Case study 2: the Ural River at Oral, Kazakhstan

What would a Geovatra Baseline have told the Dutch Risk Reduction Team before its mission, and how realistic are its rivers and catchments?

> **Draft for review, 2026-11-13.** Numbers come from `case_study_ural.json` (`manage.py case_study_ural`). The comparison rules were fixed before any score was computed. Mission facts are only those in public sources, listed at the end.

Findings

- **The streams are realistic.** Against HydroSHEDS HydroRIVERS (independent of our elevation processing), 84.0 % of HydroRIVERS' rivers lie within 500 m of a Baseline stream, and 71.0 % of Baseline's streams within 500 m of HydroRIVERS (both with at least 10 km² upstream).
- **Catchment areas: inconclusive.** Comparing upstream areas at points failed: HydroRIVERS' lines lie up to several hundred metres from Baseline's streams, so a point lands on a different river or on a hillslope. Under the rule fixed beforehand only 16.0 % of 119 points agree within 20 %; a rule revised after seeing that didn't help (25.7 % within 20 %, but the median ratio fell to 0.09). That says more about point matching than about the catchments; a fair test compares catchment outlines (e.g. HydroBASINS), which is the next step. No other rules were tried.
- **But the Ural's own catchment is out of reach.** HydroRIVERS gives the Ural 189,049 km² upstream of this area, mostly in Russia. A Baseline area is at most 5,000 km², so Baseline can say where the river floods here, but not how much water is coming. The report doesn't say this clearly today.
- **The rainfall module misleads here.** It calls Oct-Mar the "wet season" and Dec the wettest month; in this climate that precipitation falls as snow and floods come when it melts in spring.
- **Snowmelt isn't assessed, and Baseline doesn't say so.** The Ural's floods come with spring snowmelt; the Dutch team and Kazakh authorities named snowmelt runoff as the hardest part to anticipate (2024). Baseline's rainfall module (M4) counts rain only, and its data-gaps module (M7) has no snowmelt warning. This is a gap to fix.
- **The April 2024 flood couldn't be checked:** there is a single Sentinel-1 radar scene over the area for the whole flood season and none in summer to compare with, a gap in radar coverage of western Kazakhstan since Sentinel-1B failed in December 2021.
- **What Baseline adds for such a mission** is a common picture on day one, from open global data: flood-prone land, who lives on it, the rainfall regime, land cover and data gaps. What it can't replace is upstream hydrology, reservoirs and snow, the topics the mission spent most time on.

The mission (public sources only)

- **2022 scoping mission** of the Dutch Risk Reduction Team for the Netherlands Enterprise Agency (RVO), with a newly formed interdepartmental working group of the Kazakh government, after Kazakhstan "was again hit by flooding in 2022": information sessions and a field visit to the Ural River (Witteveen+Bos). Report: Bosch & Zuijderwijk (2022), DRR222KZ01, not public. Key recommendations, as summarised by RVO: data-driven decision-making tools, and an interdepartmental working group.
- **April 2024**: floods across Kazakhstan; the Dutch DRRS team was in the country 15-18 April during the emergency, then held three collaboration weeks (Astana, Aktobe) in August-November 2024 (RVO 2025).
- **What the 2024 report names as hard** (RVO 2025): estimating snowmelt runoff; reservoirs, some regulated and some not; ungauged and transboundary basins; satellite and radar data not yet used in the national hydrological models; the need for global data and remote sensing to fill gaps.

What a Baseline would have given before the mission

Run today on 4,987 km² around Oral (workflow baseline/1.0.0). The data are today's versions; in 2022 the elevation, water history (1984-2020), flood model inputs and rainfall to 2021 already existed, but OpenStreetMap, WorldPop 2025 and the JRC v2 flood maps (2024) were older or not yet published.

Module	What it says
M1 terrain	Elevation 18 to 158 m; 91.8 % of the area is flat (under 2°); 2,100 km of streams; 1,200 km² less than 2 m above the nearest stream.
Flood-prone zones	790 km² flood-prone (15.8 % of the land): zone A 240 km², zone B 550 km²; zone C 2,400 km². Confidence: high.
M2 water history	Water seen on 250 km² in 1984-2020; 16 km² permanent.
M4 rainfall	440 mm a year (1991-2020, CHIRPS); wettest six months Oct-Mar; wettest month Dec.
M5 land cover	Cropland 45.2 %; Grassland 48 %;
M6 exposure	26,000-78,000 people on flood-prone land, of 300,000-410,000 in the area; 1,100 km of main roads.
M7 gaps	M1 terrain: medium. M4 rainfall: high. M6 exposure: medium. M5 land cover: high. M2 water history: high. Flood evidence zones: medium.

The mission's topics against Baseline

Topic (from the public sources)	Baseline	How
Where the river floods, and how often water has been seen	**Covered**	Flood-prone zones (M1) and water history (M2), with a confidence level.
Who and what is exposed	**Covered**	People (two models), built-up land, roads, facilities per zone (M6), totals only.
Data for decisions; using global and satellite data	**Covered**	This is what Baseline is: open global and satellite data, with a methods log and a data package.

Rainfall regime and recent years	**Partly**	M4 gives the rainfall normal and recent years, but rain isn't what drives these floods.
Snowmelt runoff	**Not covered**	Not assessed, and not flagged as a gap: to fix.
Upstream and transboundary catchment	**Not covered**	Areas are at most 5,000 km²; the Ural's catchment above Oral is 189,049 km².
Reservoirs and their operation	**Not covered**	Only "dam failure is not assessed" is stated.
Forecasting and early warning	**Out of scope**	Baseline describes the situation before a flood; it doesn't forecast.

Rivers and catchments against HydroRIVERS

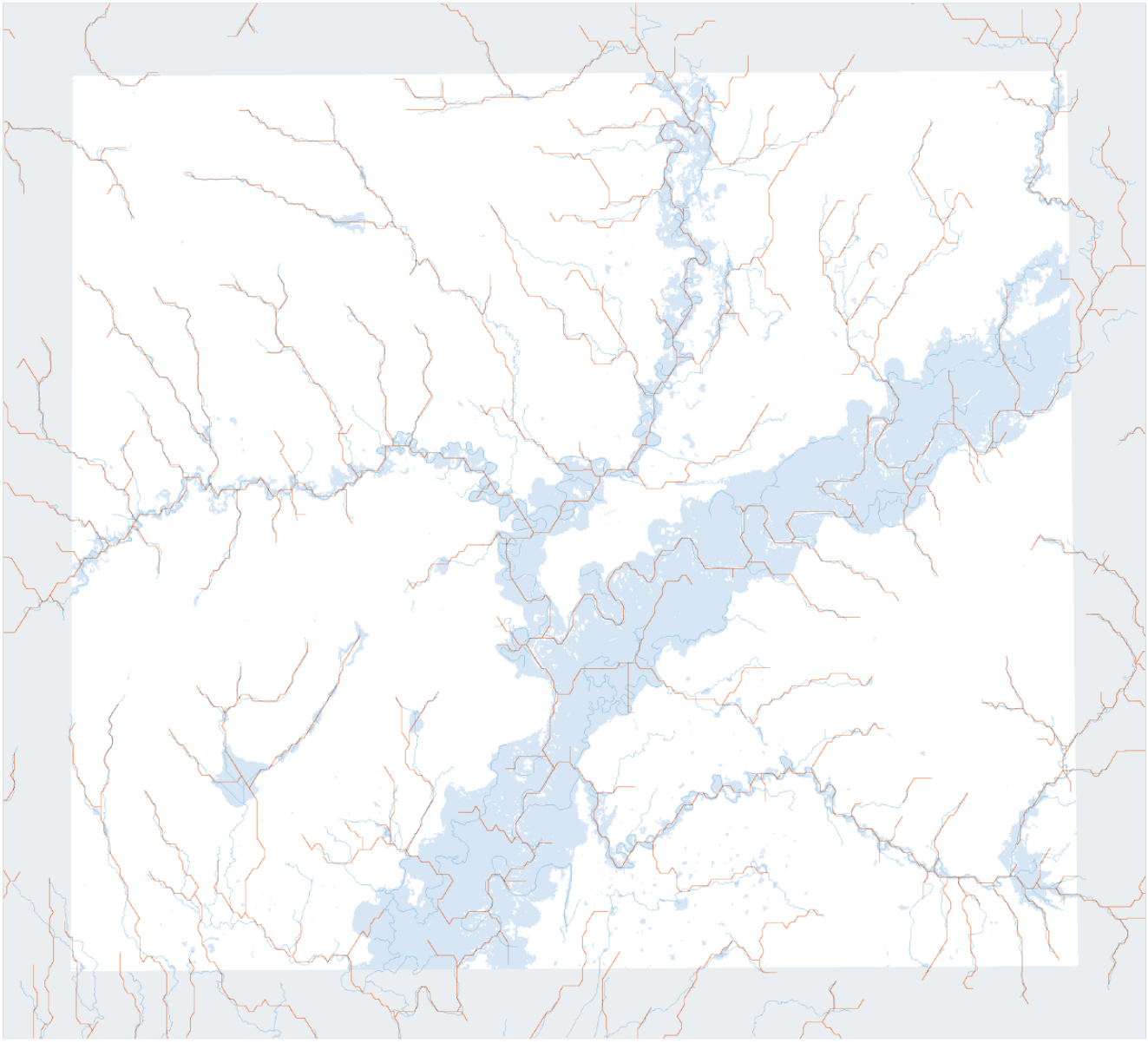
Reference: HydroSHEDS HydroRIVERS v1.0 (Lehner & Grill 2013), river lines with at least 10 km² upstream, derived from SRTM-based HydroSHEDS at 15 arc-seconds (about 300 by 460 m here), independently of Baseline's Copernicus DEM processing (GRASS r.watershed, 30 m).

Rivers with at least 10 km² upstream, inside the area	Baseline	HydroRIVERS
Length (km)	1,558	1,624
Share of the network within 250 m of the other's rivers	57.0 %	71.6 %
Share of the network within 500 m of the other's rivers	71.0 %	84.0 %

Baseline's length counts 30 m cells along the streams, so it is approximate (diagonal steps are counted short).

Upstream area, Baseline against HydroRIVERS (catchments wholly inside the grid)	Rule fixed before scoring: reach end, within 500 m	Revised rule: 300 m above the reach end, within 150 m
Points compared	119	167
Left out: catchment reaches beyond the grid	157	110
Median ratio, Baseline / HydroRIVERS	1.52	0.09
Within ±20 % / within ±50 %	16.0 % / 31.1 %	25.7 % / 31.1 %

Why the rule was revised after the first results: HydroRIVERS reach ends lie at confluences, so the largest accumulation within 500 m of an end is usually that of the river the reach flows into (e.g. a 10 km² stream matched a 33 km² river). The revised rule samples each reach 300 m upstream of its end, within 150 m. It didn't help: slightly more points agree within 20 %, but the median ratio fell to 0.09, because the lines are often more than 150 m apart and the search lands on hillslope cells. Both results are reported and the comparison is called inconclusive; no further rules were tried, to avoid choosing one that looks good. Next: compare catchment outlines (HydroBASINS sub-basins against basins drawn from Baseline's flow directions at the same outlets).



The April 2024 flood: no radar to check against

The plan was to check the zones against the April 2024 flood seen by Sentinel-1 radar. It isn't possible here: the Planetary Computer catalogue has one Sentinel-1 scene over this area for 5 April-20 May 2024 and none for the summer of 2024 (nor for spring 2023). Since Sentinel-1B failed in December 2021, western Kazakhstan has had very sparse radar coverage. This also matters for the product: radar flood maps of an event (the planned module M3) would not work here for 2022-2025.

Engine message: ValueError: no Sentinel-1 scenes for 2024-04-05/2024-05-20 or 2024-07-15/2024-09-15

Limits

- One area, one river. The mission's exact field sites aren't public; the area was drawn around Oral.
- The Baseline was run with today's data, not the data of 2022 (see above).
- HydroRIVERS is coarser than Baseline's streams and is itself modelled from elevation data, not mapped on the ground.
- The comparison with the mission is qualitative: the mission's report isn't public, so only what public sources say could be compared.

- Zone C (low-lying land to check) is very large here (2,400 km²), as expected on a flat steppe floodplain: on flat land, terrain alone says little.
- The April 2024 flood could not be checked (no radar scenes, see above).

What it means for the product

- Add a snowmelt warning to M7 for areas with cold winters, say that rain-only analysis understates spring floods there, and don't call snowy winter months the "wet season" in M4.
- State the river's upstream catchment (e.g. from HydroRIVERS) in the report, so a reader knows how much of the flood source lies outside the area.
- Test catchment outlines against HydroBASINS before claiming catchment accuracy.

Sources

Witteveen+Bos, "DRR mission to Kazakhstan" (witteveenbos.com/news/drr-mission-to-kazakhstan). Netherlands Enterprise Agency (2025), "DRRS: Collaboration on Flood Risk Management, Kazakhstan" (english.rvo.nl/files/file/2025-02/DRRS_Kazakhstan_FRMCollaboration2024_Final31012025.pdf). IHE Delft, "Dutch emergency team assists with Kazakhstan floods: an interview with Fredrik Huthoff" (un-ihe.org). Lehner, B. & Grill, G. (2013), Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems, Hydrological Processes 27, 2171-2186 (HydroSHEDS licence: free use with attribution). Run 48fa0c6a-12c5-46ae-9f9c-f8e4b473f7db, code 436d135.