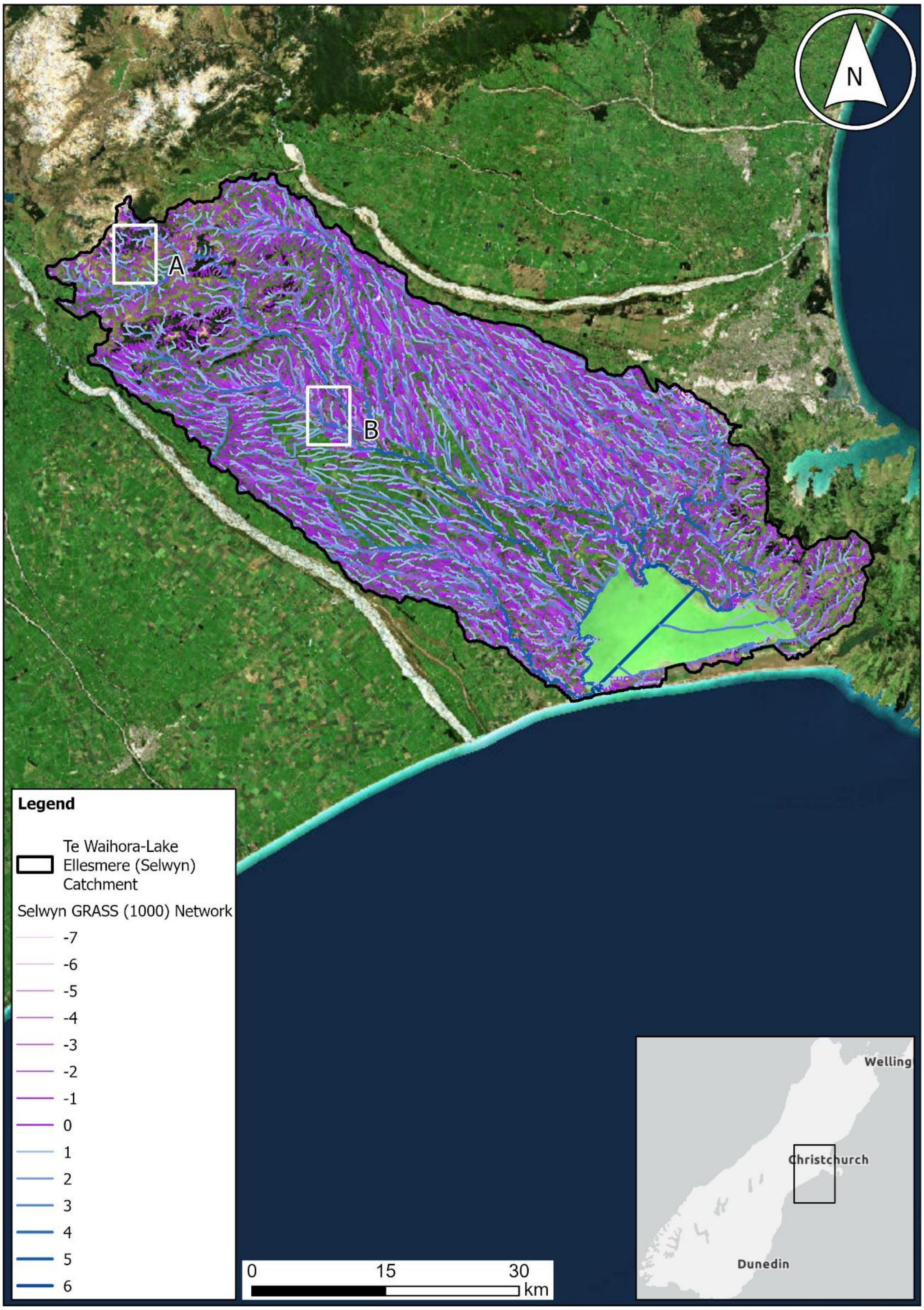


Figure 5-6: Te Waihora-Lake Ellesmere (Selwyn) Network. GRASS (1000) Network generated from GeoFabrics DEMs, coloured with the multi-coloured labelling system. Extent boxes A and B relate to the areas shown in Figure 5-7. Service layer credits: Eagle Technology, LINZ, StatsNZ, NIWA, Natural Earth, OpenStreetMap contributors, GEBCO, Community maps contributors.

Only a small proportion of the catchment upstream of Te Waihora-Lake Ellesmere in the northwestern extent comprises steep mountainous topography; the majority of the catchment is characterised by relatively flat, lowland plains topography. In general, segments were denser in lower-flatter areas than in higher-steeper areas, possibly supporting the need for a spatially varying area-threshold to be applied. The area of less dense segments toward the south-east of Box B in Figure 5-6 corresponded with the DEM area with less detailed topographic information shown in Figure 4-1. This indicates that network density is determined by a combination of input data and generation algorithm.

Magnified extents of the network in a mountainous headwater region and in a flat lowland region illustrate the alignment of the network with the local ground surface (Figure 5-7). In mountainous regions, the network appears to align closely with channelised topography visible in aerial imagery (Figure 5-7, A.1). Comparison of the network to the topographic map illustrates that the applied threshold for purple segments – which aim to represent ephemeral flow pathways (Table 2-3) – underestimates some channel initiation locations in this steeper terrain (Figure 5-7, A.2) and, in some cases, even the blue segments underestimate the locations of supposedly permanently-flowing channels. For example, the channel draining the northeastern side of spot height 949 appears in the topographic basemap as beginning at approximately 800 m elevation, and is represented with a solid blue line to indicate that the channel has permanently flowing water. The GRASS (1000) network suggests that channel has intermittently flowing water, and that it begins at approximately 700 m elevation (Figure 5-7, A.2). On the other hand, there are some instances where the GRASS (1000) network identifies channels that are not mapped as streamlines on the underlying topographic map. For example, again looking at spot height 949 (Figure 5-7, A.2), the channel to the west of the previously mentioned channel is identified by the GRASS (1000) network, but it is not mapped as a streamline on the underlying topographic map. Without field validation, it is unclear if the GRASS (1000) network has correctly identified an ephemeral channel in that location. The larger order segments, such as the mainstem of Bush Stream (Figure 5-7, A.1-A.2) appear to align well with the apparent valley floor. Although the GRASS (1000) segment does not always align with the apparent centre of the valley or with the apparent main channel of the river, it does stay within the confines of the valley walls (Figure 5-7, A.1-A.2).

In the flat lowland region (Figure 5-7, B.1-B.2), the GRASS (1000) network generally overestimates the presence of channels. Visible in both the aerial imagery (Figure 5-7, B.1) and on the topographic map (Figure 5-7, B.2), the GRASS (1000) network plots segments across agricultural land. Visual inspection of the segments indicates that none of the purple segments align with mapped streamlines on the topographic map, nor do some of the lower-order blue segments. For example, the blue segments that lie to the south of Dunsandel Road (Figure 5-7, B.1-B.2) appear in the GRASS (1000) network but not as streamlines in the topographic map. The network also failed to include the straight drain located immediately adjacent to the southern part of Dunsandel Road mapped on the topographic map (Figure 5-7, B.2). The higher order segment(s) representing the Hororata River align well with channel and appears to represent the sinuosity of the river. The GRASS (1000) segment(s) representing the Waikirikiri-Selwyn River (unlabelled, but visible as the gravel-bed river in the running diagonally in the top right corner of Figure 5-7, B.1) appear to generally follow the mapped channels in the topographic map, but because the Waikirikiri-Selwyn River is braided in this area, the network only follows one of the mapped channels at any given point. In some instances, smaller-order segments appear adjacent to the main channel and then route into the main channel, sometimes aligning with a mapped channel and then joining the higher-order segment. In general, the GRASS (1000) network in the flat lowland region aligns well with mapped channels for larger-order channels, but it overestimates the presence of smaller-order channels. The overestimation of smaller-order channels is likely due to subtle elevation differences across the landscape and captured in the DEM that may represent relict channels that existed in the landscape prior to agricultural development.

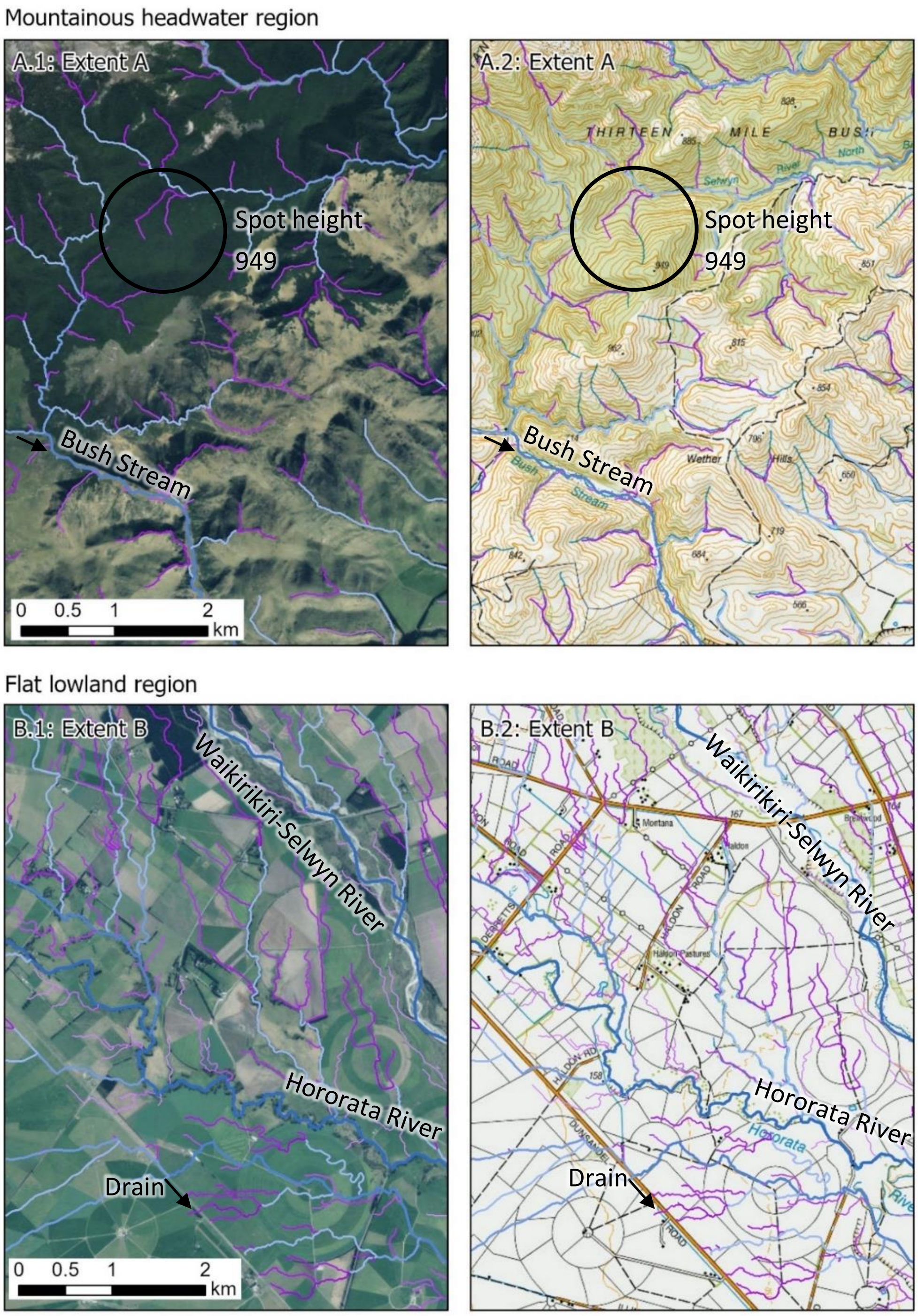


Figure 5-7: Zoomed extents of the Selwyn Network, overlaid on aerial imagery and a topographic map.

Figures A.1 and A.2 relate to Extent A in Figure 5-6 . Figures B.1 and B.2 relate to Extent B in Figure 5-6. Points of interest mentioned in the text are circled. Service layer credits: Eagle Technology, LINZ, GEBCO, Community maps contributors.

5.2.2 Te Awa o Mokotūāraro, Hawke's Bay

The GRASS (1000) network generated from GeoFabrics DEMs in the Te Awa o Mokotūāraro (Ngaruroro) catchment shows similar patterns to that in the Te Waihora-Lake Ellesmere (Selwyn) catchment, with the majority of the network comprising purple segments (61%) and relatively fewer blue segments (39%; Table 5-2). The Te Awa o Mokotūāraro (Ngaruroro) network has more higher-order (orders 4-6) blue segments (1307 segments) than the Te Waihora-Lake Ellesmere (Selwyn) network (788 segments), possibly due to a greater proportion of the catchment containing mountainous terrain with steeper, dissected topography that generate more channels that come together as opposed to flat lowlands, or more simply due to a larger catchment area (Table 5-2).

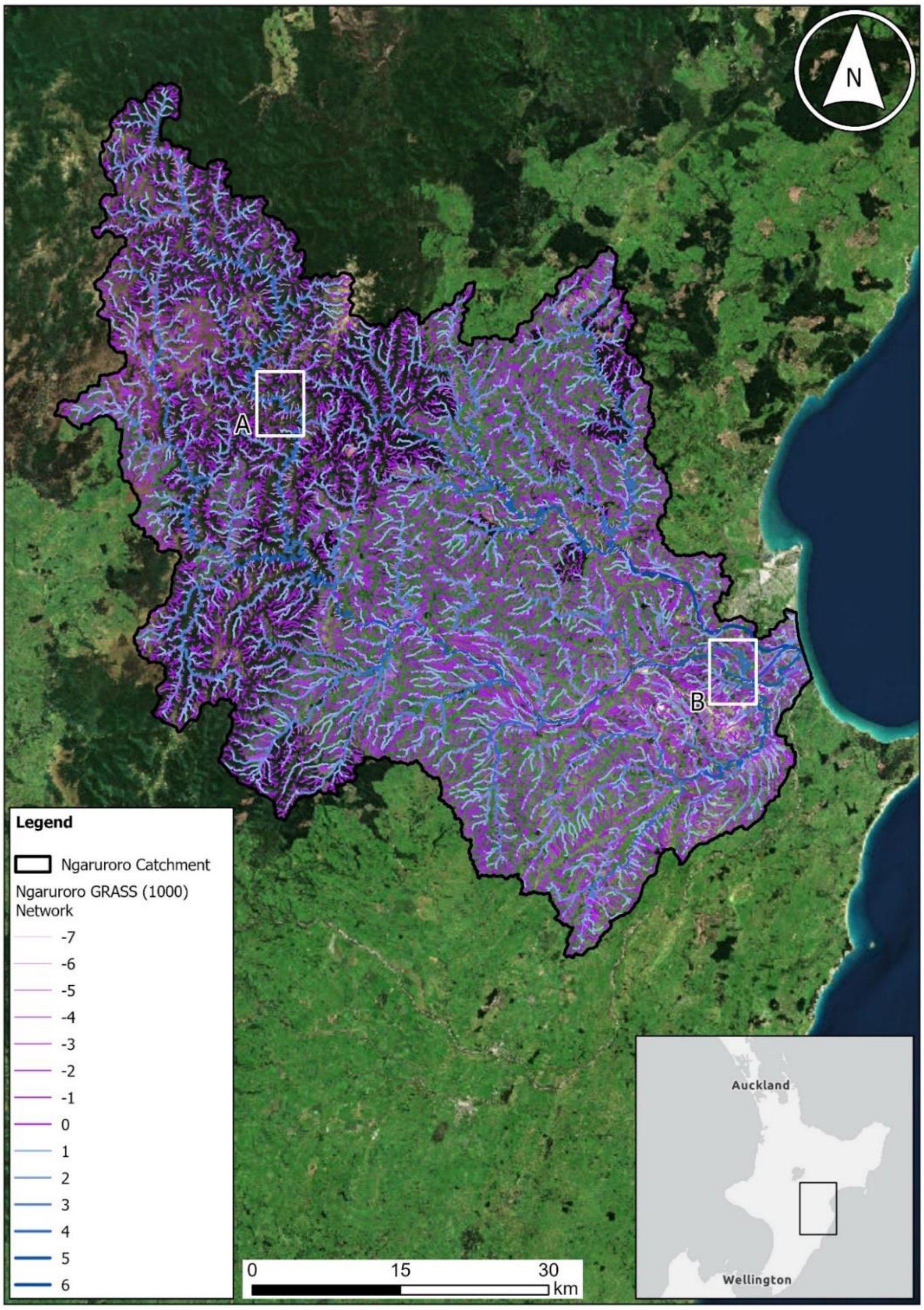


Figure 5-8: Ngaruroro Network. GRASS (1000) Network generated from GeoFabrics DEMs resolution. Extent boxes A and B relate to the areas shown in Figure 5-9. Service layer credits: Eagle Technology, LINZ, StatsNZ, NIWA, Natural Earth, OpenStreetMap contributors, GEBCO, Community maps contributors.

Magnified extents of the Te Awa o Mokotūāraro (Ngaruroro) network in a mountainous headwater region and in a flat lowland region illustrate the alignment of the network with the local ground surface (Figure 5-9). Visually, the network appears to align well with the underlying topography (Figure 5-9, A.1-A.2) but it underestimates the location of mapped channel initiation relative to the topographic map. Channelised areas upstream of mapped segments are visible in the aerial imagery, but it is unclear without field validation if there is permanent or intermittent water flowing in those channels. The mapped streamlines on the topographic map might suggest the presence of permanently flowing water, indicating that the network has underestimated the location of channel initiation with its current resolution (Figure 5-9, A.2). The higher-order segments that represent the Ngaruroro River in the centre of the map extents appear to stay within the confines of the valley floor, and, upon visual inspection, do appear to align well with the main channel of the river.

The selected flat lowland region in the Te Awa o Mokotūāraro (Ngaruroro) catchment is characterised by agricultural and urban areas (Figure 5-9, B.1-B.2). As with the Te Waihora-Lake Ellesmere (Selwyn) network, the Te Awa o Mokotūāraro (Ngaruroro) network appears to overestimate the presence of channels in the flat lowland regions (Figure 5-9, B.1-B.2). In some cases, the network shows segments routing across fields and along paved streets. The Te Awa o Mokotūāraro (Ngaruroro) network does appear to capture artificial drainage areas (realignments and possible some drains) slightly better than the Te Waihora-Lake Ellesmere (Selwyn) network, with segments routing down drains that run parallel to fields. The higher-order segments, such as those representing the Tutaekuri Waimate Stream, the Ngaruroro River (the gravel bed river running west to east in the centre of the map extents) and the Raupare Stream, align well with the channels visible in the aerial imagery (Figure 5-9, B.1) and the mapped stream lines in the topographic map (Figure 5-9, B.2). The segments representing the Ngaruroro River do not always fall directly on the mapped channel, but this is to be expected because the Ngaruroro River is a dynamic gravel-bed river that can change channel location through time. As with the Te Waihora-Lake Ellesmere (Selwyn) network, the overestimation of smaller-order channels is likely due to subtle elevation differences across the landscape that may represent relict channels that existed in the landscape prior to agricultural and/or urban development.

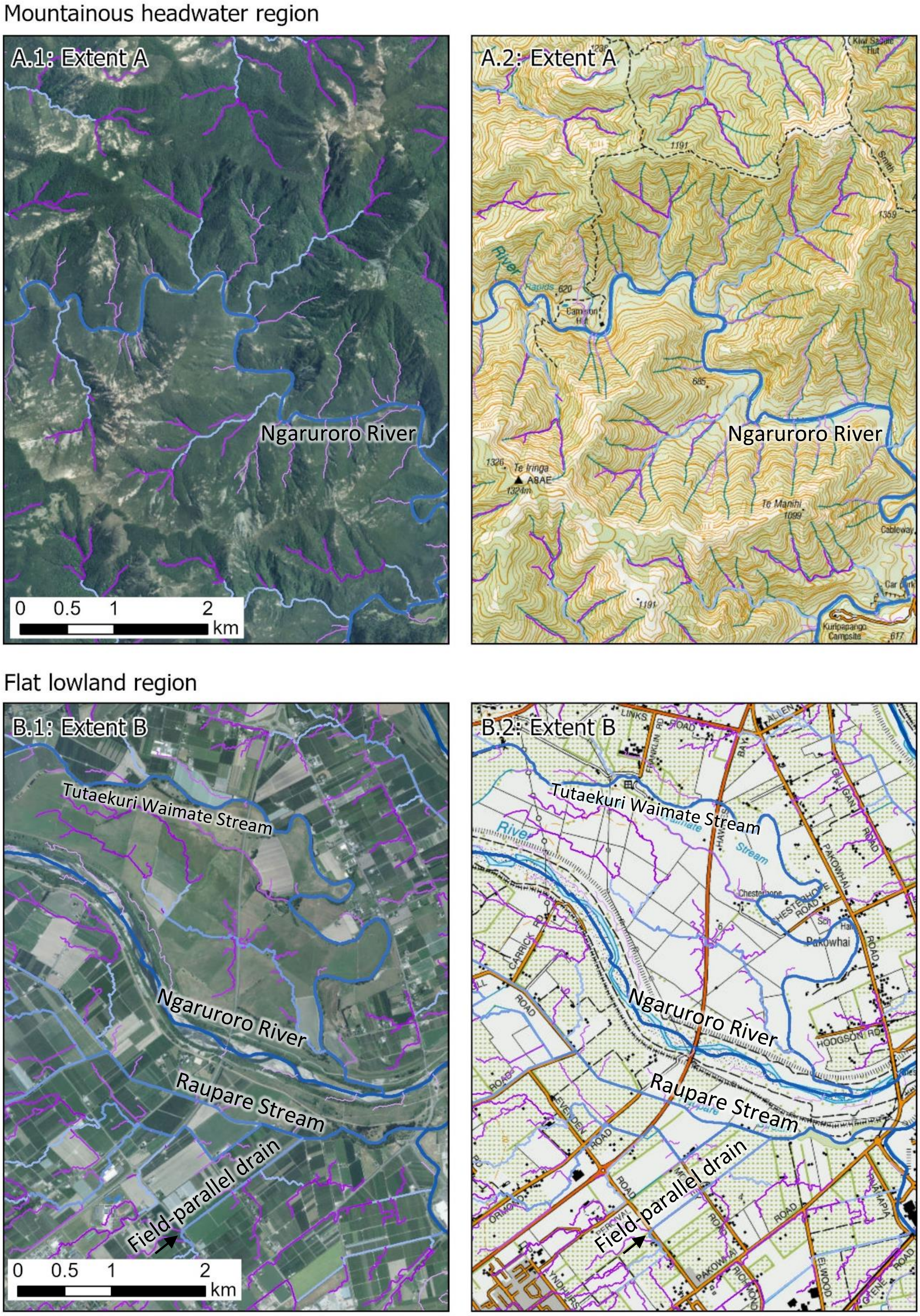


Figure 5-9: Zoomed extents of the Ngaruroro Network, overlaid on aerial imagery and a topographic map. Figures A.1 and A.2 relate to Extent A in Figure 5-8. Figures B.1 and B.2 relate to Extent B in Figure 5-8. Service layer credits: Eagle Technology, LINZ, GEBCO, Community maps contributors.

6 Discussion

6.1 Dataset coordination and missingness

Many datasets are available to feed into the generation and assessment of national DNs (see Table 3-2). Inspection of these datasets revealed inconsistencies between the same type of data represented within different datasets. For example, several datasets contained data on alignment of surface flow pathways but had different line densities and levels of resolution. Several datasets contain data on lake extent contained different numbers of lakes and levels of resolution (Schallenberg et al. 2024). The elements necessary to map wetland extent exist in two national databases (WONI and LCDB) and in sub-national datasets maintained by some regional councils (Newsome 2017). Differences between datasets may be advantageous for DN generation if a dataset is shown to be more suitable for DN generation than another. However, testing and comparing the utility of different datasets to assist with DN generation processes (particularly alignment of segments and position of segment headwaters) is technically challenging.

Different datasets had different coverage stemming from the positioning of their coastlines. Differences in data coverage between datasets presents a technical challenge for DN generation due to the increased propensity for missing data. Missing data present a technical challenge because it creates a need to interpolate (for missing data inland), extrapolate (for missing data around the coast), or track the effects of missingness within DN generation. Ideally, all datasets used in DN generation or assessment would extend up to or beyond a common coastline, but coordination of coastlines in national datasets is lacking because the data were produced by different institutions at different times using different raw inputs.

6.2 DEM resolution

High resolution DEMs can be advantageous for DN generation because they can resolve individual channels such as small streams. High resolution DEMs therefore have less need for river burning compared to coarser resolution DEMs. However, high resolution DEMs require large computational resources to process and store. High resolution DEMs also have the disadvantage that they include false or inconsequential topographic artifacts that need to be removed or ignored. Fine-scale flow directions can be difficult to detect and represent in DNs. Theoretically, there is a sweet spot of DEM resolution that is fine enough to resolve dominate surface flow pathways but not too fine to result in false or inconsequential topographic artifacts. Optimal grid resolution is likely to vary across the landscape due to differences between landscape phenomena as described in Section 3.1.

Smoothing of segment lines and watershed polygons is sometimes applied in automated DN generation algorithms when vectorising DEM grids into segments and watersheds. Smoothing would improve aesthetics of the DN and improve segment alignment when coarse grids are used. However, the finer the resolution of the DEM, the lesser the need for smoothing of segment lines and watershed polygons. The need for smoothing would be negated if a sufficiently fine grid was applied but segment lines would have a zig-zag appearance and watersheds would have a blocky appearance when highly magnified. If we consider grids of increasingly fine resolutions, at some point the need to vectorise the raster grid to create segments and watersheds is negated by the fineness of the grid. If a very fine grid is used and computational requirements are overcome, then all segment and watershed information could be represented by just labelling each cell with a flow direction, a label indicating whether the cell is a surface flow pathway, and a label indicating watershed membership.

6.3 Conventions for segment/watershed identifiers

Previous national DNs have applied conventions that allow users to quickly identify objects by island and region. Object IDs with seven digits were located in the North Island, with their first digit representing a region. Object IDs with eight digits were located in the South Island, with their first two digits representing a region. This convention could be applied in the future as long as a single region does not exceed a million segments. Catchments were not represented explicitly within object IDs of previous national DNs but could be subsequently identified using routing information. Catchment membership for each segment can easily be recorded as an object attribute.

6.4 Naming of rivers

Matching of DN objects to place names (river, catchment, region, aquifer, or administrative names) does not appear in the procedural steps outlined in Table 2-5. Attempts have been made to join catchment names to previous national DNs. For example, catchment names that flow to the sea have been joined to objects in national DNv2.4 (Whitehead and Booker 2019). Joining of DN objects to place names can add to DN utility by allowing easy searching or extracting of all segments comprising a catchment. However, matching of names to objects is challenging because:

- Not all segments will have an obvious name, leaving many objects to be labelled as "Unnamed" followed by an incrementing number. This is often the case for small first order segments of a DN that flow directly to the sea because they do not exist as defined channels in reality, or because they are genuinely unnamed.
- There are many different rivers with the same name (e.g., Stony Creek, Kaituna, etc.). However, this can be overcome if there is a catchment identifier also associated with the segment.
- Lengths of river with the same name do not always correspond to mathematical conventions for network routing.
- Some named rivers do not constitute catchments because they are named tributaries that do not flow to the sea.
- Spatial resolution and alignment of data sources with river names may not correspond with the DN being matched to. Thus, obtaining segment names by matching to the nearest named river line can produce erroneous results, especially around confluences.
- Names of rivers can change through time as official re-naming procedures are applied.

We suggest that matching a name to a set of DN objects is a feasible but labour-intensive task. For example, it is relatively easy to match a name to a contiguous sequence of DN segments if the name, starting node and finishing node are known. A data layer of polylines with river names is available via LINZ (LINZ river names data layer), these could be intersected with a DN and used as a significant head start on naming, but those lines do not have flow directions, do not necessarily connect to the sea, and would not necessarily align with segments of a functional DN.

6.5 Aquifers and underground routing

Representation of aquifers and underground routing is beyond the scope of this report. Methods have been devised to supplement surface water DN routing with groundwater routing in Aotearoa-New Zealand (Yang et al. 2017) and elsewhere (Hughes et al. 2015). Essentially these methods rely on a second underground routing layer comprising the same watersheds as the surface DN, but no segments. The underground routing layer can follow the surface routing or be amended to represent aquifer flow directions that are not parallel with surface flow pathways. The underground routing layer can be amended to represent multi-directional connections between watersheds. Underground connections are complicated by many-to-many matches because each watershed can border, and therefore route towards, an indeterminant number of neighbouring watersheds. One possible solution to simplify the underground routing conundrum is to create an underground layer comprising cells (with the same positioning as DEM cells) rather than watersheds. Use of cells is convenient because each cell can only have eight neighbours and matches to only one watershed.

6.6 Meeting recommendations for DNs from other work for MfE

Bright et al. (2022) aimed to support long-term improvements to freshwater accounting systems in Aotearoa-New Zealand. They gave 26 recommendations aimed toward: a) providing high quality information for a wide range of local, regional, and national needs; and b) increasing public trust and confidence in the environmental management system through increased transparency, completeness, and consistency of information. Bright et al. (2022) gave two recommendations specifically relating to DNs as follows.

- Recommendation 22: MfE should oversee the development of a nationally consistent digital water-body network “model” that uniquely labels each river reach and groundwater body, and their links to adjacent river reaches, adjacent groundwater bodies, and land parcels defined by the digital cadastral database[5], at a level of detail appropriate for catchment-scale application.
- Recommendation 23: Stream reaches and groundwater bodies should be labelled in a way that makes aggregation of water takes as simple as possible, particularly when aggregating upgradient from the coastline.

The guiding principles (Section 2.2), methodological strategy (Section 2.3), and technical solutions (Section 3) described in this report match well with these recommendations. We suggest there are two noteworthy issues relating to these recommendations. Firstly, we did not consider digital cadastral boundaries in this report. Secondly, segment IDs could be made to signify catchment membership using a similar convention as has been used to signify regions for each object in previous national DNs (see Section 6.3). However, such a system would create unwieldy IDs. Moreover, this information is easily extracted from routing information and can be stored in the network’s attribute table alongside catchment area, slope, length, etc.

6.7 Fit with previously recommended framework

Proposed procedural tasks outlined in Section 2.5 align well withing the previously proposed framework for DN and river classification generation (Figure 2-11). Strong correspondence between the procedural steps and our proposed technical solutions with the previously proposed framework indicates that the framework could be applied to generate a DN with a given set of characteristics

[5] LINZ digital cadastral database web page

(see Section 2.1) and therefore meet user needs (see Table 2-2) by conforming to the proposed principles (see Section 2.2).

After having considered technical challenges and potential solutions for DN generation, we propose two amendments to the DN generation framework as shown in Figure 6-1. The first amendment recognises that the multi-coloured labelling system negates the need to explicitly prune the network; a process of removing unwanted upstream segments to coarsen the network. The second amendment broadens "remotely sensed data" to include other dataset with national coverage. The previous framework implied use of "remotely sensed images" to generate river lines, lake polygons, and the coastline. Remotely sensed images can be used to detect water on the land surface if appropriate algorithms are applied, trained, and tested (Wang et al. 2020). Remotely sensed images can also be used to assess conditions in fluvial corridors (Piégay et al. 2020). However, challenges include interference from background noise, obscuration by clouds, temporal variability in waterbodies, and difficulty separating lakes from rivers or wetlands (Jiang et al. 2014; Huang et al. 2018). Application to relatively small and shallow rivers is a particular concern in the NZ setting. There are several datasets available with national coverage that could be used to obtain river, lake, and coastline information (Table 3-2).

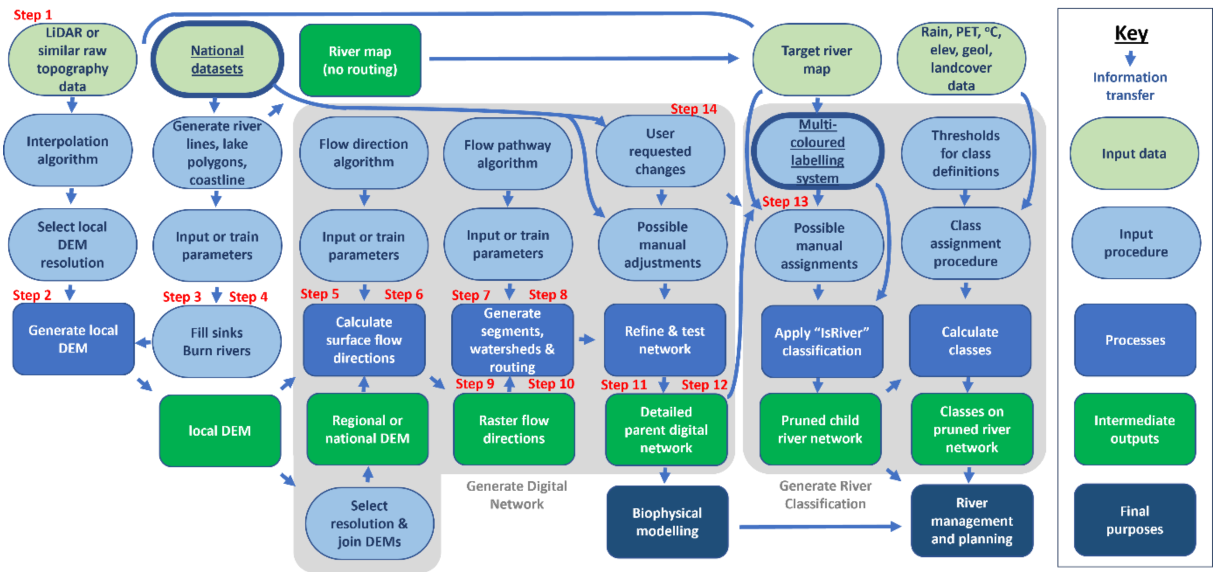


Figure 6-1: Procedural steps and amended framework for generation of DN. Original framework was from Booker (2023). Amended parts are underlined with bold borders. Technical steps shown in red text.

7 Conclusion and recommendations

We demonstrated that different DNs will result from a complicated interaction between generation procedures, algorithms, user input parameters, and DEM data, which is itself influenced by resolution and raw input data. This situation means that the presence of fine-scale DEM data, such as that generated from LiDAR surveys, is extremely useful for DN generation, but requires careful processing due to uncertainties in the data, lack of uniformity of data, and the variety of phenomena present across the landscape.

We make the following conclusions and recommendations in relation the ten principles outlined in Section 2.2 to guide DN generation in view of the breadth of user needs, likely ongoing improvements to input data, and the need for spatial consistency.

Reproduceable: We outlined a sequence of 14 procedural steps required for DN generation regardless of technical details about data or algorithms. The procedural steps align well with the proposed framework for DN generation and maintenance previously described by Booker (2023), which we amended slightly in light of technical developments and application to two pilot locations. We recommend that a national DN should be accompanied by a description of how each of these steps was applied or ignored (e.g., our procedure ignored the step for burning river lines because waterways were already incorporated into the GeoFabrics DEMs we used for DN generation).

Automatically updateable: We demonstrated that various automated routines are available for DN generation. We recommend that DN generation should be as automated as possible so that DNs can be efficiently updated following improvements to input data. However, we suggest that different DN generation algorithms and parameters should be based on landscape phenomena.

Version controlled: We used scripted code and existing algorithms to apply several DN generation procedures. Existing algorithms benefit from testing and development by a broad user community. We recommend that DN generation code is version controlled so that it can be reapplied and amended as necessary, and DN products are version controlled so that users know what data and methods were used to produce the DN they are using, and whether DN-derived products would be expected to match with each other.

Spatially consistent: We used GeoFabrics DEMs that had been previously generated using spatially consistent procedures to supplement available LiDAR data with other mapped data. We applied DN generation routines globally to the GeoFabrics DEMs to treat the entire DN domain in a spatially consistent manner. However, we also suggest that landscape phenomena such as rivers, wetlands, lakes, springs, and sinks present different challenges for DN generation. We suggest that the tension between the need for spatial consistency versus using bespoke procedures and data to best represent particular landscape phenomena in DNs can be overcome. We recommend that landscape phenomena can be delineated so that separate procedures can be consistently applied within their bounds (e.g., a lake-specific procedure can be applied within lake boundaries) to generate spatially consistent DNs. We note that spatially uniform sink-filling may compromise network routing through wetlands because it will remove wetland depressions from DEMs, but this method shows promise as a rapid and objective method of identifying potential wetland areas.

Bespoke updateable: We suggest that approaches for testing DN alignment include subjective visual comparisons against remotely sensed images or field observations. A DN user community can provide more testing and suggested improvements than can be applied at the time of DN generation. We recommend that a method to collate, vet, and approve user requests for DN alterations be developed and operationalised. Vetting of every requested DN alteration may be laborious and expensive. We suggest that some automated filtering could be applied prior to vetting to indicated if the area to be amended is small and whether the user needs to make just a few manual adjustments.

Spatially complete: We used a DEM with national coverage to generate DNs for two case study catchments. We also outlined several datasets with national coverage that could be used to improve DN generation and assessment. However, we noted inconsistencies between the coverage of these datasets, especially around the coastline. We suggest two possible solutions: a) all national datasets attempt to fill the same coverage by applying a common coastline; or b) DN procedures must include the ability to fill in all missing data. Application of a common coastline to all relevant national datasets may be unfeasible. We therefore recommend that DN procedures include filling in of missing input data. Areas that have been filled would ideally be identifiable by users.

Functionally correct: We outlined checks to verify that the network is behaving correctly from a mathematical perspective. We recommend that a network should not be released until it has passed a set of checks that confirm mathematically correct routing behaviour. Functional correctness would ideally be automated with use of an open-source toolbox.

Functionally informative: We devised a muti-coloured labelling system for generating a singular network that can viewed and used consistently across purposes with varying needs for DN characteristics (e.g., different resolutions). Blue objects within a network would represent a river network as would be used for river management purposes. The addition of purple objects representing ephemeral flow pathways would be used for biophysical modelling purposes. Green objects can be added to represent engineered flow pathways. We recommend that muti-coloured labelling system be applied to generate a single network that could be sub-sampled to be viewed and utilised for different purposes such as detailed biophysical modelling versus broader scale river management purposes.

Appropriate: We outline options for improved DN functionality by representing bifurcating channels so that artificial channels, islands, channels that split, and braided rivers could be represented if appropriate input information were available. We recommend that DN capabilities and the envisioned purposes are clearly explained and released alongside the DN.

Available, including version history: We stated that there are many legitimate uses for national DNs. We suggest that independent generation of DNs for different regions and by different institutions would cause inefficiencies and result in inconsistencies. We recommend that all users can obtain DNs, including previous versions, from a stable source. We note that small updates can be stored and communicated as deviations from parent network version, thus negating the need to store multiple network versions that exhibit only minor differences.

8 Acknowledgements

Thanks to Duncan Gray (ECan), Jeanine Topelen (ECan), Maurice Wills (ECan), Ashton Eaves (HBRC), and Phil McCartney (ES) for their discussions on DN generation and application. Thanks to Rose Pearson (NIWA) for providing DEM data. Thanks to Cyprien Bosserelle (NIWA), Christian Zammit (NIWA), and Ude Shankar (NIWA) for useful discussions about data availability and previous DN generation. Thanks to Olivia Burge (MWLR) for input on wetland mapping, Susie Wood (Cawthron Institute) for input on lake mapping, and Louise Weaver (ESR) for input on aquifer mapping.

9 References

Acreman, M., Holden, J. (2013) How Wetlands Affect Floods. Wetlands, 33. pp.773-786. DOI 10.1007/s13157-013-0473-2.

Ajvazi, B., Czimber, K. (2019) A comparative analysis of different DEM interpolation methods in GIS: case study of Rahovec, Kosovo. Geodesy and cartography, 45(1), pp.43-48.

Ashmore, P. (2013) Morphology and dynamics of braided rivers. Pages 290–312 *in* Shroder, J, ed., Treatise on geomorphology, v. 9, Wohl E, ed., Fluvial geomorphology. Academic Press.

Ausseil, A.G.E., Lindsay Chadderton, W., Gerbeaux, P., Theo Stephens, R.T., Leathwick, J.R. (2011) Applying systematic conservation planning principles to palustrine and inland saline wetlands of New Zealand. Freshwater Biology, 56(1), pp.142-161.

Ayad, M. (2021) An Assessment the Impacts of Climate Change on Mosul Dam Lake Using Satellite Data. MSc Thesis. http://dx.doi.org/10.13140/RG.2.2.28581.17121.

Band, L.E. (1986) Topographic partition of watersheds with digital elevation models. Water Resour. Res., 22, pp.15-24.

Baron, J.S., Poff, N.L., Angermeier, P.L., Dahm, C.N., Gleick, P.H., Hairston Jr, N.G., Jackson, R.B., Johnston, C.A., Richter, B.D., Steinman, A.D. (2002) Meeting ecological and societal needs for freshwater. Ecological Applications, 12(5), pp.1247-1260.

Barquín, J. (2004) Spatial patterns of invertebrate communities in spring and runoff-fed streams. Unpublished PhD thesis. Massey University, Palmerston North. 181p.

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. (2018) Quantifying the hydrological effect of permitted water abstractions across spatial scales. Environmental management, 62, pp.334-351.

Booker, D.J., Snelder, T.H. (2023) Climate change and local anthropogenic activities have altered river flow regimes across Canterbury, New Zealand. Water Resources Management, 37(6), pp.2657-2674.

Bright, J., Daughney, C., Jackson, B., McDowell, R., Smith, R., van Uitregt, B. (2022) Future Focused Freshwater Accounting. Prepared for Ministry for the Environment, ARL Report RD21011/1. Aqualinc Research Limited.

Brown, R.A., Pasternack, G.B. (2019) How to build a digital river. Earth-Science Reviews, 194, pp.283-305.

Carling, P., Jansen, J., Meshkova, L. (2014) Multichannel rivers: their definition and classification. Earth Surface Processes and Landforms, 39(1), pp.26-37.

Caruso, C., Quarta, F. (1998) Interpolation methods comparison. Computers & Mathematics with Applications, 35(12), pp.109–126. https://doi.org/10. 1016/S0898-1221(98)00101-1.

Cattoën, C., McMillan, H., Moore, S. (2016) Coupling a high-resolution weather model with a hydrological model for flood forecasting in New Zealand. Journal of Hydrology (NZ), 55(1), pp.1-23.

Coggin, D.W. (2008) Lidar in Coastal Storm Surge Modeling: Modeling Linear Raised Features. MSC Thesis. University of Central Florida.

Colson, T.P. (2006) Stream Network Delineation from High-Resolution Digital Elevation Models. PhD Thesis. North Carolina State University.

Connor-Streich, G., Henshaw, A.J., Brasington, J., Bertoldi, W., Harvey, G.L. (2018) Let's get connected: A new graph theory-based approach and toolbox for understanding braided river morphodynamics. Wiley Interdisciplinary Reviews: Water, 5(5), p.e1296.

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Boehner, J. (2015) System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. Geosci. Model Dev., 8, pp.1991-2007, https://doi.org/10.5194/gmd-8-1991-2015.

Costabile, P., Costanzo, C. (2021) A 2D-SWEs framework for efficient catchment-scale simulations: Hydrodynamic scaling properties of river networks and implications for non-uniform grids generation. Journal of Hydrology, 599, p.126306.

Datta, S., Karmakar, S., Mezbahuddin, S., Hossain, M. M., Chaudhary, B. S., Hoque, M. E., et al. (2022) The limits of watershed delineation: Implications of different DEMs, DEM resolutions, and area threshold values. Hydrology Research, 53(8), 1047–1062. https://doi.org/10.2166/ nh.2022.126.

Dietrich, W.E., Wilson, C.J., Montgomery, D.R., McKean, J. (1993) Analysis of erosion thresholds, channel networks and landscape morphology using a digital terrain model. The Journal of Geology, 101(2). https://doi.org/10.1086/648220.

Djokic, D., Ye, Z. (2000) DEM preprocessing for efficient watershed delineation. In Hydrologic and Hydraulic Modeling Support with GIS (ed. D. Maidment & D. Djokic), pp. 65–84. ESRI Press, Redlands, CA.

Djokic, D., Ye, Z., Miller, A. (1997) Efficient Watershed Delineation Using ArcView and Spatial Analyst, Proc. 17th Annual Esri User Conference, San Diego, CA.

Dudgeon, D., Arthington, A.H., Gessner, M.O., Kawabata, Z.I., Knowler, D.J., Lévêque, C., Naiman, R.J., Prieur-Richard, A.H., Soto, D., Stiassny, M.L., Sullivan, C.A. (2006) Freshwater biodiversity: importance, threats, status and conservation challenges. Biological reviews, 81(2), pp.163-182.

Dudley, B.D., Burge, O., Plew, D., Zeldis, J. (2020) Effects of agricultural and urban land cover on New Zealand’s estuarine water quality. New Zealand Journal of Marine and Freshwater Research, 54(3), pp.372-392.

Dymond, J.R., Sabetizade, M., Newsome, P.F., Harmsworth, G.R., Ausseil, A.-G. (2021) Revised extent of wetlands in New Zealand. New Zealand Journal of Ecology, 45(2),1–8.

Ehsani, A.H., Quiel, F., Malekian, A. (2010) Effect of SRTM resolution on morphometric feature identification using neural network-self organizing map. Geoinformatica, 14, pp.405-424.

Elliott, A.H., Semadeni-Davies, A.F., Shankar, U., Zeldis, J.R., Wheeler, D.M., Plew, D.R., Rys, G.J., Harris, S.R. (2016) A national-scale GIS-based system for modelling impacts of land use on water quality. Environmental Modelling & Software, 86, pp.131-144.

Ford, D. C., Williams, P.W. (2007) Karst Hydrogeology and Geomorphology, Wiley, Chichester.

Garbrecht, J., Martz, L.W. (1994) Grid size dependency of parameters extracted from digital elevation models. Computers and Geosciences, 20(1). pp.85-87.

Giachetta, E., Willett, S. D. (2018) A global dataset of river network geometry. Sci. Data 5, 180127.

Gleick, P.H. (1998) The human right to water. Water policy, 1(5), pp.487-503.

Greenlee, D.D. (1987) Raster and Vector Processing for Scanned Linework. Photogrammetric Engineering and Remote Sensing, 53(10), pp.1383–1387.

Guo, Q., Li, W., Yu, H., Alvarez, O. (2010) Effects of Topographic Variability and Lidar Sampling Density on Several DEM Interpolation Methods. Photogrammetric Engineering & Remote Sensing. 76(6), pp.701-712.

Haddadchi, A., Booker, D.J., Measures, R.J. (2018) Predicting river bed substrate cover proportions across New Zealand. Catena, 163, pp.130-146.

Harding, J.S., Winterbourn, M.J. (1997) An ecoregion classification of the South Island, New Zealand. Journal of Environmental Management, 51(3), pp.275-287.

Heritage, G.L., Milan, D.J., Large, A.R.C., Fuller, I.C. (2009) Influence of survey strategy and interpolation model on DEM quality. Geomorphology. 112, pp.334-344. http://dx.doi.org/10.1016/j.geomorph.2009.06.024.

Huang, C., Chen, Y., Zhang, S., Wu, J. (2018) Detecting, extracting, and monitoring surface water from space using optical sensors: A review. Reviews of Geophysics, 56(2), pp.333-360.

Hughes, J.D., Langevin, C.D., White, J.T. (2015) MODFLOW-based coupled surface water routing and groundwater-flow simulation. Groundwater, 53(3), pp.452-463.

Jenkins, D. G., McCauley, L.A. (2006) GIS, SINKS, FILL, and Disappearing Wetlands: Unintended Consequences in Algorithm Development and Use. In Proceedings of the 2006 ACM Symposium of Applied Computing (Track on Computer Ethics and Human Values), Dijon, France, April 23–27, 2006. ACM Digital Library 1-59593-108-2/06/0004, 227–282.

Jenson S.K., Domingue, J.O. (1988) Extracting topographic structure from digital elevation data for geographic information system analysis. Photogrammetric Engineering & Remote Sensing, 54, pp.1593–1600.

Jiang, H., Feng, M., Zhu, Y., Lu, N., Huang, J., Xiao, T. (2014) An automated method for extracting rivers and lakes from Landsat imagery. Remote Sensing, 6(6), pp.5067-5089.

Kashani, A.G., Olsen, M.J., Parrish, C.E., Wilson, N. (2015) A review of LiDAR radiometric processing: From ad hoc intensity correction to rigorous radiometric calibration. Sensors, 15(11), pp.28099-28128.

Kennish, M.J. (2002) Environmental threats and environmental future of estuaries. Environmental conservation, 29(1), pp.78-107.

Kilroy, C., Snelder, T.H., Floerl, O., Vieglais, C.C. and Dey, K.L. (2008) A rapid technique for assessing the suitability of areas for invasive species applied to New Zealand's rivers. Diversity and Distributions, 14(2), pp.262-272.

Kleinhans, M., van Kreveld, M., Ophelders, T., Sonke, W., Speckmann, B., Verbeek, K. (2019) Computing representative networks for braided rivers. Journal of Computational Geometry, 10(1), pp.423-443.

Larned, S.T., Moores, J., Gadd, J., Baillie, B., Schallenberg, M. (2020) Evidence for the effects of land use on freshwater ecosystems in New Zealand. New Zealand Journal of Marine and Freshwater Research, 54(3), pp.551-591.

Leathwick, J., West, D., Moilanen, A., Chadderton, L. (2012) Development of a systematic, information-based approach to the identification of high value sites for river conservation in New Zealand. In River conservation and management (ed. P.J. Boon and P.J. Raven), Ch. 15, pp.183-191.

Lehner, B., Grill, G. (2013) Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. Hydrol. Process, 27, pp.2171-2186. DOI: 10.1002/hyp.9740.

Li, J., Wong, D. (2010) Effect of DEM sources on hydrologic applications. Computers, Environment and Urban Systems, 34(3), 251-261.

Li, J., Li, T., Zhang, L., Sivakumar, B., Fu, X., Huang, Y., Bai, R. (2020) A D8-compatible high-efficient channel head recognition method. Environmental Modelling and Software, 125, art. no. 104624.

Lindsay, J. B. (2016a) Efficient hybrid breaching-filling sink removal methods for flow path enforcement in digital elevation models. Hydrol. Process., 30, pp.846–857. doi: 10.1002/hyp.10648.

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

Lindsay, J., Creed, I.F. (2005) Removal of Artefact Depressions from Digital Elevation Models: Towards a Minimum Impact Approach. Hydrological Processes, 19, pp.3113–3126. doi:10.1002/hyp.5835.

LINZ (2023) NZ 8m Digital Elevation Model (2012). Available online from https://data.linz.govt.nz/layer/51768-nz-8m-digital-elevation-model-2012/.

Makaske, B. (2001) Anastomosing rivers: a review of their classification, origin and sedimentary products. Earth-Science Reviews, 53(3-4), pp.149-196.

Mark, D. M. (1983) Automated detection of drainage networks from digital elevation models, Auto-Carto, 6, pp.169-178.

Mark, D. M. (1988) Network models in geomorphology. In Modelling Geomorphological Systems (ed. M. G. Anderson), John Wiley, New York.

Ministry for the Environment (2022) Method for deriving the stock exclusion low slope land map for the Resource Management (Stock Exclusion) Regulations 2020. Report reference INFO 1125.

Mitsch, W.J., Gosselink, J.G. (2000) The value of wetlands: importance of scale and landscape setting. Ecological Economics, 35(200), pp.25-33.

Mizukami, N., Clark, M.P., Gharari, S., Kluzek, E., Pan, M., Lin, P., Beck, H.E., Yamazaki, D. (2021) A Vector-Based River Routing Model for Earth System Models: Parallelization and Global Applications. Journal of Advances in Modeling Earth Systems, 13, e2020MS002434. https://doi.org/10.1029/2020MS002434.

Montgomery, D.R., Dietrich, W.E. (1989) Source Areas, Drainage Density, and Channel Initiation. Water Resources Research, 25(8), pp.1907-1918.

Montgomery, D.R., Foufoula-Georgiou, E. (1993) Channel Network Source Representation Using Digital Elevation Models. Water Resources Research, 29(12), pp.3925-3934.

Moreau, M., Bekele, M (2015) Groundwater component of the water physical stock account. GNS Science Consultancy Report 2014/290.

Morel, M., Booker, D.J., Gob, F., Lamouroux, N. (2020) Intercontinental predictions of river hydraulic geometry from catchment physical characteristics. Journal of Hydrology, 582, p.124292.

Morris, D. G., R. G. Heerdegen (1988) Automatically derived catchment boundaries and channel networks and their hydrological applications. Geomorphology, 1, pp.131-141.

Moshiri, G.A., Ed. (1993) Constructed Wetlands for Water Quality Improvement. Lewis Publishers, Boca Raton, FL.

Newsome, P.F. (2017) Considering a future spatial framework for wetland mapping and monitoring in New Zealand. Landcare Research Contract Report LC3000 for Ministry for the Environment. 29pp.

O'Callaghan, J., Mark, D.M. (1984) The extraction of drainage networks from digital elevation data, Comput. Vision Graphics Image Process., 28, pp.323-344.

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. https://doi.org/10.1016/j.envsoft.2023.105842

Petras, V., Petrasova, A., McCarter, J.B., Mitasova, H., Meentemeyer, R.K. (2023) Point Density Variations in Airborne Lidar Point Clouds. Sensors, 23, 1593, pp.1-18. https://doi.org/10.3390/s23031593.

Piégay, H., Arnaud, F., Belletti, B., Bertrand, M., Bizzi, S., Carbonneau, P., Dufour, S., Liébault, F., Ruiz-Villanueva, V., Slater, L. (2020) Remotely sensed rivers in the Anthropocene: State of the art and prospects. Earth Surface Processes and Landforms, 45(1), pp.157-188.

Planchon, O., Darboux, F. (2002) A fast, simple and versatile algorithm to fill the depressions of digital elevation models. Catena, 46(2-3), pp.159-176.

Poff, N.L., Allan, J.D., Bain, M.B., Karr, J.R., Prestegaard, K.L., Richter, B.D., Sparks, R.E., Stromberg, J.C. (1997) The natural flow regime. BioScience, 47(11), pp.769–784.

Polidori, L., El Hage, M. (2020) Digital Elevation Model Quality Assessment Methods: A Critical Review. Remote Sens., 12(21), 3522, https://doi.org/10.3390/rs12213522.

Qin, C.Z., Wu, X.W., Jiang, J.C., Zhu, A.X. (2016) Case-Based Knowledge Formalization and Reasoning Method for Digital Terrain Analysis–Application to Extracting Drainage Networks. Hydrology and Earth System Sciences, 20, pp.3379–3392. doi:10.5194/hess-20-3379-2016.

Qin, C., Zhu, A. X., Pei, T., Li, B., Zhou, C., Yang, L. (2007) An adaptive approach to selecting a flow partition exponent for a multiple flow direction algorithm. International Journal of Geographical Information Science, 21(4), pp.443-458.

Richardson, D.C., Holgerson, M.A., Farragher, M.J., Hoffman, K.K., King, K.B., Alfonso, M.B., Andersen, M.R., Cheruveil, K.S., Coleman, K.A., Farruggia, M.J., Fernandez, R.L. (2022) A functional definition to distinguish ponds from lakes and wetlands. Scientific Reports, 12(1), p.10472.

Rieger, W. (1998) A phenomenon-based approach to upslope contributing area and depressions in DEMs. Hydrol. Process, 12, pp.857-872.

Sangireddy, H., Stark, C.P., Kladzyk, A., Passalacqua, P. (2016) GeoNet: An open source software for the automatic and objective extraction of channel heads, channel network, and channel morphology from high resolution topography data. Environmental Modelling & Software, 83, pp.58-73.

Scarsbrook, M.R., Barquín, J., Gray, D.P. (2007) New Zealand coldwater springs and their biodiversity (Vol. 278). Wellington: Department of Conservation. 73pp.

Schallenberg, L., Thomson-Laing, G., Steiner, K., Shchapov, K., Hampton, H., Brasell, K., Pearman, J., Wood, S. (2024) Review of 'lakes' in the Freshwater Ecosystem New Zealand database. Nelson: Cawthron Institute. Cawthron Report 4039. Prepared for Ministry for the Environment.

Shahzad, F., Gloaguen, R. (2011) TecDEM: A MATLAB based toolbox for tectonic geomorphology, Part 1: Drainage network preprocessing and stream profile analysis. Computers & Geosciences, 37(2), pp.250-260.

Shankar, U., Griffiths, J., Kitlasen, W., Lilburne, L. (2022) NZWaM - Hydro-Geofabric guide. NIWA client report 2022135CH. 30pp.

Shutes, R.B.E. (2001) Artificial wetlands and water quality improvement. Environment International, 26(5-6), pp.441-447. https://doi.org/10.1016/S0160-4120(01)00025-3.

Singh, S.K., Zeddies, M., Shankar, U., Griffiths, G.A. (2019) Potential groundwater recharge zones within New Zealand. Geoscience Frontiers, 10(3), pp.1065-1072.

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.

Snelder, T., Lilburne, L., Booker, D., Whitehead, A., Harris, S., Larned, S., Semadeni-Davies, A., Plew, D., McDowell, R. (2023) Land-use suitability is not an intrinsic property of a land parcel. Environmental Management, 71(5), pp.981-997.

Sofi, M.S., Bhat, S.U., Rashid, I., Kuniyal, J.C. (2020) The natural flow regime: A master variable for maintaining river ecosystem health. Ecohydrology, 13(8), p.e2247.

Soille, P. (2004) Optimal removal of spurious pits in grid digital elevation models. Water Resources Research, 40(12), 9pp. https://doi.org/10.1029/2004WR003060.

Stewart-Harawira, M.W. (2020) Troubled waters: Maori values and ethics for freshwater management and New Zealand's fresh water crisis. Wiley Interdisciplinary Reviews: Water, 7(5), p.e1464.

Stoffels, R.J., Booker, D.J., Franklin, P.A., Snelder, T.H., Clapcott, J.E., Fragaszy, S.R., Wagenhoff, A., Hickey, C.W. (2021) Estimation of policy-relevant reference conditions throughout national river networks. MethodsX, 8, p.101522.

Tagliapietra, D., Sigovini, M., Ghirardini, A.V. (2009) A review of terms and definitions to categorise estuaries, lagoons and associated environments. Marine and freshwater Research, 60(6), pp.497-509.

Tarboton, D. G. (1997) A new method for the determination of flow directions and upslope areas in grid digital elevation models. Water Resources Research, 33(2), pp.309-319.

Tarboton, D. G. (2003), Terrain Analysis Using Digital Elevation Models in Hydrology, 23rd ESRI International Users Conference, San Diego, California, July 7-11.

Tarboton, D.G., Ames, D.P. (2001) Advances in the Mapping of Flow Networks from Digital Elevation Data. Bridg. Gap. 1–10.

Tarboton, D.G., Bras, R.L., Rodriguez-Iturbe, I. (1991) On the Extraction of Channel Networks from Digital Elevation Data. Hydrol. Process., 5, pp.81-100.

Turcotte, R., Fortin, J. P., Rousseau, A. N., Massicotte, S., Villeneuve, J. P. (2001) Determination of the Drainage Structure of a Watershed Using a Digital Elevation Model and a Digital River and Lake Network. Journal of Hydrology, 240, pp.225–242. doi:10.1016/S0022-1694(00)00342-5.

Tarboton, D.G., Mohammed, I.N. (2013) TauDEM 5.1 quick start guide to using the TauDEM ArcGIS toolbox. Utah State University (USU), United States.

Wang, L, Liu, H. (2006) An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling. International Journal of Geographical Information Science, 20(2), pp.193-213.

Wang, G., Wu, M., Wei, X., Song, H. (2020) Water identification from high-resolution remote sensing images based on multidimensional densely connected convolutional neural networks. Remote sensing, 12(5), p.795.

Wang, Y.J., Qin, C.Z., Zhu, A.X. (2019) Review on algorithms of dealing with depressions in grid DEM. Annals of GIS, 25(2), pp.83-97.

White, P.A. (2001) Groundwater resources in New Zealand. In Rosen, MR & White, PA (Eds), Groundwaters of New Zealand (pp45–75). Wellington: New Zealand Hydrological Society.

White, P.A., Moreau, M., Mourot, F., Rawlinson, Z.J. (2019) New Zealand groundwater atlas: Hydrogeological-unit map of New Zealand. GNS Science consultancy report 2019/144.

White, W.B. (2005) Springs. In Encyclopaedia of caves (eds. D.C. Culver and W.B. White). Elsevier Academic Press, London. pp.565–569.

Whitehead, A.L., Booker, D.J. (2019) Communicating biophysical conditions across New Zealand's rivers using an interactive webtool. New Zealand journal of marine and freshwater research, 53(2), pp.278-287.

Whitehead, A.L., Leathwick, J.R., Booker, D.J., McIntosh, A.R. (2022) Quantifying the relative contributions of habitat modification and mammalian predators on landscape-scale declines of a threatened river specialist duck. PLOS One, 17(12), p.e0277820.

Wohl, E. (2018) The Upstream Extent of a River Network: A Review of Scientific Knowledge of Channel Heads. ERDC/CRREL CR-18-1. Report prepared for Wetlands Regulatory Assistance Program (WRAP), U.S. Army Corps of Engineers.

Wu, S., Li, J., Huang, G.H. (2007) Modeling the effects of elevation data resolution on the performance of topography-based watershed runoff simulation. Environmental Modelling & Software, 22(9), pp.1250-1260. https://doi.org/10.1016/j.envsoft.2006.08.001.

Wu, S., Li, J., Huang, G.H. (2008) A study on DEM-derived primary topographic attributes for hydrologic applications: Sensitivity to elevation data resolution. Applied Geography, 28(3). pp.210-223. https://doi.org/10.1016/j.apgeog.2008.02.006.

Yang, J., McMillan, H., Zammit, C. (2017) Modeling surface water–groundwater interaction in New Zealand: model development and application. Hydrological Processes, 31(4), pp.925-934.

Yang, J., Dudley, B.D., Montgomery, K., Hodgetts, W. (2020) Characterizing spatial and temporal variation in 18O and 2H content of New Zealand river water for better understanding of hydrologic processes. Hydrological Processes, 34(26), pp.5474-5488.

Zeiringer, B., Seliger, C., Greimel, F., Schmutz, S. (2018) River hydrology, flow alteration, and environmental flow. Riverine ecosystem management: Science for governing towards a sustainable future, pp.67-89.

Zhang, W., Montgomery, D. (1994) Digital elevation model grid size, landscape representation, and hydraulic simulations. Water Resources Research., 30(4), 1019-1028. 0043.1397/94/93WR-0355.