A Technical Overview of HR_Catchments

Introduction

The purpose of this document is to provide a brief high-level technical overview of the Geofabric Hydrological Report Catchments product (HR_Catchments). The document's main goal is to convey enough of a general understanding of the HR_Catchments product to enable a technically savvy GIS user to comprehend and make use of the product. A more comprehensive set of documentation, providing detailed explanations of the concepts and methods used during its construction (including illustrative graphics), is planned for release during 2013. Although a brief outline of the conceptual logic lying behind the HR_Catchments product is included within this document, this is not the document's main focus. A more extensive explanation of the conceptual logic can be found in related documentation (see below and on the Bureau of Meteorology's website).

Overall aim

The vision for HR_Catchments was to build a scale-independent national set of terrain-based catchments representing the natural surface hydrology of Australia. This set of catchments, together with their associated node-link network, would be constructed in such a way as to be consistently reproducible at all finer levels of detail and would thus form a stable 'Catchment Hierarchy' referencing framework.

Conceptual model

The HR_Catchments product is an attempt to implement some of the conceptual theory that forms part of a broader conceptual model of hydrological features known as HY_Features. Further detail concerning this proposed international standard conceptual model for sharing references to hydrological features is available in related documentation (Atkinson et al., 2012).

Contracted Nodes

The necessary first step in building a consistently reproducible framework was to identify a set of natural surface hydrology features possessing some form of reliable identity. For this purpose, a list of business rules were compiled and then applied to the features in the stream network datasets found within two sister Geofabric products: Surface Hydrology Network (SH_Network) and Surface Hydrology Cartography (SH_Cartography).

For a detailed explanation of these business rules, please consult related Geofabric documentation, however, in brief, the business rules that were used to identify a set of candidate features were:

- Named stream confluences or bifurcations.
- Named stream coastal outlets.
- Named inland sinks (coincidental with expert DEM-9S data sinks).
- Inflows and outflows of named water bodies (coincidental with expert DEM-9S data sinks).
- Expert knowledge.

Excluding cases where expert input was used, for a feature to become 'contracted' and considered for inclusion in the referencing framework, there needed to be evidence of the feature's existence within two distinct data sources. For example, a

confluence of two named streams in the AHGFNetworkStream feature class had to be matched with an identically named confluence in the AHGFMappedStream feature class.
The resulting set of contracted nodes arising from the application of this business logic can be identified in the Geofabric by the presence of persistent identifiers within the AHGFNetworkNode (SH_Network) and AHGFMappedNode (SH_Cartography) features classes. Where a logical contracted node feature is identified in both of the two stream networks, this is indicated via a common ConNodeID value in both of the aforementioned feature classes and an entry in the feature's MapNode/NetNode field pointing to the related alternate representation (it's HydroID). Contracted node features also have a confidence level value of 1 or 2, which is stored in the ConLevel field. Please note, that not all contracted node features have a representation in the AHGFMappedNode feature class.
The features contained within the AHGFNode (HR_Catchments) feature class are the subset of the total set of contracted SH_Network features that form the stable referencing framework.

Full Node-link Network

Once all of the nodes were assessed against the above criteria and the resulting set of contracted nodes had been identified, the next step was to build a node-link network from the SH_Network representations of the contracted nodes. In order to do this, the underlying stream network was traced for each contracted node until each possible trace path either reached another contracted node or an end-point in the network. Links were generated between the node being processed and each contracted node that was found. The end result of carrying this process out for all of the contracted nodes was a 'full' node-link network.
In many areas, particularly areas of low-relief, the resulting full node-link network was often found to be non-dendritic and in certain areas also very complex in nature. In other areas, such as in areas immediately adjacent to the coast, a lack of any contracted features resulted in a lack of any representation for groups of streams discharging to the sea. As a result, the necessary next step in the process was to apply further business logic in order to simplify the full node-link network to form a dendritic 'Catchment Hierarchy'.

Diffused Contracted Nodes

In addition to the contracted node features defined by the business logic already outlined above, an additional process has been used in order to identify a supplemental set of 'diffused' contracted nodes. These nodes are used to represent the diffused passage of water between two areas where the underlying stream network is either complex and/or unreliably defined. The three distinct types (DiffusType) of diffused contracted nodes used are as follows:

Diffused coastal node

In coastal areas, where drainage is often directed along short, poorly defined and unnamed channels to the sea, it is difficult to reliably contract stream features. In order to handle these sections of the coastline and to include them in the referencing framework, a notion of a 'coastal diffused node' has been utilised. Groups of adjacent, non-contracted coastal outlet nodes lying between a pair of coastal contracted nodes (as defined by aforementioned business logic) have been grouped together to represent a diffused flow to the sea. Each node in a given group has been assigned with a common identifier (DiffusGrp) and then the node closest to the spatial mean centre of the group was flagged (DiffuseKey) to become the diffused node representing flow to the sea within the referencing framework. This representative node is also contracted (ConLevel=2) to assign the logical diffused outflow for the stretch of the coastline a persistent identifier.

Diffused delta node

While the diffused coastal node type's focus is on representing near-direct discharge from short, poorly defined channels into the sea, the diffused delta node's purpose is to cater for distributed outflows originating from more well-defined sections of the stream network. When a stream travels through a flat area of land, the drainage network of the stream often spreads out to form a divergent flow pattern. In cases where this happens, and the flow continues out into the sea, this can be called a coastal delta system. Such systems can be very complex and their structure is often volatile and affected by such things as tidal flows and sediment build-up.

The business logic used in the Geofabric to determine whether a divergent flow pattern draining to the coast was to be attributed and treated as diffused delta node (as opposed to forming part of a diffused coastal node) involved the examination of adjacent nodes located along the coast. Wherever two or more coastal contracted nodes had flows that originated from topologically connected parts of the network (SH_Network), a diffused delta group was created. All coastal outlet nodes (contracted or non-contracted) in a diffused delta group have been attributed with a common identifier (DiffusGrp). Out of the subset of these nodes that were contracted (by aforementioned business logic), the node with the highest total upstream drainage area (DArea) was chosen and flagged (DiffuseKey) to become the representative diffused contracted node for the referencing framework.

Diffused inter-catchment node

In order to achieve the goal of producing a stable and dendritic referencing framework, while at the same time representing key reporting boundaries, it was necessary to introduce a third type of diffused node. This third type, the inter-catchment diffused node, represents flow between two or more inland areas where a group of nodes are to be treated as a single diffused outlet to a logical catchment.

As with the other diffused node types, all of the nodes (contracted or non-contracted) in a diffused inter-catchment group have been attributed with a common identifier (DiffusGrp). The key node (DiffuseKey), which is used to represent the diffused flow in the referencing framework, was chosen manually via expert input and wherever possible an existing contracted node has been used. However, for a few cases, where a suitable contracted node was not available at the inter-catchment outlet, the most prominent node was chosen and subsequently contracted manually (ConLevel=2).

Note: Currently (v2.1), the nodes within a given inter-catchment diffused node group can also contain a special type of 'blocker node'. These nodes, which are often located separate to the main cluster of diffused inter-catchment outlet nodes, were included in the group for use during catchment processing. These nodes serve the role of blocking or allowing a stream trace during contracted catchment construction, depending upon the origin of the trace. In the next minor update (v2.1.1) these nodes will be flagged separately via the use of additional attribution.

Network simplification

In the version 2.0 release of the Geofabric product suite, the HR_Catchments product comprised of a full node-link network produced by the application of the aforementioned processes. At the release of version 2.1, the full node-link network has been simplified to form a 'Catchment Hierarchy' referencing framework containing only a subset of the full set of contracted nodes. Through the combined application of a number of network simplification rules and the enabling concepts that have been outlined above, it was possible to reduce the complexity of the stream network (and its associated catchments) to form a simplified version of the node-link network. This network, which uses special logic to encapsulate divergent flow patterns, forms a fully dendritic network of stable contracted features terminating either at the coast or at known inland sinks. The largely automated process of simplifying the full node-link network of contracted features is described in brief below:

1. **Collapse & remove anabranches**

   The first step in the simplification of the node-link network involved the removal of anabranches from the node-link network. Anabranches in a stream network are sections of the network where flow diverts from the main channel and then later rejoins the same main channel at a point further downstream.

   Wherever an anabranched stream section in the SH_Network product contained one or more contracted nodes, this would result in the same anabranch being present in the full node-link network. Since, by definition, a fully dendritic network cannot include divergent nodes (flow splits), the corresponding sections of the node-link network containing anabranches needed to be simplified to a degree where the network became fully dendritic. The first step towards this goal began with the extraction of the node-link network's set of divergent nodes. Once identified, these divergent nodes were processed one at a time with each node's outflows being tested to look for pairings of pathways that ultimately lead to a common coastal terminus or inland sink. Any pairings of outward flowing links that returned a positive result for this test were then processed further to collapse down the node-link network a single link level, by redirecting each of the paired links' 'to-nodes' to instead point to their topological successor(s) (i.e. the next downstream node). Once redirected, any other inflows to these now by-passed to-nodes were also redirected to the same topological successors, before removing the nodes altogether from the node-link network. This process continued until all of the divergent nodes from the original list were processed. After one full pass was complete, the fresh set of divergent nodes was then extracted and compared against the set from the previous pass. If the two sets of nodes were different, then the same test-and-collapse process was repeated, with this loop continuing until the set of divergent nodes became static.

   At this point, the resulting node-link network was now simplified to a level where it was free from anabranched structures. However, the network was not necessarily dendritic, since it still had the potential to contain divergent nodes with distributed outflows terminating at the coast (coastal delta systems) or at inland sinks 'dangling' off of the main channel.

2. **Identify and attribute delta diffused node groups**

   The next step in the simplification process was the identification of coastal deltas. To this end, coastal terminus nodes were processed one at a time with the first step being an upstream trace in the underlying stream network (SH_Network) in order to locate any divergent nodes lying upstream of the node. For each upstream divergent node found, a downstream trace was then performed, and where the trace finished at the coast the corresponding terminus node(s) were recorded. Each of the coastal termini found in this initial downstream trace(s) were then checked in the same fashion (trace upstream & downstream from divergent nodes) in order to uncover any further coastal termini belonging to the same delta system. Once this iterative process had been run to completion, the resulting set of coastal termini was then examined for contracted nodes. If at least one of the nodes in the resulting set was found to be a contracted feature, then a logical diffused delta system was created, with each terminus node in the group being assigned a common identifier (DiffusGrp). As a final step, the contracted terminus node in the group with the largest total upstream drainage area (DArea) was then chosen and flagged (DiffuseKey) as the representative node for the diffused coastal delta system. These flagged nodes would be important in a later step when simplifying the deltas to be represented by a single link (see below).

3. **Identify and attribute coastal diffused node groups**
Once all of the coastal deltas were identified and flagged, the next step in the simplification process was to identify the coastal diffused node groups.
The starting point for the delineation of these groups was to generate a spatial buffer for the entire coastline. Once created, this buffer was then sub-divided by splitting the buffer at each linear stream feature ending with a contracted terminus node at the coast (SH_Network). Next, the resulting set of adjacent polygons was spatially intersected with the complete set of non-contracted coastal terminus nodes (taken from the underlying stream network), and all nodes falling within the same polygon were assigned with a common diffused group identifier (DiffusGrp). Finally, each node closest to the spatial mean centre of the diffused groups were flagged (DiffuseKey) and contracted to become the diffused nodes representing flow to the sea for their particular area of coastline.
Note: Nodes that had previously been assigned as belonging to a delta diffused node group (contracted and non-contracted) were treated as contracted nodes in this process i.e. they were used during the sub-division of the coastal buffer.

4. **Simplify diffused coastal deltas to single diffused outlets**
Since coastal deltas by definition are non-dendritic, the representations for these systems in the node-link network needed to be simplified to a single representative outlet. Once again, the process used to achieve this involved traversing the full stream network (SH_Network).
For each node in a given delta diffused node group (identified in an earlier step), an upstream trace was performed in order to extract a list of all of the upstream divergent nodes for the delta. This total list of nodes was then reduced to consist of only those nodes that had no further divergent nodes upstream i.e. the set of nodes representing the upstream extent of the coastal delta. Next, by using these remaining divergent nodes as starting points for traces downstream to the coast, it was possible to identify the nodes (and links) that needed to be simplified out (removed) from the node-link network. The final step in the process was to create links between each of the divergent nodes at the upper boundary of the delta and the representative delta outlet node at the coast (DiffuseKey). Any inflows to nodes that were removed from the node-link network were also redirected to the

representative outlet node for the diffused coastal delta system.

5. **Remove all ‘headwater-bypass’ nodes**

With all anabranches having been removed from the node-link network in an earlier step, it goes without saying that it is not possible to bypass topologically connected nodes whilst traversing the network. For example, if nodes A, B and C are connected in sequence (A->B->C), it is not possible to traverse from node A to node C without first passing through node B. This holds true whether traversing the node via the links in the node-link network or by the linear features connecting the same nodes within the underlying stream network (SH_Network).

However, in headwater regions of the stream network, although this rule still holds, it may well be possible to traverse around a node-link node in the underlying stream network and end up in an area of the network that is upstream of the node that was avoided (bypassed). While this has no significance when tracing the node-link network itself, it does however have consequences for the process used during the creation of the contracted catchments (see later). The method used for the creation of a set of non-overlapping ‘contracted’ catchments, which correspond to the areas drained by individual or pairs of contracted nodes, relies upon a foreign key (DrainID) to primary key (HydroID) relationship between the features in the AHGFNetworkStream and AHGFCatchment feature classes. Business logic used to create contracted catchment features, applies a first-come-first served principle when building catchments in order to avoid catchment overlaps. Therefore, if contracted nodes can be bypassed to any given degree in an upstream trace, the resulting contracted catchments may be significantly different depending upon the order in which the catchments are built. Using the same example as before, if a headwater area upstream of node B in the underlying stream network (but not connected to A) can be reached from node C via a network trace that doesn’t first pass through node B, then the catchment for node C will be different depending on whether it is processed before or after the catchment for node B.

Hence, in order to remove this unwanted sensitivity to process order, the next step in the simplification process was to check for occurrences of these ‘headwater bypass nodes’ in the stream network. Any nodes in the node-link network that were found to fall in this category were simply removed from the node-link network, with all of their inflowing links redirected to instead point to the next node downstream in the node-link network.

6. **Reassign and remove ‘dangling sinks’**

While the earlier steps of anabranch and coastal delta simplification greatly reduced the number of divergent flows contained within the node-link network, a number of these divergent network nodes still remained. Wherever a portion of stream flow diverged (split) from the main channel without later rejoining further downstream (anabranch) or continuing all of the way to the coast (delta), a branch that terminated at some form of inland sink would be evident ‘dangling’ off of the main channel.

In order to achieve the goal of a dendritic node-link network, these ‘dangling sinks’ would also need to be simplified to remove them from the network.

The first step in the simplification process involved determining which outflow from the divergent node represented the main channel (i.e. the flow path out of the divergent node that was to remain after simplification). In order to determine this, a series of depth-first stream traces, originating from the divergent node, were performed and the terminus for each path was recorded. For cases where one of the paths traced to a coastal terminus node, this path was immediately designated as the main channel. For instances where all of the paths from the divergent node terminated at inland sinks, a different method for determining the main channel was employed. In such cases, the 'main channel' was designated as the path with the largest total upstream drainage area at its terminus node.

Having decided upon a main channel, the next step was to remove all other paths (to inland sinks) flowing out of the divergent node from the node-link network. This involved removing all nodes and links on each path, except for the terminus (dangling sink) nodes themselves, which were necessary to preserve in order to allow the catchment area for the dangling sink to be later merged together with the catchment for the main channel. The remaining terminus nodes were all assigned (via MergedSink attribute) with the persistent identifier of the first node along the main channel path downstream of the divergent node.

Since the removal of divergent nodes was an iterative process and certain areas of the node-link network contained recursive dangling node paths (i.e. dangling node paths which themselves contained divergent dangling node paths and so on…), this also needed to be handled. Therefore, each time a path to a dangling sink was flagged for removal a check was performed to see whether the dangling node already had pointers from other processed dangling nodes pointing to it. Any nodes that were found to do so were updated (MergedSink adjusted) to instead point to the alternate node for the divergent node currently being processed.

After performing this process for all of the divergent nodes, the node-link network was then entirely free from divergent nodes and by definition could now be considered to be dendritic.

7. **Add any missing links to coastal/delta diffused nodes**

   The main input into the process used for contracting nodes, the source stream network data, contained inconsistencies around its attribution which sometimes affected the resulting node-link network created from the contracted nodes. These artefacts lead to instances of 'false termini', whereby the node-link network failed to continue all the way to a coastal terminus, even though this was apparent from evidence contained in the underlying geometry of the stream network(s). In an attempt to rectify this issue (as best as possible), an additional process was utilised to force any such false termini to connect through to the coast. The process by which this was achieved was by tracing the underlying stream network (SH_Network) upstream from each diffused coastal node or delta node looking for any occurrences of these false terminus nodes. Any false termini, encountered on an upstream trace from a given diffused node, were simply connected through to the diffused coastal outlet by the addition of a link in the simplified node-link network.

Contracted Catchments

Upon the completion of the network simplification steps outlined above, the node-link network was then ready to be used as the starting point for the construction of a hierarchy of AHGFContractedCatchment features (HR_Catchments). The catchments would be delineated from the nodes contained within the simplified network, with each node corresponding to a catchment outlet or diffused outlet key and a starting point for the construction of a single catchment or group of sub-catchments.

The catchment creation process began by tracing upstream from an outlet node or an outlet node group (diffused group) in the underlying stream network (SH_Network), with the trace continuing until either another simplified contracted node was reached or the stream network was exhausted (i.e. stream head reached). If fewer than two simplified contracted nodes were encountered during this initial trace, then the traced set of streams was ready to be used to build a single catchment for the outlet node. However, if two or more upstream simplified contracted nodes were present in the results of the trace, then it was necessary to further process the catchment into sub-catchment areas, with a sub-catchment defined for each of the links flowing into the outlet node. The process used to determine sub-catchment areas involved taking the stream features of the initial trace and then tracing downstream from each of the upstream terminal contracted nodes. These downstream traces were performed in descending order based on the contracted nodes' total upstream drainage area, so that the sub-catchments for the most major streams were always processed first. This order of processing was significant because the resulting stream traces were not allowed to contain overlaps i.e. once a stream segment was traced the first time it was allocated and acted as a blocker for subsequent traces. The next step of the sub-catchment creation process involved taking the sets of stream features from the downstream traces (in the same order) and then tracing upstream (in the full set of features from the initial trace) from each quasi-terminus stream contained within each set. Again, overlapping of traces was not allowed, with sets of streams from other downstream traces acting as trace blockers for the current set being processed.

The end result of tracing the stream network for each catchment outlet node was either a single set of stream features or a collection of stream feature sets delineating a catchment or group of sub-catchments respectively. In order to turn these traced sets of linear features into catchments, the relationship between AHGFNetworkStream features (DrainID) and AHGFCatchment features (HydroID) was utilised to return set(s) of corresponding polygonal features.

Before the final catchments could be created, one further preparatory step was required. This penultimate step involved using the 9" DEM to merge any 'NoFlow' catchments (AHGFCatchment features without related AHGFNetworkStream features) with nearby 'Flow' catchment features (catchments with related streams). Once the process of merging was complete and all of the NoFlow catchments were merged into catchments with a related stream segment, the final step was simply to dissolve the sets of stream segment-level catchments into a catchment or collection of sub-catchments. These contracted catchment(s), representing the area draining to a contracted outlet node, would form part of the stable 'Catchment Hierarchy' referencing framework.

River Region Enforcement

While the application of the business logic and processes described above achieved the goal of creating a simplified dendritic node-link network and a set of hierarchical contracted catchments, further work was still required. As a result of encapsulating complexities of the network in some areas of low relief, key areas, such as a large part of the Murray-Darling basin, had been simplified to such an extent that single catchments spanned across large areas. Since the minimum use case for the HR_Catchments product was to produce catchments that could be aggregated into the Bureau of Meteorology's set of River Regions, more business logic and further processing were required. It was here that the third diffused node type, the diffused inter-catchment node (see above), was applied. Using the simplified network and the set of contracted catchments as a visual reference, expert input was used to identify and attribute groups of nodes at the outlets of River Regions to form diffused inter-catchment nodes. Similarly, expert input was also used in order to force the splitting of coastal diffused node groups at River Region boundaries along the coast. Once these changes were identified and attributed, the network simplification and catchment generation processes were simply re-run using the extra information as additional inputs. The end result this second time round was a simplified dendritic node-link network and a set of hierarchical contracted catchments enforced to align with the River region boundaries.

Contact

For more information on the Geofabric please visit our website:
http://www.bom.gov.au/water/geofabric,
Or, contact us via email: ahgf@bom.gov.au

References

Atkinson, A., Dornblut, I., Smith, D. (2012). An international standard conceptual model for sharing references to hydrologic features. Journal of Hydrology (Elsevier), 424–425, (2012), 24-36.