Differences Between ArcGIS and QGIS Implementations of SCS Toolbox

This note documents all intentional and unavoidable differences between the original ArcGIS SCS Modules (Rusnák) and the QGIS port.

Module 1 – Centreline Extraction

Summary
The modules are conceptually equivalent—both generate centrelines from channel polygons in per-year and union/segmentation modes—but differ in algorithms, geometry cleaning methods, and attribute handling.

---

1. Centerline Algorithm (Major Difference)

ArcGIS

- Implements a custom Thiessen-based medial axis:
  - Densifies polygon boundary to regularise vertices.
  - Builds Thiessen polygons from boundary points.
  - Extracts Thiessen edges forming the medial axis.
  - Angle filtering: removes edges nearly parallel to channel banks (DIFF 50–130°).
  - ExtendLine forces the centreline to reach polygon boundaries.

QGIS

- Primary method: GRASS v.voronoi.skeleton (true medial axis transform).
- User-controlled parameters: smoothness, thin, snap, minarea, region.
- Fallback: QGIS Voronoi-clip workflow (no angle filtering).
- Does not reproduce ArcGIS angle filtering or ExtendLine.
- The resulting centreline is therefore geometrically similar, but not identical, to ArcGIS output.

Implication
Centrelines in QGIS and ArcGIS typically follow the same channel axis but differ in fine-scale detail because the algorithms are not exactly the same.

---

2. Hollow / Island Handling

ArcGIS

- Removes interior rings by manually reconstructing polygons using cursor logic.

- Produces POL_YYYY.shp and union_channel.shp (no hollows).

QGIS

- Removes hollows using:
  - fixgeometries
  - deleteholes
  - optional healing via buffer(0)
- Produces:
  - union_channel.gpkg (holes preserved)
  - union_channel_noholes.gpkg (holes removed)
  - union_channel_byYear_noholes.gpkg (optional)

Implication
Channel envelopes are conceptually identical, but QGIS may smooth tiny ring defects or heal slivers that ArcGIS preserves.

---

3. Year / Date Handling

ArcGIS

- Expects a valid DATE field.
- Names outputs as centro_YYYYMMDD.shp.
- Stores full date in field cnt.

QGIS

- Accepts integer, text, or date year fields.
- Extracts years using regex (18xx–20xx).
- Standardises dates:
  - Names outputs as centro_YYYY0101.gpkg.
  - Stores cnt = YYYY-01-01.
- Adds a dedicated integer Year field.

Implication
QGIS is more tolerant of inconsistent source year fields but does not reproduce the original full-date naming.

---

4. Segmentation / Union Mode

ArcGIS

- Unions all channel layers → union_channel.shp (no hollows).
- Extracts a segmentation centreline using the same Centro() procedure.
- Output: SegCenterline.shp.

QGIS

- Builds several union variants:
    - union_channel.gpkg (dissolve per-date, holes preserved).
    - union_channel_noholes.gpkg (cleaned envelope).
- Selects the best polygon for skeletonisation (priority: noholes → healed → fixed → boundary polygon).
- Skeleton created using GRASS or fallback QGIS method.
- Output: SegCenterline.gpkg with metadata (Centerln, cnt, length).

Implication
QGIS provides more robust union handling but produces centrelines that may differ slightly from ArcGIS because of skeleton method differences.

---

5. Attribute Output Differences

ArcGIS

- Per-year centrelines:
    - cnt (DATE), geometry length.
- SegCenterline:
    - geometry length only.

QGIS

- Per-year centrelines:
    - Centerln = 'centerline'
    - cnt = standardised date
    - length
    - Year (int)
- SegCenterline:
    - Same schema as per-year centrelines.

Implication
QGIS outputs are more richly annotated and consistent with downstream Module 2 and Module 3.

---

6. Geometry Robustness and Safety

ArcGIS

- Uses %ScratchWorkspace% shapefiles.
- Assumes tools succeed; limited fallback logic.
- Uses manual row deletes and ring reconstruction.

QGIS

- Forces all GRASS inputs to SHP to avoid memory leaks.
- Robust fallbacks:
    - Multiple skeleton methods.
    - Multiple line-merge strategies.
    - Geometry repair using fixgeometries/buffer.
- Automatic cleanup of temporary directories.
- More defensive against invalid geometries and mixed CRS inputs.

Implication
The QGIS implementation is substantially more robust to corrupt geometries and mixed CRS inputs.

---

7. Output Formats

ArcGIS

- All outputs are shapefiles (.shp).

QGIS

- Processing intermediates: SHP (for GRASS compatibility).
- Final outputs: GPKG (GeoPackage):
    - centro_YYYY0101.gpkg.
    - SegCenterline.gpkg.
    - Optional union layers.

Implication
QGIS provides modern, multi-layer GPKG outputs suitable for Module 3.

Module 1 – Overall Summary

The QGIS version faithfully replicates the workflow logic of the ArcGIS module—per-year extraction, union-mode segmentation, hollow removal, centreline extraction, and attribute generation—but differs in:

- underlying centreline algorithm,
- geometry repair methods,
- year/date parsing,
- output formats, and
- robustness mechanisms.

These differences are necessary due to the QGIS/GRASS Processing framework and ensure the tool works reliably on diverse datasets, but they inevitably produce centrelines that match the ArcGIS outputs in function, not in exact geometry.

Module 2 – Segmentation

Summary
The modules are conceptually equivalent—both generate equal-length longitudinal segments from a channel centreline and union polygon—but differ in how they construct or source unions, discover and simplify the centreline, and build and clip Thiessen/Voronoi polygons.

---

1. High-level equivalence

Shared purpose
Both implementations:

- Generate longitudinal channel segments at a fixed spacing (interval) along a segmentation centreline.
- Use a Thiessen / Voronoi polygon construction around midpoints on the centreline.
- Clip those polygons to the channel union polygon to form the final Segments dataset.
- Attach two key attributes to segments:
  - ID_SEQ – reverse sequence number along the centreline.
  - Distance – nominal segment length (in metres).

Functionally, both create an equal-length segmentation of the channel corridor suitable for linking to EA rates and other statistics (Module 3).

---

2. Inputs & outputs

ArcGIS version

Inputs

- output_folder – workspace for SHP outputs.
- inputLayer – one or more channel polygon layers (multi-date allowed).

- inputCenterline – the segmentation centreline (typically SegCenterline.shp from Modul1).
- field_year – date field in polygons.
- interval – segment spacing (distance).
- simplification – optional centreline simplification distance.
- deleteTF – flag for deleting intermediate SHPs.

Outputs

- SHP:
  - union_channel.shp – union of all channel polygons, without hollows.
  - Segments_<interval>m.shp – final segments with Distance and ID_SEQ.

QGIS v10 (use-existing-union)

Inputs

- CENTERLINE_DIR – folder containing Module 1 outputs (e.g. SegCenterline.gpkg, union_channel_noholes.gpkg).
- INTERVAL – segment spacing (m).
- SIMPLIFICATION – optional smoothing offset (m) for centreline.
- OUTPUT_FOLDER – folder for the final GPKG.

Outputs

- GPKG:
  - Segments_<INTERVAL>m.gpkg
    - sublayer Segments with only Distance and ID_SEQ fields.

The QGIS Module 2 **does not rebuild** a union; it assumes Module 1 has already produced a hole-free union and will fail if it cannot be found.

---

3. Union / channel polygon handling

ArcGIS

- Rebuilds the union from scratch every time:
  1. Copies each input polygon layer to CH_YYYYMMDD.shp.
  2. Builds union_pol via Union.
  3. Dissolves into union_pol2 on a dummy DISS field.
  4. Manually removes hollows by reconstructing polygons from outer rings.

5. Writes union_channel.shp with projection SR.

- union_channel.shp is both the clipping mask and the persisted union for segmentation.

QGIS

- Reuses a union created in Module 1:
  - Prefers a polygon layer in the project whose name includes both "union" and "nohole".
  - If not found, searches CENTERLINE_DIR for likely candidates:
    - union_channel_noholes.*
    - union_channel_byYear_noholes.*
    - union_channel.gpkg sublayers.
    - Any polygon layer named like *union*noholes*.

Implication

- ArcGIS: Module 2 is self-contained for union creation.
- QGIS: Module 2 is tightly coupled to Module 1 outputs and expects consistent naming.

---

4. Centreline handling & simplification

ArcGIS

- Centreline is explicitly provided (inputCenterline).
- Optional simplification:
  - If simplification == 0: copy as is.
  - Else: copy to centro_simple_<simplification>m.shp, then Integrate with that tolerance.
- CRS consistency is assumed from the ArcGIS environment.

QGIS

- Centreline is **discovered**, not directly passed:
  1. Attempts to find a line layer named like "SegCenterline" in the project.
  2. If not found, searches CENTERLINE_DIR for line layers (e.g. SegCenterline*.gpkg, centro_*.gpkg, *centerline*).
  3. Ranks candidates by name and spatial overlap with the union.
- Simplification / smoothing:
  - SIMPLIFICATION interpreted in metres.

- Attempts multiple algorithms (native:smoothgeometry, qgis:smoothgeometry, qgis:chaikinsmoothing) and falls back gracefully if smoothing fails.
- Enforces projected CRS in metres and line-only geometry before proceeding.

Implication
QGIS adds robustness and auto-discovery, but demands valid projected CRS and consistent naming of outputs from Module 1.

---

5. Midpoints / segmentation logic

ArcGIS

- Generates points at exact multiples of interval along the centreline.
- Splits the centreline at these points, then builds segments and midpoints from split lines.
- Sets Distance = interval and ID_SEQ as a reversed sequence derived from object IDs.

QGIS

- Does not physically split the centreline.
- Uses native:pointsalonglines with:
  - DISTANCE = interval,
  - START_OFFSET = interval/2,
  - END_OFFSET = 0
    to generate **true midpoints** of segment-length interval.
- Adds a sequential SEQ field, then sets ID_SEQ = max(SEQ) – SEQ.
- Sets Distance = int(round(interval)) on the points.

Net effect
Logical segmentation is equivalent (evenly spaced segment centres and reversed ID sequence), but QGIS uses a virtual segmentation via midpoints only.

---

6. Thiessen / Voronoi polygons & clipping

ArcGIS

- Sets analysis extent from channel polygons.
- Uses CreateThiessenPolygons on segment midpoints, then clips Thiessen polygons with union_channel.

QGIS

- Uses qgis:voronoipolygons on midpoints plus **guard points** around the union extent to stabilise outer Voronoi cells.

- Heals the union via buffer(0) before clipping.
- Clips Voronoi cells to the healed union to form the raw segments.
- Transfers attributes from midpoints to polygons using joinattributesbylocation with multiple fallbacks (CONTAINS, INTERSECTS, WITHIN, buffered points) to ensure robust matching.

Implication
Both implement Thiessen/Voronoi-style segmentation, but QGIS is more defensive (guard points, healing, join fallbacks) and may produce slightly different segment boundaries compared to ArcGIS.

---

7. CRS, units and robustness

ArcGIS

- Assumes projected, consistent inputs; does not explicitly enforce CRS.
- Uses environment settings for extent and default error handling.

QGIS

- Explicitly checks that the centreline CRS is projected in metres.
- Reprojects datasets where needed.
- Uses feature-count checks and detailed error messages when intermediate steps fail.
- Writes final output to GPKG; intermediates are TEMPORARY_OUTPUT.

Module 2 – Overall Summary

When given the same union polygon, centreline and interval in a projected CRS, the ArcGIS and QGIS implementations are designed to produce functionally equivalent segmentations (similar segment geometry, matching Distance and ID_SEQ semantics). Differences in Voronoi construction, union healing, CRS checking and file formats mean that segment outlines will not be exactly identical, but they can be used interchangeably in downstream EA and statistics workflows.

Module 3 – EA (Erosion–Accretion) Classification

Summary
The modules are conceptually equivalent—both classify erosion, deposition, island dynamics and migration rates between dated channel states—but differ in year handling, polygon cleaning, construction of side masks, the granularity of classes, and the richness of rate outputs and robustness checks.

---

1. Year Handling

- **ArcGIS** assumes strict year matching between polygons and centrelines and relies on DATE fields being properly configured.
- **QGIS** automatically extracts years from integer, string, or date fields, ignores unmatched years, and is more tolerant of heterogeneous attribute formats.

---

2. Polygon Cleaning & Islands

- Both workflows generate envelope-without-hollows polygons and derive island polygons.
- **ArcGIS** uses cursors and ring-trimming logic to reconstruct outer rings.
- **QGIS** uses Processing tools (dissolve, deleteholes, difference, fixgeometries) to achieve conceptually identical results, with additional geometry repair.

---

3. Orientation Masks

- Both implement side masks (LEFT/RIGHT) based on buffering the centreline and partitioning space.
- **ArcGIS** uses FeatureToPolygon and SpatialJoin.
- **QGIS** mirrors the logic with singlesidedbuffer, polygonize, and joinattributesbylocation, plus geometry fixing and explicit handling of LEFT/RIGHT ties using a BOTH category.

---

4. EA Classification

- Core erosion/deposition rules match between ArcGIS and QGIS.
- **QGIS** refines categories, especially for “stable” areas:
    - Splits ArcGIS’ broad “hollow”/“stable” into:
        - stable_channel,
        - stable_floodplain,
        - inactive_floodplain.
    - Preserves island classes (island_erosion, island_deposition).

---

5. Direction & Migration

- Both use side masks to assign LEFT/RIGHT direction and define migration sign.
- **QGIS**:
    - Handles LEFT+RIGHT ties as BOTH.
    - Treats stable_channel as zero-migration.

- Applies consistent sign rules to islands and stable floodplain units.

---

6. Rates

- **ArcGIS** computes erosion/deposition rates broadly consistent with area-over-time metrics.
- **QGIS** reproduces these core rate calculations and adds:
  - EA_disp_m – displacement distance.
  - EA_rate_m_py – migration rate (m per year).

---

7. Outputs

- **ArcGIS** produces multiple shapefiles per period and class.
- **QGIS** outputs to a single GeoPackage with standard layer naming per period:
  - EA_processes_*, EAsegments_*, EA_rate_*.

---

8. Geometry Robustness

- **ArcGIS** largely trusts input geometries and uses scratch shapefiles.
- **QGIS** chains multiple safety steps (fixgeometries, makevalid, buffering, grid snapping, 2D coercion) to cope with complex or corrupted datasets.

Module 3 – Overall Summary

The QGIS EA module is designed to be drop-in compatible with the conceptual logic of the ArcGIS Modul3—same erosion/deposition meaning, similar island behaviour, comparable directions and rates—while adding finer class distinctions, richer rate metrics, and substantially more robust geometry handling, all packaged into a single GeoPackage for each period or period-pair.

Module 4 / 4a – Floodplain Age, Assemblage, HACH & CHM

Summary
The modules are conceptually equivalent—both derive surfaces describing floodplain age (or assemblage), height above channel and canopy height, and link them to longitudinal segments—but differ in how floodplain age is represented (channel-based FAM vs EA-based assemblage), in interpolation methods for HACH, in file formats, and in how statistics are attached to segments.

---

1. Processing scope and modes

ArcGIS Modul4_FloodplainStat

- Builds a **Floodplain Age Map (FAM)** directly from channel polygons and a date field.
- Then derives:
    - Height Above Channel (HACH) via DEM + flow path.
    - Canopy Height Model (CHM) via DSM – DEM.
    - Segment-based statistics combining FAM, HACH and CHM.
- Uses k modes driven by DEM/DSM/flow availability (k = 1, 2, 3).

QGIS Module 4a (v2.3)

- Builds an EA-based **floodplain assemblage** from multiple EA polygon layers (Module 3 outputs) using a chronological "newest-wins" overwrite chain.
- The composite output is Depositional_Composition with:
    - EA_Class_final – final EA class at each location.
    - Period_final – final period (yyyy_mm_dd).
    - Year – integer year.
- Optionally adds:
    - HACH (DED_m5.tif) via DEM + flow.
    - CHM (veget_CHM_m5.tif) via DEM + DSM + flow.
    - Segment-based statistics if SEGMENTS is provided.

Key conceptual change
ArcGIS: floodplain age is stored as a dedicated **FAM** field derived from channels.
QGIS 4a: age-like information is carried via Period_final / Year in an EA-based assemblage layer, not as a standalone FAM shapefile.

---

2. Inputs, outputs, and formats

ArcGIS Modul4_FloodplainStat

- Inputs: dated channel polygons, year/date field, DEM, DSM, flow path, segments.
- Outputs (SHP + GeoTIFF):
    - fam_layer.shp.
    - DED.tif.
    - veget_CHM.tif.
    - M4stattistics_all.shp, M4stattistics_hach.shp, M4stattistics_FAM.shp.

QGIS Module 4a v2.3

- Inputs: up to 15 EA polygon layers, class and period fields, optional DEM, DSM, flow, segments.
- Outputs (GPKG + GeoTIFF):
    - <output>.gpkg with layer Depositional_Composition.
    - DEM_clipped_m5.tif.
    - DED_m5.tif (if DEM + FLOW).
    - veget_CHM_m5.tif (if DEM + DSM + FLOW).
    - M5statistics_all.gpkg / M5statistics_hach.gpkg / M5statistics_FAMlike.gpkg (if segments provided).

QGIS standardises on GeoPackage for vector outputs, consistent with Modules 1–3 and 5.

---

3. Floodplain Age Map (FAM) vs EA Assemblage

ArcGIS FAM

- Constructs per-date channel copies (CH_YYYYMMDD.shp) with yYYYYMMDD fields.
- Unions them, then sets:
    - FAM = max(yYYYYMMDD) across all dates.
- Outputs fam_layer.shp as a single polygon layer with a FAM field.

QGIS 4a Assemblage

- Sorts EA layers oldest → newest using end date parsed from the period field.
- For each EA layer i, builds a minimalist schema:
    - EA_Class_i, Period_i (period text).
- Applies a SHP-based “newest-wins” overwrite chain with optional stable_floodplain carry-forward.
- Coalesces to:
    - EA_Class_final, Period_final, Year.
- Does not output a separate fam_layer; instead, age is inferred from Period_final / Year in Depositional_Composition.

Implication
QGIS 4a trades a simple FAM field for a richer assemblage representation that combines time and EA process class.

---

4. HACH (Height Above Channel)

ArcGIS

- Uses DEM + flow:
    - Densifies flow.
    - Samples DEM along the densified line.
    - Uses TopoToRaster to build a smooth trend (outTrend).
    - Computes DED = DEM – outTrend.

QGIS 4a

- Clips the DEM to a buffered AOI around the assemblage (DEM_clipped_m5).
- Generates points along flow at ≈5× pixel spacing.
- Samples DEM_clipped_m5 at these points (dem_1).
- Builds a trend surface via IDW:
    - Prefer gdal_grid CLI (invdistnn) with explicit extent and resolution.
    - Fallback: provider-based IDW + warpreproject.
- Computes DED_m5 = DEM_clipped_m5 – trend_idw_m5.

Implication
Both produce a height-above-channel surface. QGIS replaces TopoToRaster with IDW plus careful grid control, so values differ numerically but preserve the broad pattern.

---

5. CHM (Canopy Height Model)

ArcGIS

- CHM = DSM – DEM, then SetNull(chm, chm, "VALUE <= 0").
- Outputs veget_CHM.tif with NoData in non-vegetated areas.

QGIS 4a

- CHM_raw = DSM – DEM_clipped_m5.
- Uses a raster calculator to set non-positive values to -9999 and uses -9999 as nodata.
- Outputs veget_CHM_m5.tif.

Conceptually, both represent positive canopy height above terrain; the difference is in nodata encoding and DEM clips.

---

6. Segment-based statistics

ArcGIS

- Intersects fam_layer with segments, converts to singleparts.
- Runs ZonalStatisticsAsTable for HACH and CHM.

- Renames stats to e_* (HACH) and v_* (CHM) and joins them back into the polygon layer.
- Saves as M4stattistics_* shapefiles.

QGIS 4a

- Intersects Depositional_Composition with SEGMENTS via native:intersection, converts to singleparts.
- Adds zonal statistics directly with qgis:zonalstatistics:
    - DED_m5 → e_* fields.
    - veget_CHM_m5 → v_* fields.
- Writes a single GPKG with the appropriate name (M5statistics_all, _hach, _FAMlike).

Implication
The same statistics (min/max/range/mean/std/sum) are produced, but tied to the EA assemblage rather than a pure FAM surface.

---

7. Geometry robustness and safety

ArcGIS

- Uses %ScratchWorkspace% for temporary rasters and shapefiles.
- Explicitly deletes intermediate datasets but does not add many defensive checks.

QGIS 4a

- Uses _safe_polygon_difference and _freeze_to_shp to:
    - Filter geometries to polygons only.
    - Fix invalid geometries.
    - Strip PK-like FID fields.
- Relies heavily on TEMPORARY_OUTPUT, GPKG writing, and explicit clean-up of existing files in the output folder to avoid Windows/GPKG locking issues.

Module 4 / 4a – Overall Summary

ArcGIS Modul4 and QGIS Module 4a both aim to provide a vertical and temporal characterisation of the floodplain and to attach those metrics to longitudinal segments. ArcGIS does so through a FAM + HACH + CHM workflow based on channel polygons, whereas QGIS 4a constructs a richer EA-based floodplain assemblage, then derives equivalent HACH/CHM surfaces and segment statistics on top. The two are conceptually consistent, but the QGIS implementation is more tightly integrated with Modules 3 and 5 and uses GeoPackage + IDW-based interpolation for greater robustness and interoperability.