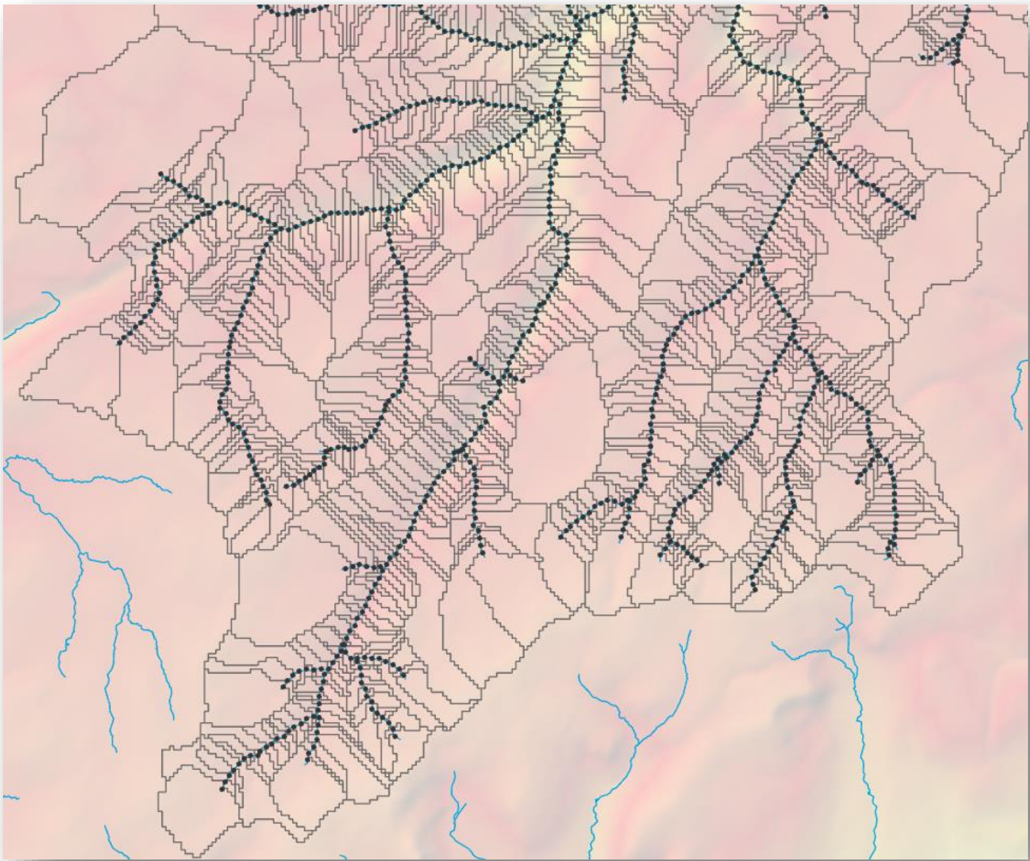
OpenRiverNetwork (ORN) Catchments

Supplemental: catchment polygons for the topologically aligned sampling points of the ORN



Report of Joint Project: GeoData® Institute, University of Southampton, River Restoration Centre

March 2025

OpenRiverNetwork (ORN) Catchments

Supplemental: catchment polygons for the topologically aligned sampling points of the ORN

Suggested citation: Hornby, D.D., Naura, M., and Hill, C.T. (2024) OpenRiverNetwork Catchments Supplemental: catchment polygons for the topologically aligned sampling points of the ORN

Online resource: http://openrivers.net/
GeoData Institute, River Restoration Centre (RRC)
Version 2.0
March 2025

Third Party Data Acknowledgements:
The development of OpenRiverNetwork Catchments is built upon the open dataset OS Open Rivers, Ordnance Survey 2023.

Disclaimer

This report has been compiled based on the processing and analysis of a modified version the Ordnance Survey's Open Terrain 50$^{(TM)}$ Digital Elevation Model. The product is provided as a research output; the University of Southampton, RRC and the OS do not accept any liability for any errors in the delineation of the catchment boundary

IMPORTANT: Not all sampling points have a catchment polygon as they failed to reliably generate a catchment area. In areas that are low in elevation, little topographic relief and drainage was subject to human alteration then catchment data may be sparse due to a consistent failure to delineate logical catchment areas.

RRC is a national centre for information and advice and holds a dataset of river restoration and best practice management works. To inform the further development of this network, please let us know of any issues you may find or if you any suggestions how we could make it an even more valuable resource (rrc@therrc.co.uk).

GeoData Institute is a research and enterprise group within the School of Geography and Environmental Sciences at the University of Southampton. It specialises in the use of geospatial information for a sustainable environment and society (geodata@soton.ac.uk)

Contents

Background 4

Introduction 4

Methods 6

Step 1: preparing the OS Terrain 50 DEM 6

Step 2: delineating catchments 7

Quality Check 1: testing if source is in catchment 7

Quality Check 2: using a distance ratio to identify candidates for rejection 10

Quality Check 3: counting sources in candidate rejections 10

Quality Check 4: resolving cross catchment errors 11

Step 3: final merge of data, check and export 12

Quality Check 5: a manual scan along the coastline 12

Step 4: preparing the final release 13

Other reasons for catchment delineation failure 14

Results 16

Conclusion 18

User guidance 19

Accessing catchment data in ESRI ArcPro 20

Accessing the catchment data using the supplied File GeoDatabase relationship class 20

Create your own map relationship and access the catchment data 22

Accessing catchment data in QGIS 24

Ensure fields are indexed 24

Creating the relationship between sampling point and catchment polygon 26

Using the relationship 27

An example of automating processing using python in ArcPro 29

References 31

OpenRiverNetwork (ORN) Catchments

Supplemental: catchment polygons for the topologically aligned sampling points of the ORN

Background

Hornby et al. (2023) took a 2023 release of the Ordnance Survey Open Rivers dataset and created a fully connected river network. The development of the OpenRiverNetwork was supported by Ordnance Survey. This allowed the network to be attributed with a range of metrics and was sampled at 100m intervals, generating a dataset of 1,472,056 points across the entire river network to manage multiple attributes. Key to this first release of the Open River Network (ORN) were the resolution of many topological errors that existed in the OS Open Rivers dataset.

Further details on the editing, processing, quality control and preliminary results can be found in the Hornby et al. (2023) report that be accessed from https://openrivers.net/.

Introduction

The ORN is an open dataset that can be used for catchment analysis. To add further value and analytical processing power, catchment areas have been delineated for many of the sampling points. These catchment polygons are an addition to the version 1 release of the ORN. Using open data ensures these catchment areas are free to access and use without restrictions.

Code was developed to call standard hydrological tools in ESRI ArcGIS Pro to delineate a catchment polygon for every sampling point (1,472,056) and write this into a single dataset. The result is a large dataset with overlapping polygons that have a 1-to-1 relationship to the sampling point ID.

It is well known that delineation of catchment boundaries can fail in areas of low topographic relief or where humans have significantly altered the natural drainage pattern of an area. As part of the code each generated catchment polygon was quality controlled and flagged as accept or reject.

To improve the success rate of the catchment delineation, the OS Terrain 50[†] underwent DEM conditioning using the ArcHydro tool to stream burn the ORN into the DEM, a technique often used to improve delineation results.

Due to the generalisation of the ORN, resolution of the DEM (50m pixel size) and proximity of neighbouring catchment tributaries, in some cases the stream burning failed and introduced an illogical cross-catchment scenario. A second round of quality control was developed and run on the initial output to further reject illogical catchments.

A final quality control was to manually scan, following the coastline, and identify further anomalous catchments which were tagged to be filtered out.

[†] = https://osdatahub.os.uk/downloads/open/Terrain50

With 5 quality control steps and the use of the stream burning technique to improve success it is still possible for unusual shape catchments to exist within this dataset. These typically exist within areas of modified drainage and are a likely product of the fact that the OS Terrain 50 dataset is an interpolation of a TIN and that the process of converting the TIN to a raster did not enforce drainage as it was not the intention of OS Terrain 50 to map the surface of the Great Britain as a hydrological surface.

The methods section of this report details the source of various problems and the quality control steps taken.

The final catchment dataset is a large feature class with 1,299,024 overlapping polygons! It is not the intention of this dataset to be displayed in full as it is a highly complex dataset that would be very difficult to interpret to non-specialists.

With each catchment polygon linked to its sampling point by ID it is very easy to filter the dataset either as a selection, definition query or export. This manual provides step-by-step advice in such techniques using ArcPro or QGIS under the **User guidance** section of this report.

The **true power of this dataset** can be exploited by advanced users, scripting or using model builder to automate a sequential step through sampling points to select up their respective catchment to be used in a spatial query. A sample script can be found at the end of this report under the User guidance section.

An example could be the change in land use and its impacts on diffuse pollution inputs into the river downstream. In fact, anything that intersects the catchments and make sense in and accumulating fashion could be extracted with minimal scripting.

IMPORTANT: Only catchment polygons that passed the quality controls exist in this release of the dataset which means not all sampling points have a catchment polygon as they could not be reliably generated.

Methods

The creation of this catchment dataset used a series of steps starting with the OS Terrain 50 dataset being prepared using the technique of stream burning. A Python script was developed to automate the delineation of the catchment boundaries, results were passed through a further 3 quality control scripts then a final manual scan. The bulk of the delineation processing was distributed over 2 computers and took approximately 1 month to complete.

Step 1: preparing the OS Terrain 50 DEM

It is well known that delineation of catchment boundaries can fail in areas of low topographic relief or where humans have significantly altered the natural drainage pattern of an area. Standard GIS tools using a DEM can delineate a catchment for a location along a river. This relies on the generation of a flow direction raster along with a flow accumulation raster to aid in the snapping of a point to a pour point location used in the delineation step. If the base DEM was not built with surface hydrology in mind, then the flow direction/accumulation rasters can deviate from the reality of where the water course is seen in a topographic map. To improve the success of delineation a technique known as stream burning was employed.

ESRI ArcHydro was downloaded and installed. Although this is a free to use tool it requires the spatial analyst extension and higher license levels to operate.

The ORN was converted to a raster using the ArcPro tool Polyline to Raster. Cell size was set 50m and snap raster environment set to the DEM. This creates a raster version of the river network aligned to the DEM.

The ArcHydro DEM reconditioning tool was used, with default settings, to burn the raster version of the ORN into the DEM (Figure 1). This tool uses the AGREE method to burn in the river course. Further details on the AGREE method can be found on a web page by Maidment (1997).

With the AGREE DEM created, it was run through the standard fill sinks, flow direction and flow accumulation tools, accepting default settings to create the rasters necessary for delineation.

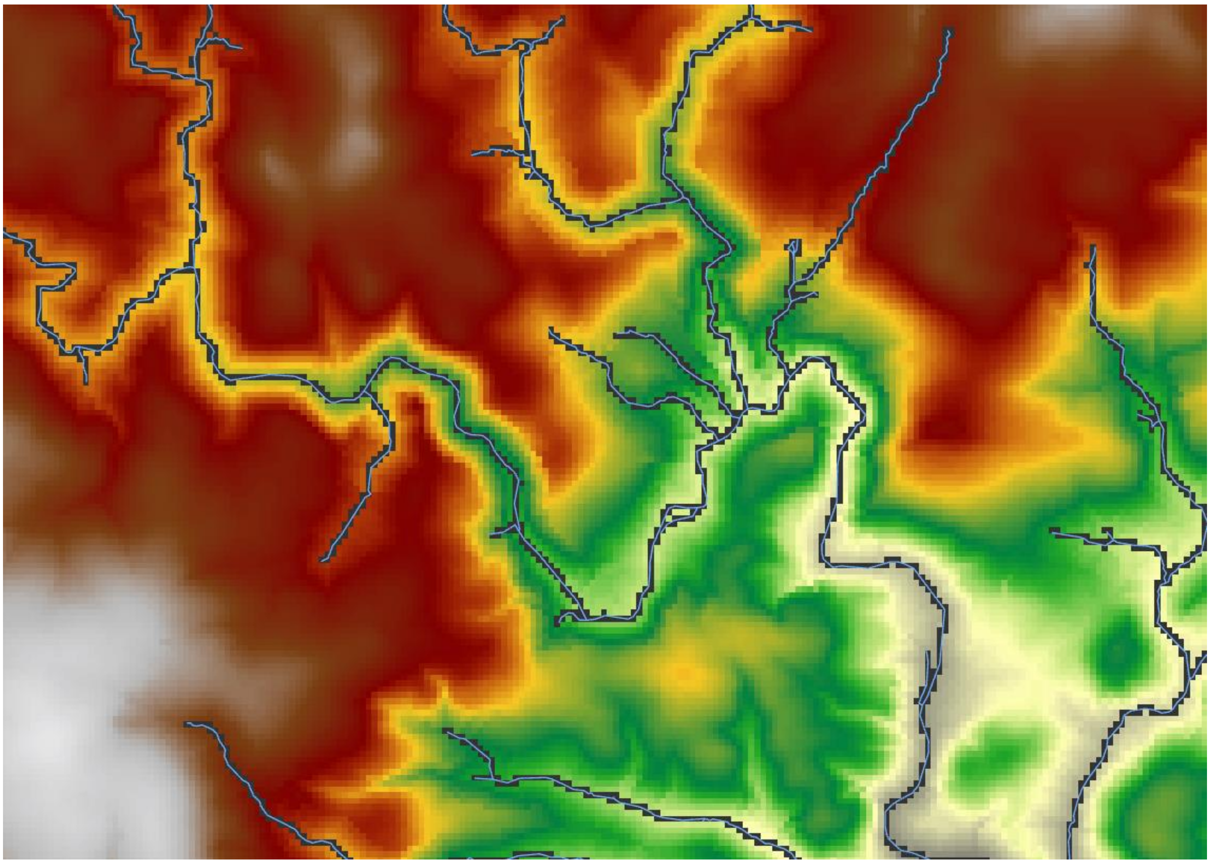


Figure 1 The result of stream burning; where the vector network exists, the pixels have been dropped in elevation by 1000m. The resulting flow direction/accumulation rasters should align better with the network.

Step 2: delineating catchments

A python script was developed to automate the production of catchment areas. To reduce processing time the script was run on two computers each processing a batch of sampling points. The script was run in IDLE, the python interpreter that comes packaged with ArcPro. As each batch of sampling points were processed, IDLE was restarted to release memory to ensure the next batch of points would run to completion. It took approximately 1 month to complete the initial catchment delineation for the 1,472,056 sampling points.

Catchment delineation, even with a stream burnt DEM, can generate erroneous boundaries due to lack of topography, imperfections within the DEM or spatial alignment issues caused by datasets captured at different scales and resolution. This means not all sampling points would generate a logical catchment boundary and as such should be excluded to avoid misuse. A set of quality control checks were devised to find candidate catchment areas for rejection.

Quality Check 1: testing if source is in catchment

An important step in the generation of the catchment boundary is to test its validity. The ORN is a richly attributed network and one of the attributes is the source ID. All sampling points are attributed with their source so a simple, yet powerful test, of *is the source in the catchment polygon just generated,* can be used to identify sampling points that failed to generate a logical catchment boundary. Such sampling points will have their erroneous catchment boundary flagged to be rejected.

The method of catchment delineation is to snap the sampling point to the nearest largest flow accumulation value and that becomes the starting pixel for delineation. This is discussed in detail in Figure 2.

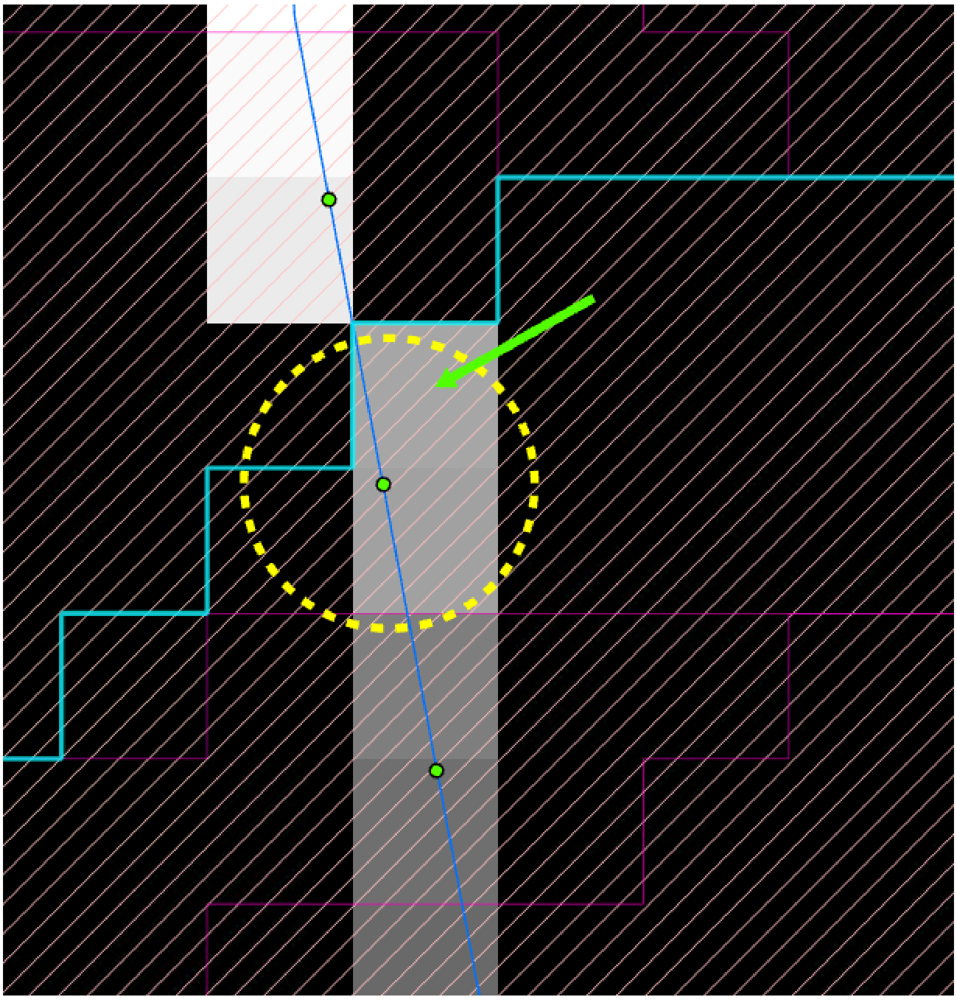


Figure 2 The green point is a sampling point on the ORN (blue line). The dashed circle indicates a 50m search distance. The white/grey/black squares are the flow accumulation raster. The lighter the pixel colour, the larger the flow accumulation value. The green arrow points to the grey cell that is used to start catchment delineation. This was the cell with the largest value within 50m of the green point. Although the green point happens to be on a sensible flow accumulation cell, the process of snapping the point has shifted it downstream. Finally, the hatching are the generated overlapping catchment polygons, the selected polygon (cyan) is the catchment polygon for the point, it was created from the starting cell indicated by green arrow.

Potential issues: For most sites, delineation works as expected, yet if a sampling point is on a parallel channel or very near a tributary junction then snap processes will identify a starting pixel on a larger channel as defined by the flow accumulation raster. This leads to the erroneous situation of a catchment boundary being generated for the much larger channel which also contains the source of the sampling site on a different branch of the network. An example of this false positive snap is shown in Figure 3 which leads to the erroneous catchment in Figure 4.

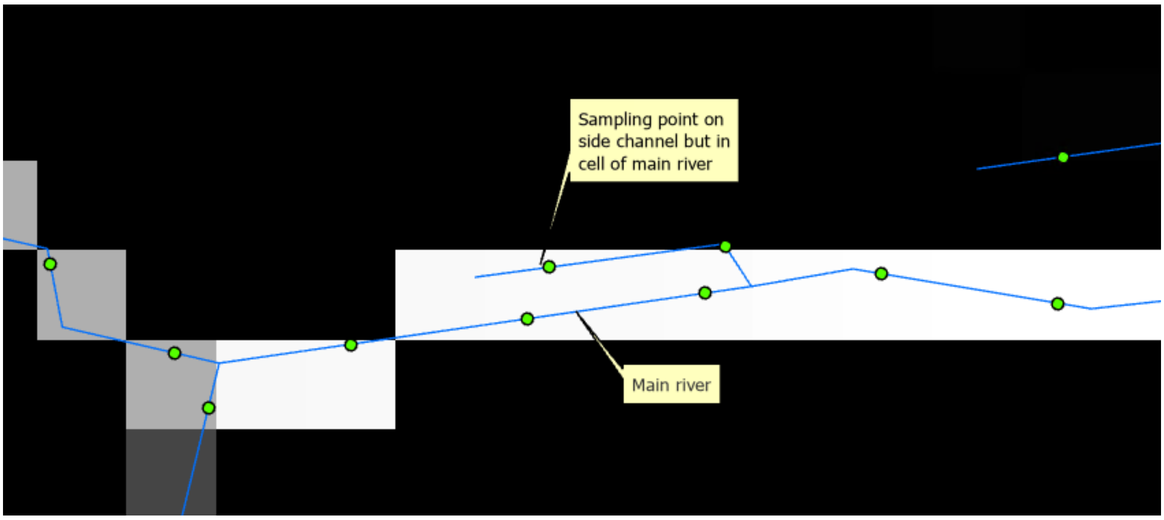


Figure 3 A sampling point on a side channel, it is intersecting the flow accumulation raster cell (white cells) for the main river.

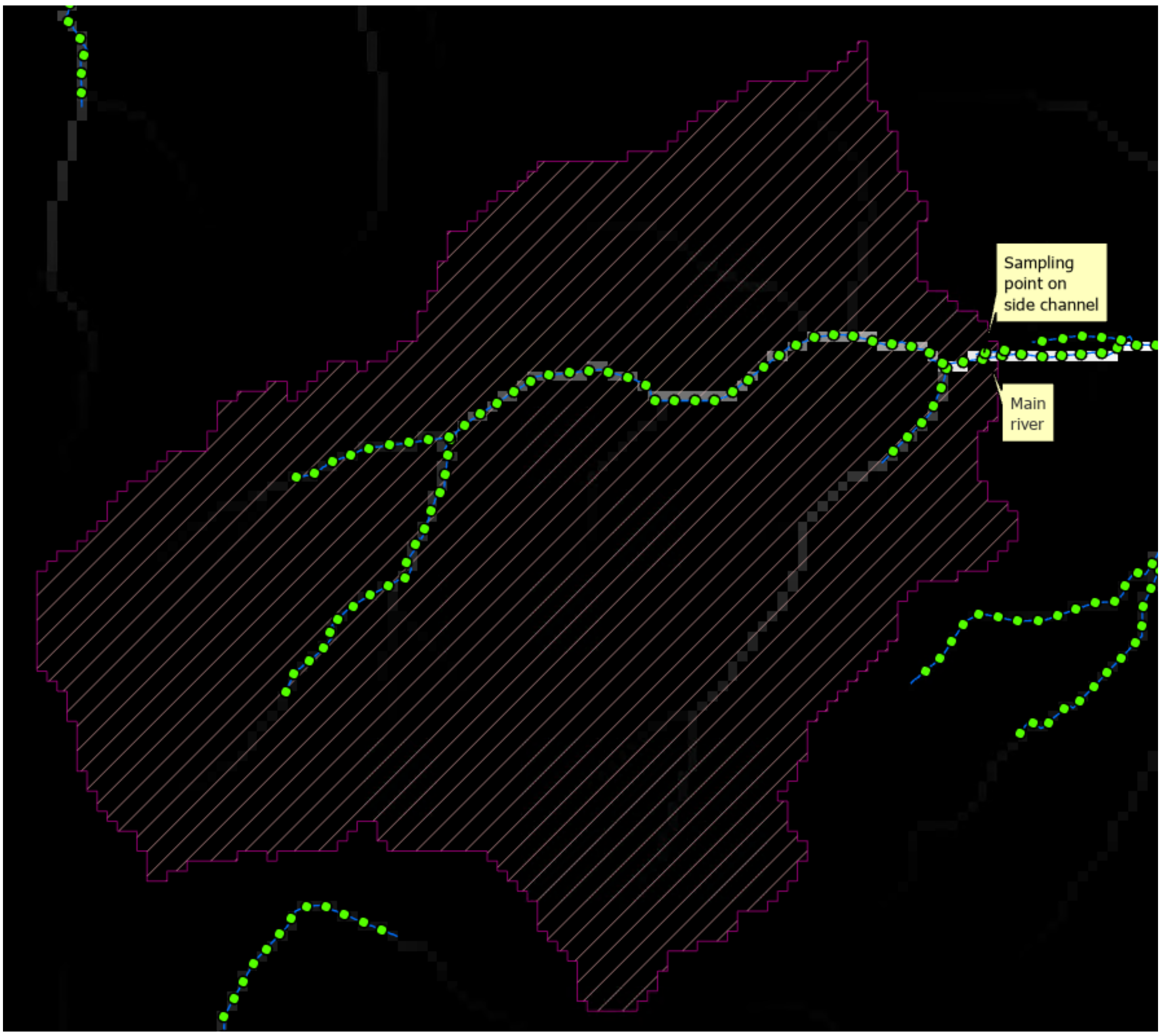


Figure 4 The catchment area for the sampling point on the side branch (Figure 3), we see that due to the snapping process catchment delineation started on the main stem and generates a catchment area. Yet this illogical catchment fulfils the primary test of is the source in the catchment?

Quality Check 2: using a distance ratio to identify candidates for rejection

A second quality control step was introduced to identify those sampling points on tributaries but wrongly associated with the main stem. (Figure 4). A ratio is calculated of the distance to source for the sampling point (an attribute encoded into the network by RivEX), to the diameter of the circle that encloses the catchment boundary, (which is a crude approximation of distance to source within the catchment polygon (Figure 5)). A small ratio value would indicate a potential candidate for rejection. The inverse scenario is a large distance to source divided by a diameter from a small circle enclosing a tiny catchment area, such ratios will typically be above 1.0 but these catchments always fail the '*is source in polygon*' test.

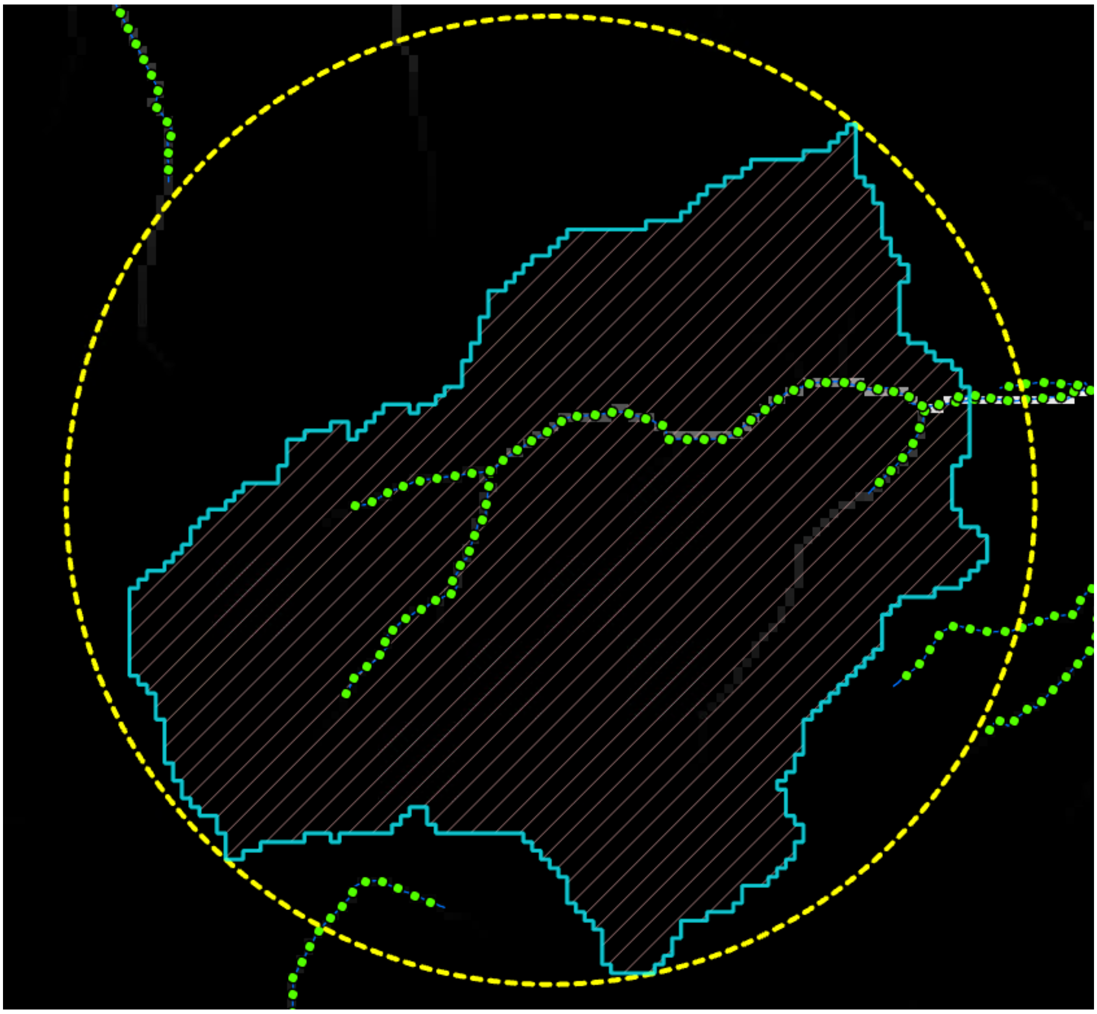


Figure 5 An example dividing the distance to source as encoded in the ORN by the diameter of the circle that encloses the catchment polygon. Using the failing site example in Figure 4 the ratio for the point = 41m / 5533m = 0.0074. A short distance to source divided by a larger diameter from a large catchment.

Quality Check 3: counting sources in candidate rejections

The use of a ratio (from Quality check 2) to identify erroneous catchments can return false positives, such as sampling points *very close* to their source (see example in Figure 6).

To reduce the number of false positives, we set a minimum ratio value of 0.1 for sampling points with their source in the catchment.

The value of 0.1 was chosen after a manual inspection of the values across a set of catchments to identify a general threshold. A larger value would have selected a greater number of valid catchments that would be tested and erroneously identified as false positive.

For a sampling point very near its source and creating a logical catchment it would have one and only one source point in the polygon. A sampling point with a small ratio value but clearly generating a much bigger catchment off a nearby main stem, this would have multiple sources and thus a candidate for rejection.

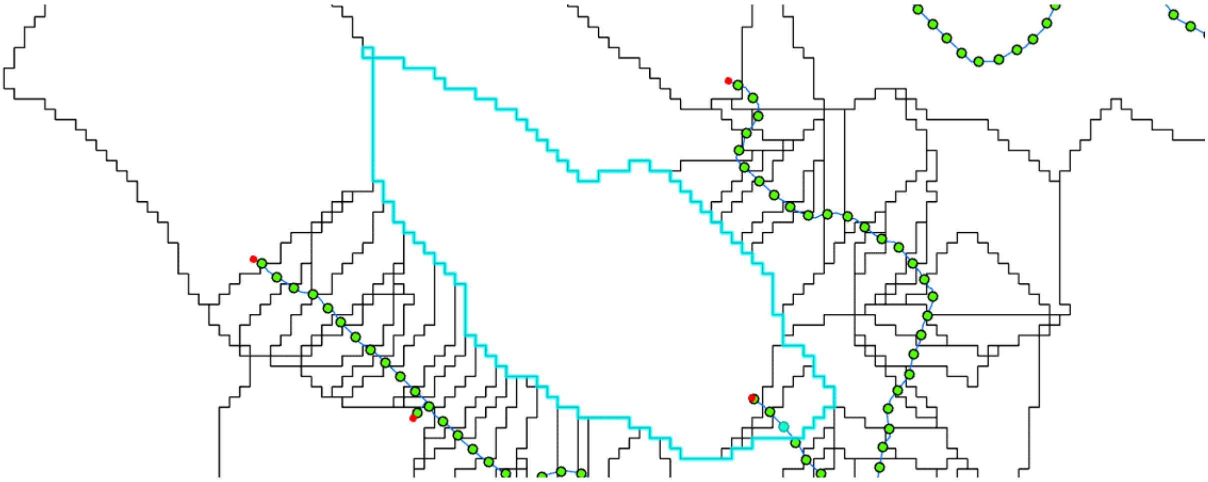


Figure 6 A genuine catchment for a sampling point (blue) very close to its source (red point) will generate a ratio with a small value from the distance along network/diameter of enclosing circle for catchment.

Quality Check 4: resolving cross catchment errors

The vector lines of the ORN are used to stream burn into OS Terrain 50 which has a cell resolution size of 50m. Channels in the ORN network from neighbouring catchments that are within 50m of each other create a flow accumulation raster that connects across river catchments. This creates an illogical scenario of a catchment being generated which merges into the downstream direction of the neighbouring catchment. This error was significant between the River Dart and Tamar, the River Clyde and Tweed and the Yorkshire Derwent and sea cut near Scalby. This error has a compounding effect causing all downstream sites to fail to generate a logical catchment.

Figure 7 shows tributaries of the Tamar (blue) and Dart (green) passing each other so closely that they occupy the same pixel in the underlying AGREE DEM. This causes the catchment delineation process within the Tamar to erroneously track downstream into the neighbouring River Dart catchment (Figure 8).

For these three regions an AGREE DEM was rebuilt using only the ORN lines for the specific catchment, rather than the entire network. The flow direction and accumulation rasters were worked up and the subset of failing sampling points were re-run through the script logic. This allowed for further 2800 catchments to be delineated logically and merged back into the master dataset.

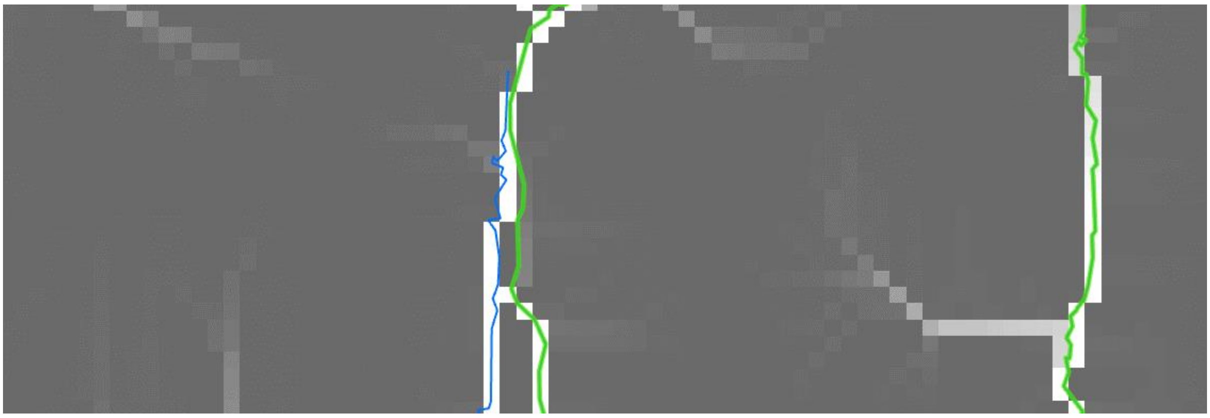


Figure 7 River Tamar (blue) and River Dart (green) are within 50m of each other on a tributary source, this causes the flow accumulation raster to route across catchment.

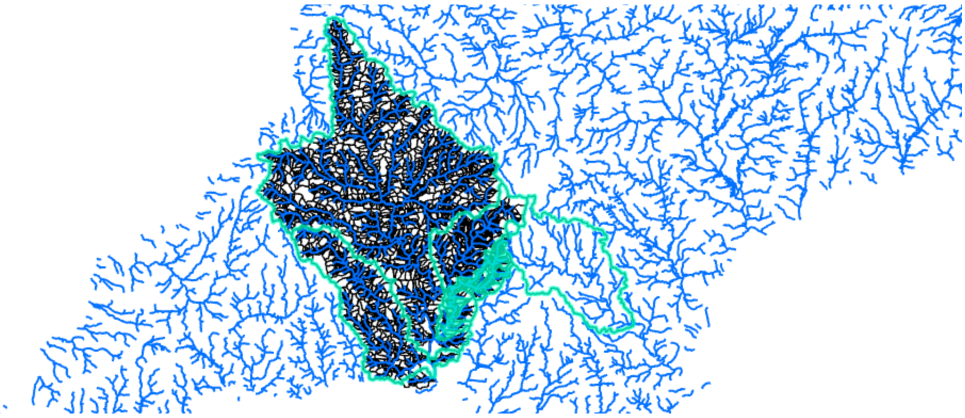


Figure 8 The selected catchments (cyan) identify cross catchment inclusion of neighbouring river catchment caused by sources within 50m of each other. Here we see the Tamar including a significant region of the River Dart.

A script that searched for this cross-catchment error was also created and run on the full catchment data with any found flagged to be rejected.

Step 3: final merge of data, check and export

As noted above, processing was distributed over two computers, their processing running locally for best performance. A script was created to merge the catchment data into a single dataset and drop all catchments where quality control field qcOK equalled "N".

Quality Check 5: a manual scan along the coastline

With catchments merged into a single dataset and the initial results filtered to reject those that had been tagged as failing the quality control checks, a manual scan of the remaining catchments was done by tracking around the coastline of the UK and spotting issues with catchments generated in estuarine regions. These catchments had passed all the logical checks yet had nonsensical boundaries (Figure 9). These were manually selected up and the quality control field (qcOK) updated to tag them as rejected.

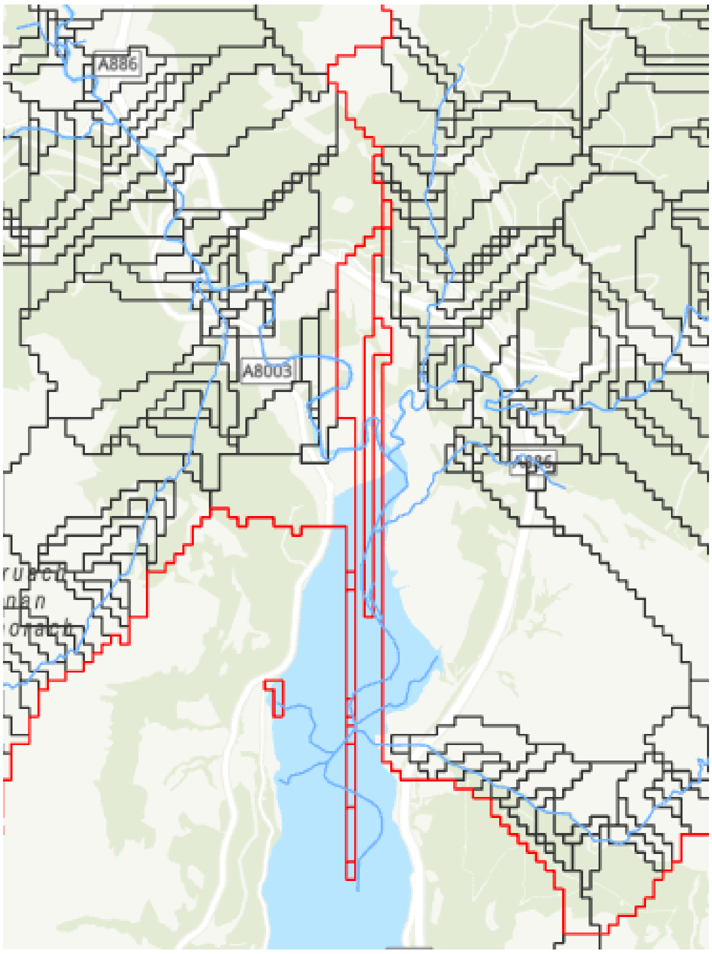


Figure 9 Estuarine section of the river Ruel, Scotland created unusual catchments (red) but fulfilled all previous logic checks. Such catchments were identified during a manual scan and tagged as rejected.

The manual scan would have flagged further catchments to be rejected, i.e. field qcOK was reset to "N".

A simple model was built that did a final export from the merged data exporting only polygons where qcOK equalled "Y", dropped the quality control fields and built a file geodatabase relationship class between the sampling points and their catchments.

Step 4: preparing the final release

The dataset was built and is maintained using ESRI ArcPro. Attribute indices and metadata in inspire format were added to the catchment layer. The Feature Class was compressed to save space, this makes it a read only dataset. If you want to edit the data, you need to uncompress the Feature Class[ƭ].

Table 1 lists the fields in the final dataset, note all quality control fields used to track quality control/processing have been removed to simplify the final release version.

A GeoPackage was prepared for QGIS users and user instructions developed under the **User guidance** section of this report.

[ƭ] = https://pro.arcgis.com/en/pro-app/latest/tool-reference/data-management/uncompress-file-geodatabase-data.htm

Table 1 Fields in the ORN catchment dataset.

Field name	Field Type	Description
OBJECTID	Object ID	A unique row ID maintained by the Geodatabase.
Shape	Geometry	The polygon geometry field, has no Z or M values.
Shape_Length	Double	The perimeter length of the catchment polygon in m, maintained by Geodatabase.
Shape_Area	Double	The area of catchment polygon in m^2, maintained by Geodatabase.
SampleID	Long	A unique sequential integer assigned to the sampling point as it is created by RivEX. This ID is attached to the polygon and has a 1-to-1 relationship.
Country	Text	ISO country codes used to identify which country the water course is within. This code was assigned to the water course and joined to this dataset from the sampling point. A catchment polygon may cross a national boundary, so this code is for the network not the catchment boundary, but can be used as a useful filter on the dataset if needed.
CatchID	Long	The catchment ID the polyline belongs to. This is built by RivEX. A catchment are all the connected lines that drain to a single mouth. Mouth is typically the tidal limit of the network but can be internal drainage networks that do not connect to the wider river basin network.

Other reasons for catchment delineation failure

Significant numbers of sites can fail to generate logical catchments due to poorly defined flow path in the flow accumulation raster; the River Tamar is a good example. The reason for this is that a significant section of the centre line in the ORN falls within the estuary (Figure 10). In the OS Terrain 50 dataset the estuary area is given an elevation value of -1.6m. This area is highlighted in red in Figure 11. Figure 12 shows the poor alignment of the ORN vector network with the stream burnt flow accumulation raster within the flat negative elevation region of the OS Terrain 50. With such poor alignment many sites fail to generate logical catchment boundaries. These were excluded from the final dataset.

Figure 10 River Tamar catchment, its lower section falls within the estuary.

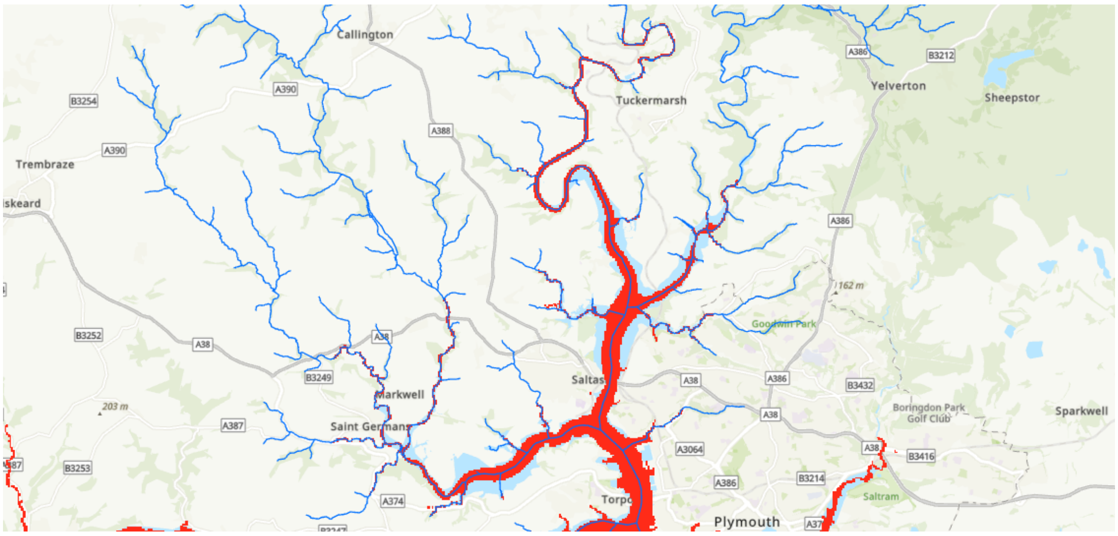


Figure 11 Red pixels are where OS terrain 50 records -1.6m or less.

Figure 12 Zoomed in section of the lower Tamar. Flow accumulation raster (white pixels) overlaid with the red mask identifying OS Terrain 50 pixels with an elevation of -1.6m or less and the ORN vector blue line. It is clear that the ORN and flow accumulation have poor alignment within the negative elevation region of the OS Terran 50.

Results

This section reviews some preliminary statistics of the final dataset to understand where catchment generation failed. 1,472,056 points were fed into the delineation scripts and after all the automated and manual checking 1,299,024 catchment polygons remain. Thus 11.8% of the sampling points failed to create a sensible catchment and were rejected. The remaining catchments are guaranteed to contain the source of the river for the sampling point as defined by the RivEX source ID field, although sampling points in areas of low topographic relief and human altered drainage can, and do, have unusual catchment shapes.

There are 3,558 unique catchments coded by RivEX into the ORN. 2 of these are shorter than 100m and consequently have no sampling points mapped to them, so 3,556 catchments were fed into the delineation script. The length of network for each catchment defines the number of sampling points that fall within it. The least number of sampling points was 3 whilst the Yorkshire Ouse had the most at 130,677.

Figure 13 plots the number of catchments by the % failed within a catchment. For example, 13 catchments out of 3,556 had between 70-75% of their sampling points fail to delineate a logical boundary and were rejected. We can see that 59 catchments had between 95-100% failed catchments. 2,356 (66%) of catchments had between 0-10% of their sampling points fail to delineate logical catchments.

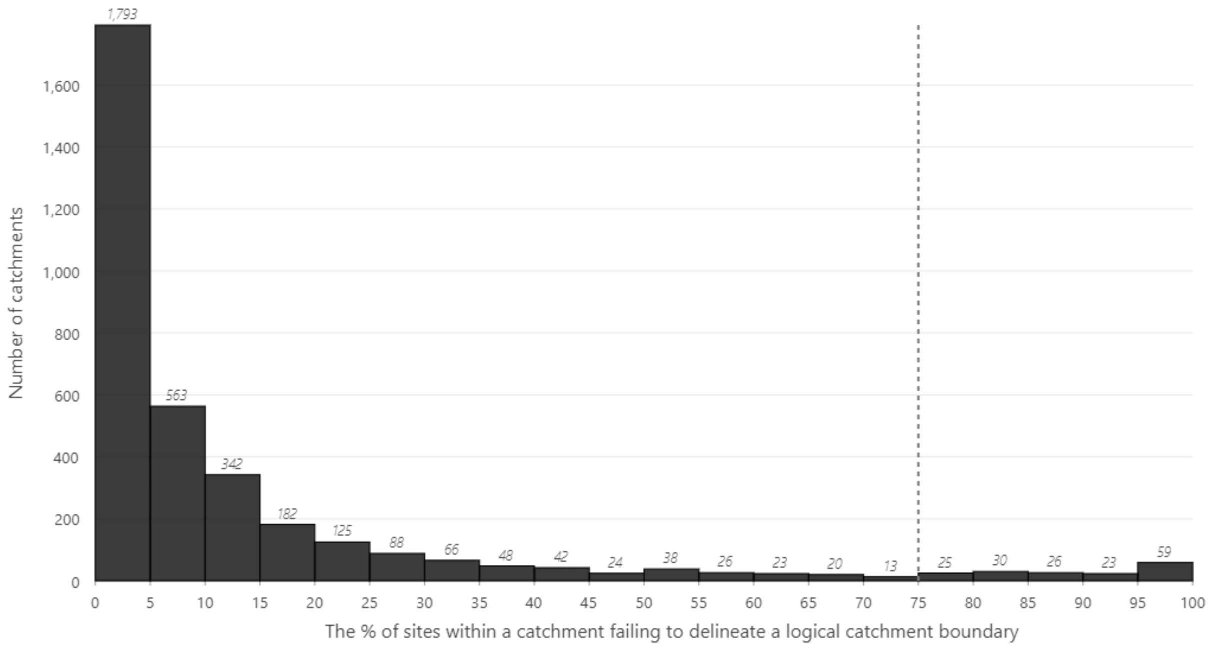


Figure 13 The number of catchments plotted by the % number of sites failing to delineate a logical catchment boundary.

Reviewing the 163 catchments where 75% or more of their sampling sites failed to generate a logical catchment (Figure 13) we can get a sense of where the worst catchments were. Surprisingly these catchments, where more than 75% of sampling points failed, were typically small coastal streams and distributed nationally (Figure 14).

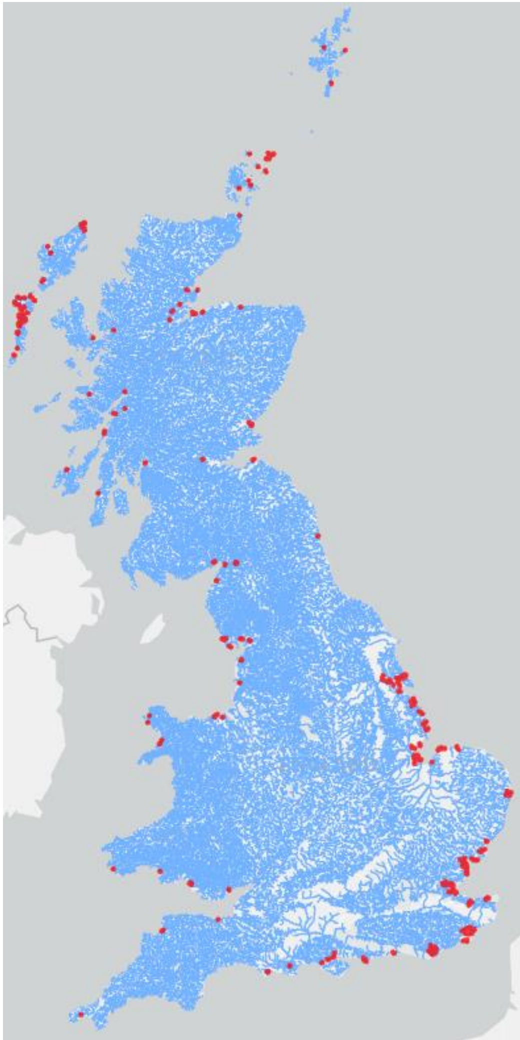


Figure 14 Catchments within the ORN highlighted in red are catchments which had more than 75% of their sampling points fail to generate logical catchments.

An alternative way to look at where failure to generate logical catchments were occurring most is to plot the mean % failed by WFD river basin (Figure 15). For example, the South West WFD river basin had a mean % failure within a catchment of about 9%, whereas in the Anglian WFD river basin catchments had an average of about 56% failure within a catchment.

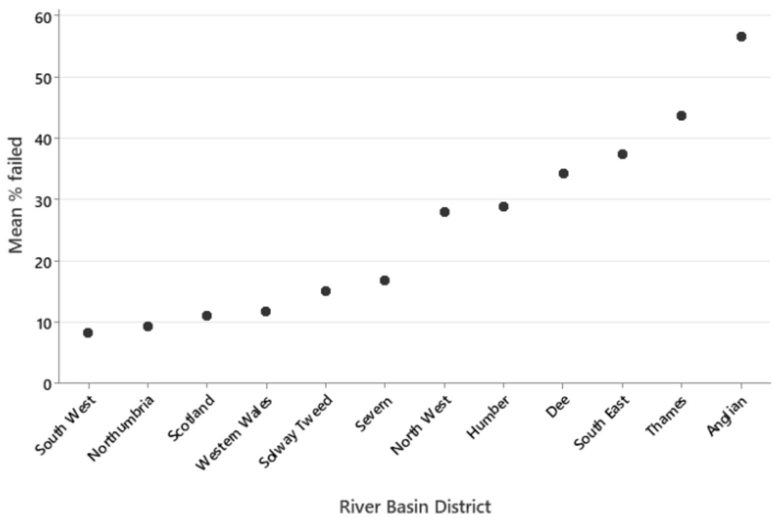


Figure 15 The mean % failed within a catchment grouped by WFD river basin.

Unsurprisingly, the 3 three river basins that had the highest average percentage failure to generate logical catchment boundaries were the South East, Thames and Anglian; regions known to have had significant drainage alteration.

A count on the number of catchments within each WFD river basin categorised by their % failure rate is shown in Figure 16. For example, Northumbria WFD river basin has 39 catchments, just over 80% of these had the least failure rate of less than 10% of the sampling points failing to generate a logical catchment. The other extreme is the Anglian WFD basin where just under 10% of the 61 catchments had the least failure rate of less than 10% of the sampling points failing to generate a logical catchment. Anglian stands out as having the greatest mix of failure rates (a % of catchments in each 10% group) whilst the Dee\Severn have the least (a % of their catchments within 5 10% groups).

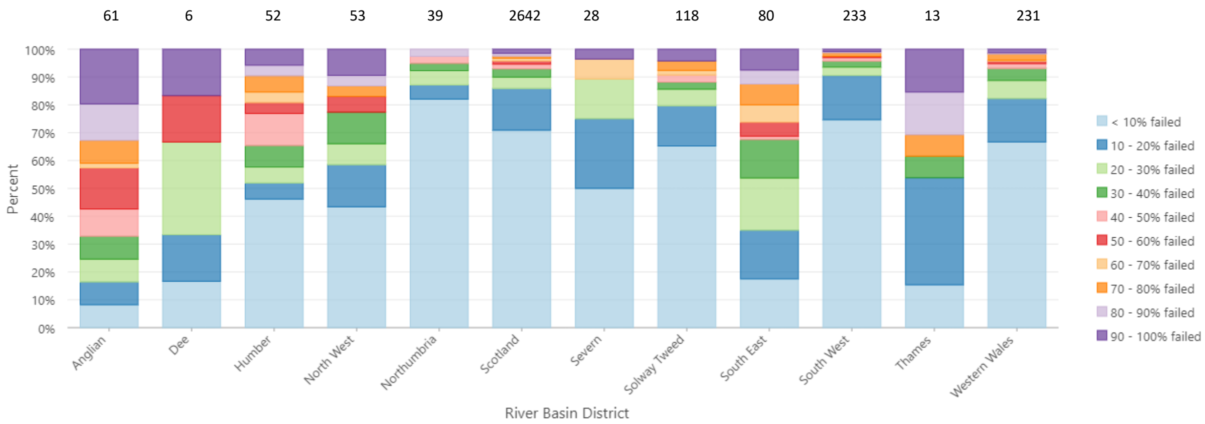


Figure 16 Percentage of catchments for WFD river basins broken down into 10% categories of their % failure. Numbers at top of stack are the number of catchments within the WFD river basin, Scotland is an order of magnitude large than all other basin regions.

Conclusion

1,299,024 catchment polygons were generated, which means 11.8% of the sampling points failed to create a logical catchment area. Preliminary analysis indicates 66% of all catchments in the ORN network had between 0-10% of their sampling points fail to delineate logical catchments. 163 catchments where 75% or more of their sampling sites failed to generate a logical catchment were typically small coastal streams distributed nationally. When viewed in the context of the WFD river basins South East, Thames and Anglian, regions known to have had significant drainage alteration, contained the catchments which had the highest mean failure rate for generating logical catchments. Northumbria, South West, Western Wales and Scotland contained the most catchments which had the least failure rates.

These results indicate that, if you are working on rivers in the South East, you are most likely to experience “sparse data”, which could bias automated aggregation of data if you are stepping along a particular river.

The 1-to-1 relationship through the SampleID field provides a powerful look up of catchment boundary for the 1,299,024 sampling points across the entire Open River Network. Minimal scripting could drive aggregation of catchment wide statistics at a 100m interval across the river network. A sample script can be found at the end of the User guidance section of this report.

User guidance

The catchment dataset is a large and complex layer with over a million overlapping polygons. It is not meant for general display but to be a source of quality controlled, pre-computed catchments which you can select from. You might only be interested in a single boundary to help display data in a report, or, with minimal scripting, step through a sequence of sampling points and their catchments to query and aggregate data at catchment level. An example script is provided at the end of this guide.

This section provides guidance on using the ORN catchment data in ESRI ArcPro 3.4 or QGIS 3.28. Guidance shows how you can take advantage of the 1-to-1 relationship of the sampling point ID with its catchment polygon.

Accessing catchment data in ESRI ArcPro

If you are an ArcPro user you can access the catchment data using the supplied relationship class or if you are working from a copy of the original download, build your own map relationship. With the relationship in place, it is easy to pass selections in the point data through to the catchment data.

Accessing the catchment data using the supplied File GeoDatabase relationship class

1. Add the ORN, ORN_Sampled_100m and ORN_Catchments to a map, ensure you map coordinate system are in British National Grid.

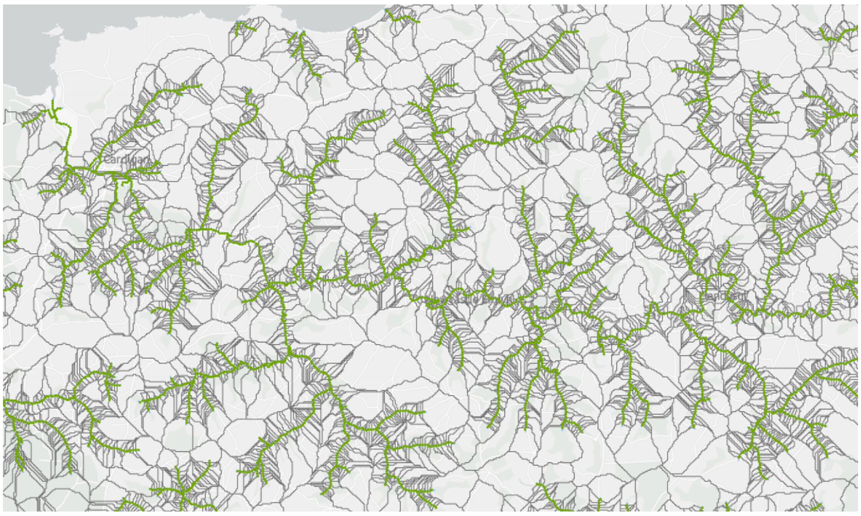


Figure 17 A zoomed in section of the lower reach of the Afon Teifi. ORN network, sampling points and catchment areas are displayed.

2. Within the ORN_v2 File GeoDatabase is a relationship class called **relORN**. You do not need to interact with it. This relationship class establishes a 1-to-1 link between the sampling points and their catchment polygons. You can verify the relationship exists by right clicking on ORN Sample Points in your map > Properties > Relates.

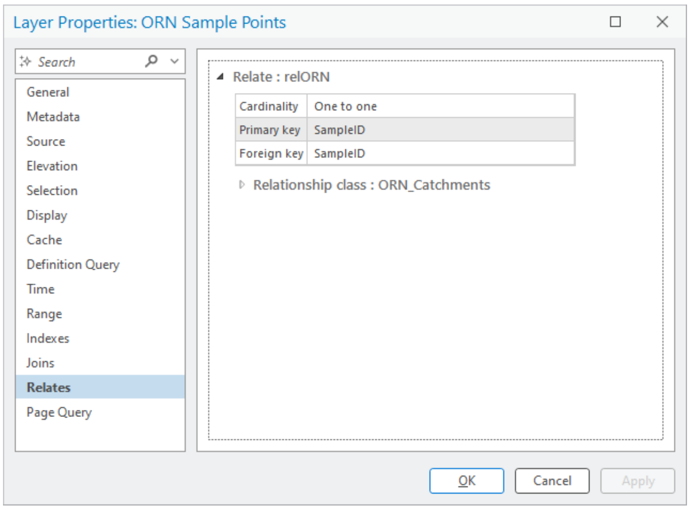


Figure 18 The layer properties of ORN sampling points clearly shows a valid relationship exists between the points and their catchment polygons.

3. Select any sampling point, go to the Data tab for the sampling point layer, note the Related Data button is active and clicking on it reveals the ORN_Catchments relate.

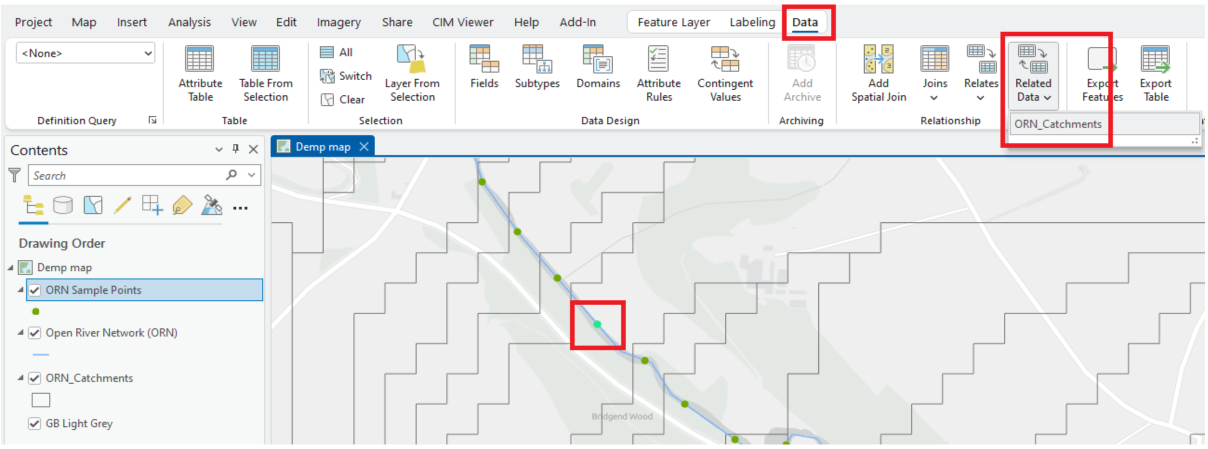


Figure 19 With a point selected the Related Data tool becomes active under the Data tab.

4. Selecting the relate will pass the selection in the point layer through to the polygon with the same ID. If no polygon exists, then no selection occurs within the catchment dataset. If you had multiple points selected then if the polygons exist, multiple polygons are selected.

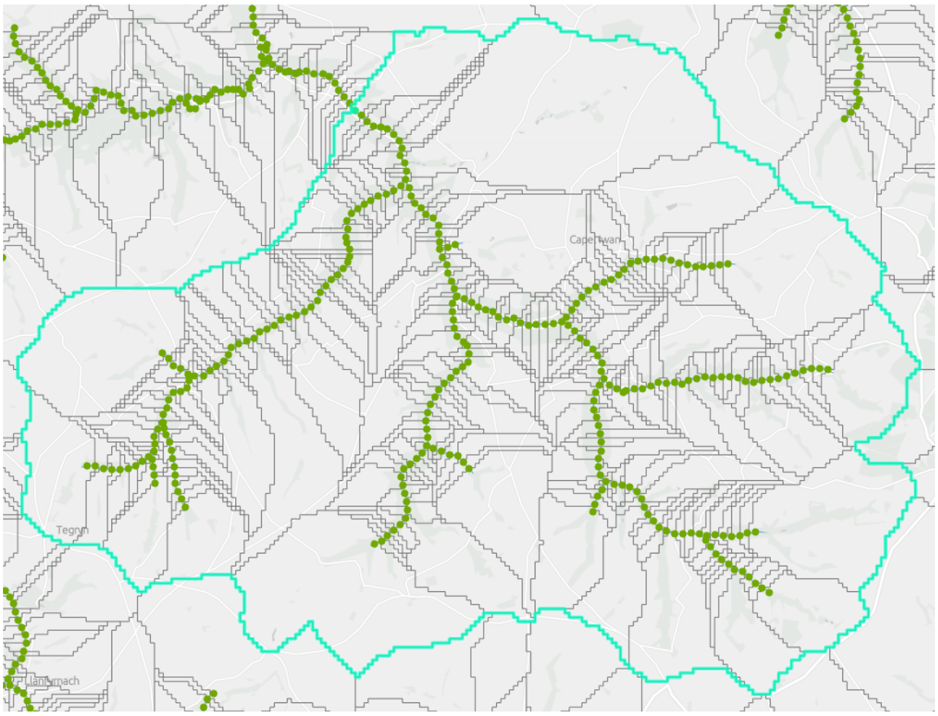


Figure 20 The selection on the point is passed through to the catchment polygon.

5. You can now export this selected polygon to a separate dataset, use the ArcPro 3.4 Generate Definition Query from Selection[†] tool to filter the layer or feed the selected polygon directly into a geoprocessing tool to aid you in further analysis.

[†] = https://pro.arcgis.com/en/pro-app/latest/tool-reference/data-management/generate-definition-query-from-selection.htm

Create your own map relationship and access the catchment data

If you have exported the ORN network, its sampling points and catchment areas to a new database it's quite possible you have lost the original File GeoDatabase relationship class that came supplied with the data. If your license level allows it, you can rebuild such a relationship. Alternatively, and not restricted by license level, you can quickly build an in-map relationship using the relate tool. This section guides you through the process.

1. Add the ORN, ORN_Sampled_100m and ORN_Catchments to a map, ensure you map coordinate system are in British National Grid.

2. Ensuring the ORN Sample points layer is highlighted in the Contents Pane, click on the Data tab > Relates > Add Relate.

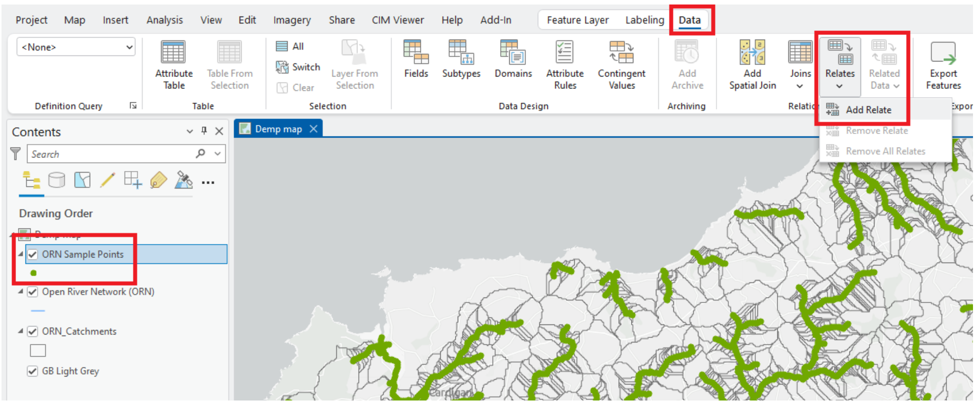


Figure 21 Running the Add Relate tool.

3. Complete the tool as below and press OK to run it. The Relate Name can be anything sensible. Nothing will change on the screen; all you have done is establish a relationship between the point and polygon layer.

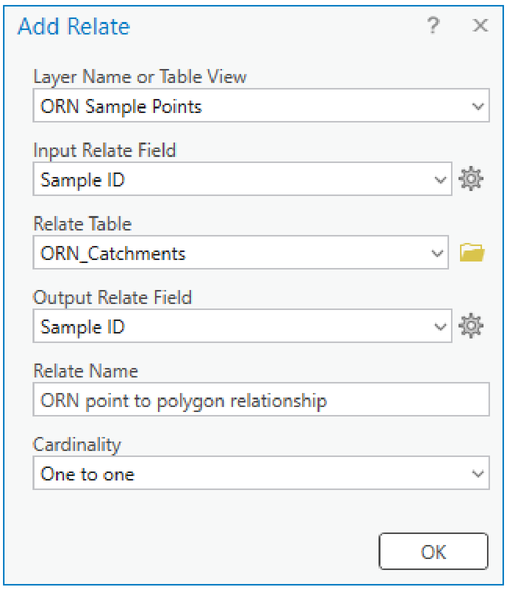


Figure 22 The completed Add Relate tool. This will establish a 1-to-1 relation between the sampling point and its catchment polygon.

4. Select any sampling point, go to the Data tab for the sampling point layer, note the Related Data button is active and clicking on it reveals the ORN_Catchments relate (Figure 19).

5. Selecting the relate will pass the selection in the point layer through to the polygon with the same ID (Figure 20). If no polygon exists, then no selection occurs within the catchment dataset. If you had multiple points selected then if the polygons exist, multiple polygons are selected.

Accessing catchment data in QGIS

If you are a QGIS user, you are likely accessing the data from the GeoPackage that can be downloaded from the ORN website. This section guides you through the process of establishing a relationship between the point and polygon layers.

Ensure fields are indexed

To ensure relationships process fast it is important to ensure the SampleID fields in the sampling points and catchment layers have been indexed. If you are accessing the ORN data directly from the GeoPackage downloaded from the ORN website then this has been done for you, otherwise you need to confirm an index exists and if not create one.

Checking field for existing index

1. Go to menu Database > DB Manager…
2. Expand GeoPackage, if ORN_v2.gpkg is not visible then right click on GeoPackage and select New Connection. Navigate to ORN_v2.gpkg and select OK. DB Manager should look as below.

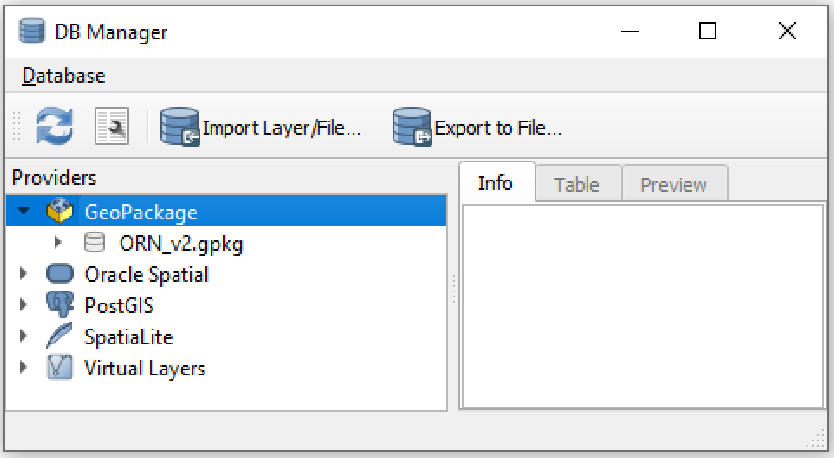


Figure 23 DB Manager in QGIS showing that it is connected to ORN_v2.gpkg.

3. Expand the ORN_v2.gpkg database and select layer ORN_Catchments. Under the Info tab scroll to the bottom and see if SampleID is indexed.

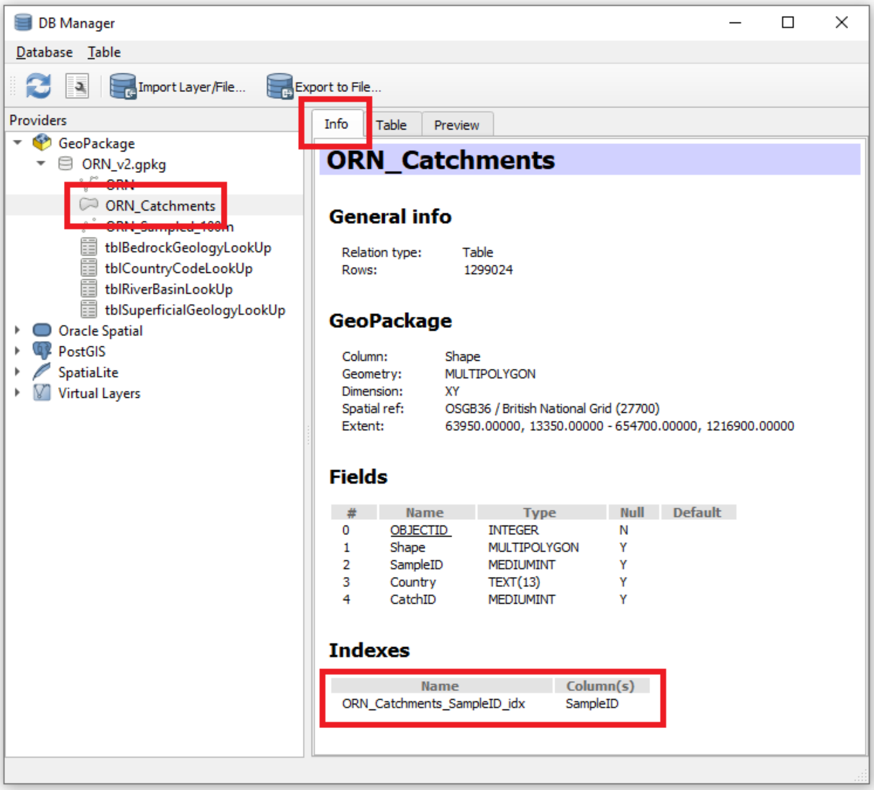


Figure 24 This layer has the field SampleID indexed.

4. Now verify if ORN_Sampled_100m has its SampleID field indexed.

5. If both SampleID fields are indexed, then it's OK to close DB Manager.

Adding an index to SampleID field

If you have found that SampleID field is missing its index in either point or polygon data, it is strongly recommended you add an index as this significantly improves performance.

1. Go to menu Processing > Toolbox.

2. Type index into the search box on the Processing Toolbox panel and double click on the Create attribute index tool found under Vector general.

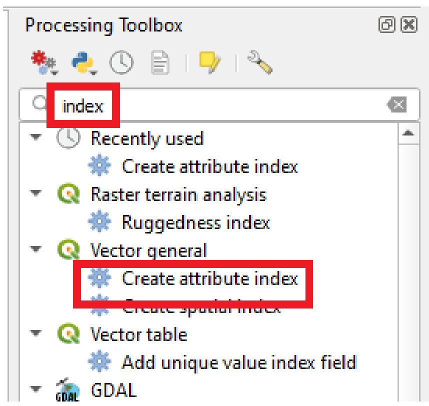


Figure 25 Accessing the Create attribute index tool in Toolbox.

3. Select the layer and set SampleID field to be the field that will be indexed, press Run to execute the tool. If required do this for the ORN_Sampled_100m and ORN_Catchments layers.

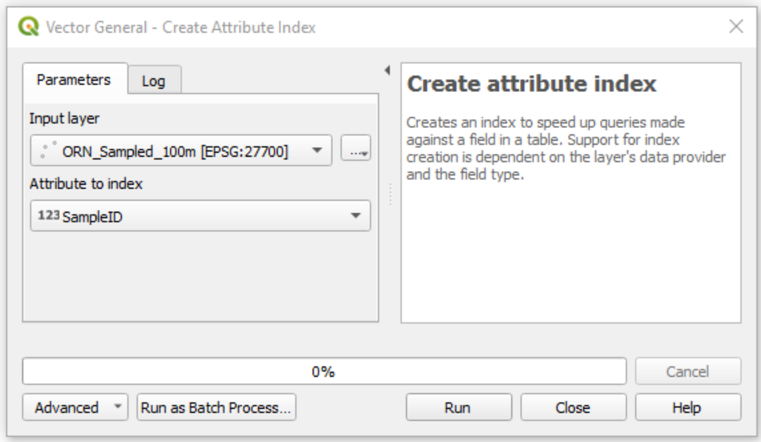


Figure 26 Running the Create attribute index tool on the ORN_Sampled_100m layer for the SampleID field.

Creating the relationship between sampling point and catchment polygon

In QGIS, relations are stored at the project level and need to be defined. This means the relationship is saved in the QGIS project and it is not a property of the source data in the GeoPackage.

1. Start a new QGIS project, navigate to the ORN_v2.gpkg in the browser panel, right click on it and select Add Connection.

2. Select up the ORN, ORN_Catchments and ORN_Sampled_100m layers, right click and select Add Selected Layers to Project.

3. Optionally, change symbology from defaults, we recommend for the catchment data to set the simple fill style from solid to no brush.

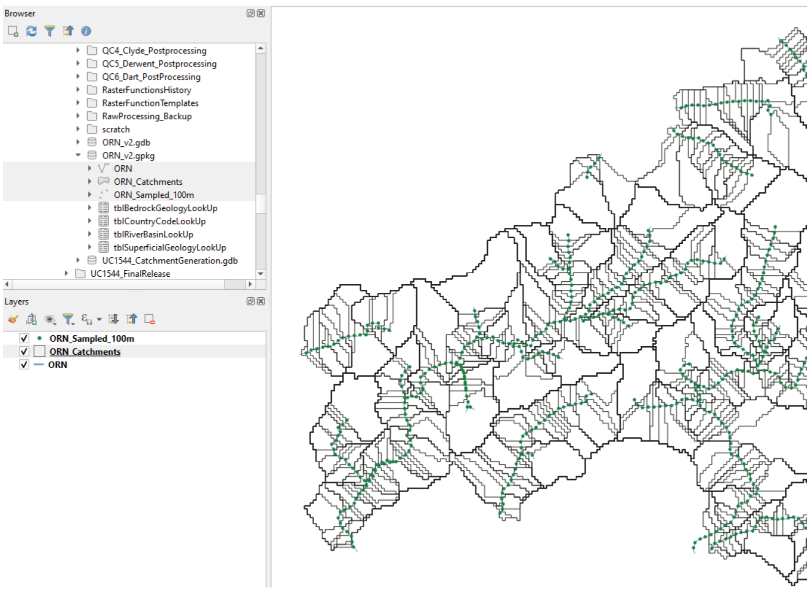


Figure 27 The three ORN layers loaded and symbolised in QGIS, location zoomed into St. Davids in Pembrokeshire, Wales.

4. Go to menu Project > Properties…

5. Select Relations tab, click on the Add Relation button and complete the Add Relation tool as shown below, press OK to accept.

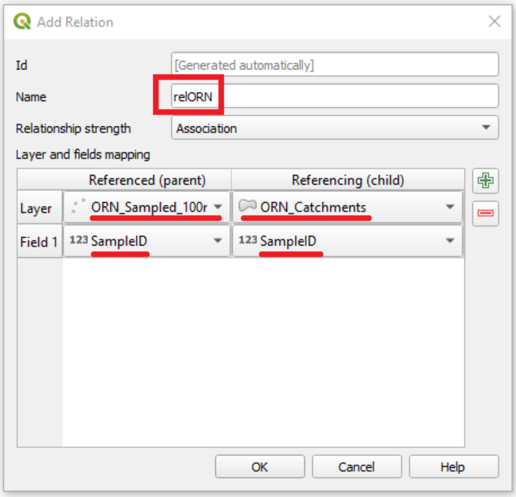


Figure 28 Adding a relationship, it is the Sampling points that are the parents and catchments are the children.

6. In the project properties dialogue, click on Apply then OK.

7. If you do not save the project file, then the relationship will be forgotten, and you would need to rebuild it next time you access the datasets.

Using the relationship

To use the relationship, you need to have preinstalled a plugin which installs a new tool bar. With the plugin activated you can easily pass selections in the point layer through to the catchment polygon layer.

Installing the Select by relationship plugin

1. Go to menu Plugins > Manage and Install Plugins.

2. Type *select by relationship* into the search bar to filter the plugins. Select the plugin and click on install plugin.

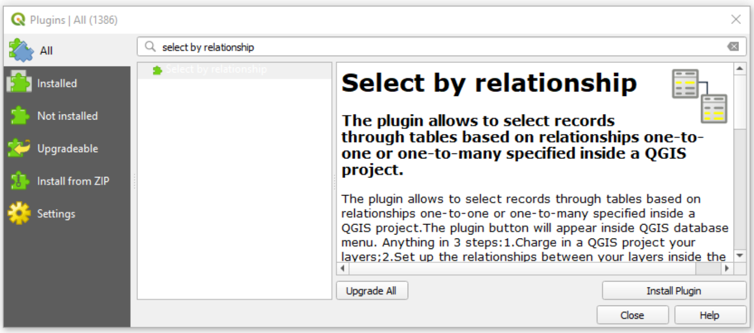
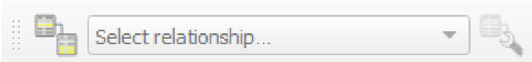


Figure 29 Installing the Select by relationship plugin.

3. A new tool bar appears in the toolbar section of QGIS. It contains the Allow relation, a drop-down to choose which relation you wish to use and a settings button.



Running a relationship

1. Ensure ORN_Sampled_100m layer is selected in the Layers panel.

2. Click the Allow relationship button to enable relationships. This button allows you to quickly turn on/off the use of relationships in your project.

3. In the drop-down tick **on** the relationship.

4. Click the relationship settings button and ensure **Select Features(s) of Child from Parent Layer** is ticked **on**. You only need to do this once.

5. Use the select tool and select 1 or more sampling points. Note their corresponding catchments are selected based on the shared ID in the SampleID field (Figure 30).

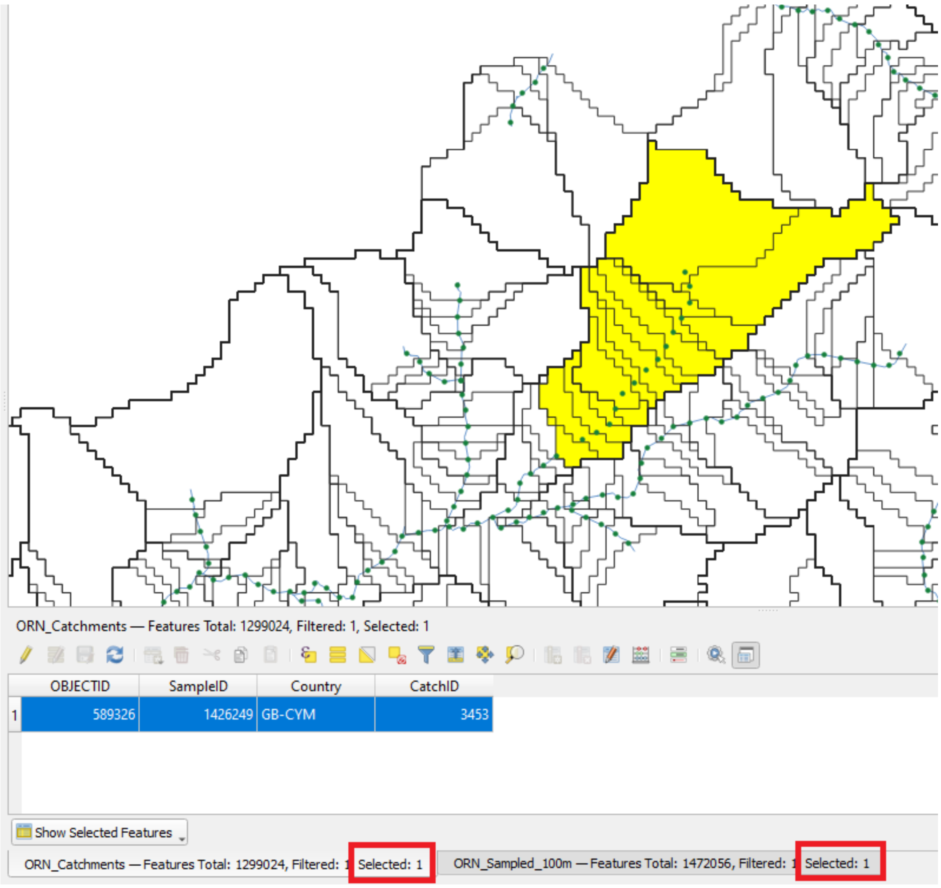


Figure 30 The select by relationship has passed the selection from the sampling point layer to the catchment area. Although it looks like multiple catchments are selected there is in fact just 1 as indicated by the selection numbers in the table tabs. This multiple selection appearance is simply due to the no brush fill applied to the catchment symbology.

An example of automating processing using python in ArcPro

A user can take advantage of a sampling points 1-to-1 relationship with its catchment area through python scripting. This section shows a minimal sample script written in python using arcpy in ArcGIS Pro 3.4.

Figure 31 shows a data scenario that you might want to replicate, extracting habitat data from each catchment area for a subset of selected sampling points. This example uses habitat data to demonstrate the principal, but it could be *any spatial dataset* you need to collect information at catchment level.

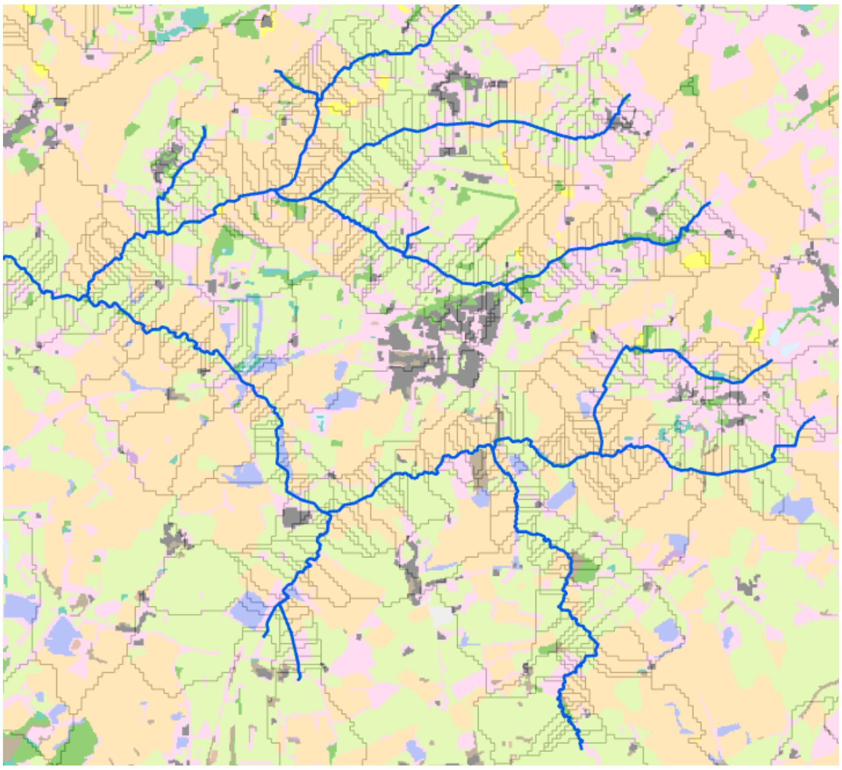


Figure 31 The ORN displayed along with its catchments (sampling points not shown for clarity) on top of the Natural England (2025) Living England Habitat Map (phase 4). The minimal sample script shown below, will automate the summation of area by habitat type, using the catchment area of each selected sampling point.

```
"""
  Description:
    A sample script to show how to exploit the 1-to-1 relationship of sampling point with catchment polygon.
    In this example the catchment area is used to clip out habitat data and summarize the results. The
    habitat data is the Living England Habitat Map Phase 4 accessed via the Living Atlas portal.

  Usage:
    [1] - Load layers into map
    [2] - Make a section on sampling points
    [1] - Run this code in the python console within ArcPro

  Created:
    28/02/2025

  Author:
    Duncan Hornby (ddh@geodata.soton.ac.uk)
"""
```

```
import arcpy

def main():
  try:
    # Declare input layers, these must exist in the map
    lyrCatchments = "ORN Catchments"
    lyrSamplePoints = "ORN Sample Points"
    lyrLEHMap = "Living England Habitat Map Phase 4"

    # Declare a temporary output feature class for clipped data
    fcTempOut = r"memory\TempClip"

    # Check Sampling points layer has a selection (avoids process everything!)
    resObj = arcpy.GetCount_management(lyrSamplePoints)
    n = int(resObj.getOutput(0))
    if n == 0:
      print("ERROR: Sampling point layer has no selection!")
      return

    # Cycle through selection and store Sample ID's in a list
    lstIDs = list()
    with arcpy.da.SearchCursor(lyrSamplePoints, "SampleID") as cursor:
      for row in cursor:
        lstIDs.append(row[0])
    print("Number of sampling points to process = " + str(len(lstIDs)))

    # Now use ID in list to step through the catchment layer.
    # Remember not all sampling points will have a catchment polygon
    for sampleID in lstIDs:
      print("Processing ID: " + str(sampleID))
      sQuery = "SampleID = " + str(sampleID)
      geom = None

      # This search will return either nothing (as there is no catchment polygon) or exactly 1 polygon
      with arcpy.da.SearchCursor(lyrCatchments, ["SampleID", "SHAPE@"], sQuery) as cursor:
        for row in cursor:
          geom = row[1]

      if geom is not None:
        # geom is a polygon so use it to clip habitat layer into a temporary in-memory layer
        resObj = arcpy.Clip_analysis(lyrLEHMap, geom, fcTempOut)
        if resObj.status != 4:
          print("ERROR: clip failed!")
          return
        else:
          # Add a new field of type Double
          arcpy.AddField_management(fcTempOut, "NEW_AREA", "Double")

          # Populate new field with area of clipped polygons
          arcpy.CalculateField_management(fcTempOut, "NEW_AREA", "!shape.area!", "PYTHON3")

          # Summarize area by habitat class and store as a dBase file in local folder
          with arcpy.EnvManager(addOutputsToMap=True):
            tblOut = r"C:\Temp\HabData_" + str(sampleID) + ".dbf"
            arcpy.Statistics_analysis(fcTempOut, tblOut, [["NEW_AREA", "SUM"]], "A_pred")

  except Exception as e:
    print("ERROR: A fatal error occurred in main()")
    print(str(e))

if __name__ == '__main__':
  # Set some environment settings
  arcpy.env.overwriteOutput = True
  arcpy.env.addOutputsToMap = False
  main()
  print("DONE!")
```

References

ESRI. (2024). ArcGIS Pro Desktop: Release 3.3 Redlands, CA: Environmental Systems Research Institute.

ESRI. (2024). ArcHydro. Environmental Systems Research Institute. https://www.esri.com/en-us/industries/water-resources/arc-hydro/downloads.

Hornby, D.D., Januchowski-Hartley, S., Naura, M., Hill, C.T., and Cavazzi, S. (2023) OpenRiverNetwork (ORN): a topologically correct and attributed open-source river network for Great Britain.

Hornby, D. D. (2024). RivEX (Version 1.11) [Software]. Available from https://www.rivex.co.uk

Maidment (1997). https://www.caee.utexas.edu/prof/maidment/gishydro/ferdi/research/agree/agree.html

Natural England. (2025). Living England Habitat Map (Phase 4) [Data Set]. https://naturalengland-defra.opendata.arcgis.com/datasets/Defra::living-england-habitat-map-phase-4/about