Hydrogeology Digitizing Handbook

For Zimbabwe and Mozambique.

Kartoza Pty (Ltd.) and IGRAC

2022

Contents

1 IGRAC Digitizing Project for Zimbabwe and Mozambique

International Groundwater Resources Assessment Centre

1 IGRAC: Georeferencing Methodology

1.1 Mozambique

1. Create a QGIS project for Mozambique and import the Mozambique Hidrogeological Map_North Region and Mozambique Hidrogeological Map_South Region tifs into the project.
2. Find the GAUL dataset, as this was the suggested reference dataset. The dataset was downloaded from here.
3. The Coordinate Reference System (CRS) for the Mozambique tifs was unknown. In the "Legenda" (legend) on the Mozambique Hidrogeological Map_South Region image, it is stated "Projeccao Conica Conforme de Lambert" (Lambert Conic Conformal Projection). The Lambert Conic Conformal Projection requires two parallels, a central meridian, and a Datum. The two parallels and the central meridian were obtained from the scanned maps, and then through research the datum was discovered to be the Tete datum (discovered through this column by Clifford J. Mugnier for ASPRS.org). A custom CRS was the made using that information.
4. The custom CRS using the Tete datum worked but, due to lack of information on the scanned images, was not accurate for georeferencing. A new custom CRS based on the WGS84 datum was made using the proj4 string:

Code:

```
+proj=lcc +lat_0=0 +lon_0=35.5 +lat_1=-14 +lat_2=-24 +x_0=0 +y_0=0 +datum=WGS84 +untis=m +no_defs
```

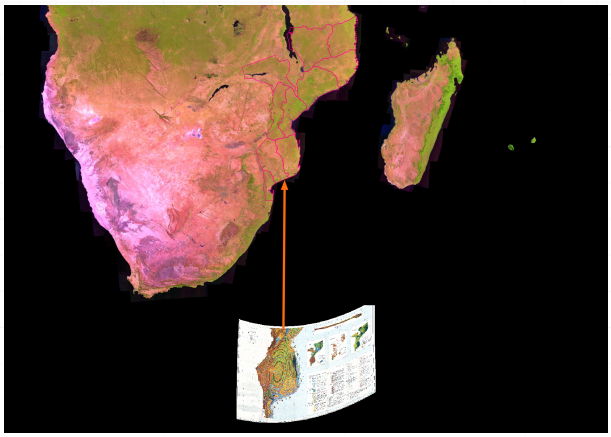
The new CRS was made to save having to datum transformations in the future. It was then decided that georeferencing would be done using the GAUL dataset for reference points because a graticule transformation was not possible.

1.1.1 Georeferencing Mozambique Southern Region

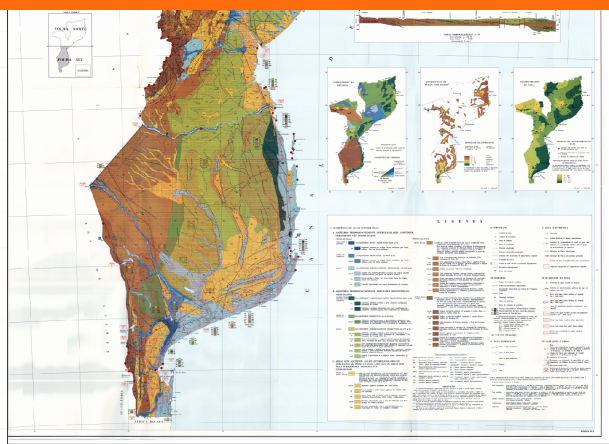
Note:

Multiple iterations were required as the georeferencing was done against a reference dataset and could not be done by simply taking the corner graticules from the tifs and projecting them into the custom CRS.

1. The first iteration of georeferencing the Southern Region of Mozambique was done using a linear transformation. The resulting image ended up being too different to the reference dataset spatially and so was immediately discarded.



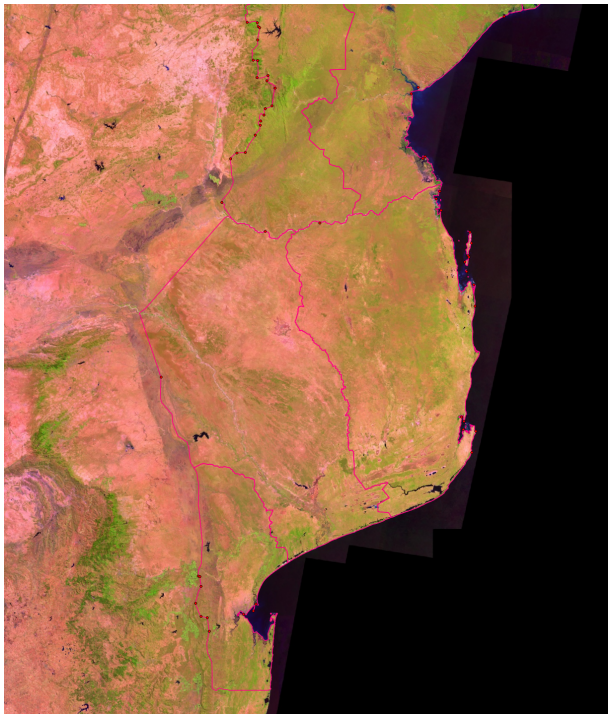
2. The second iteration was done using a Helmert transformation and gave a decent result but there were far too many discrepancies between the reference layer and the georeferenced image.
3. The third iteration was done using the Helmert Transform again but with all Residual pixels for the Ground Control Points (CPs) being under 10. 9 GCPS were used for the referencing. There were too many discrepancies between the reference dataset and the georeferenced image so it was disregarded as a viable image.
4. The fourth iteration was done using a Polynomial 1 Transformation (with all residual pixels less than 10) and 12 GCPs. Again, there were too many discrepancies between the reference dataset and the georeferenced image but there was minimal warping on the polygons.
5. The fifth iteration was done using a Polynomial 2 Transformation (with all residual pixels less than 10) and 17 GCPs. There were fewer discrepancies between the reference dataset and the georeferenced image than in the previous iterations but there was slight warping on the polygons.
6. The sixth iteration was done using a Polynomial 3 Transformation (with all residual pixels less than 10) and 18 GCPs. There were fewer discrepancies between the reference dataset and the georeferenced image than in the previous iterations but there was warping on the polygons.
7. The seventh iteration was done using a Thin Plate Spline transformation. All residual pixels were zero but this was likely a false result. This iteration had the best results for lining up the georeferenced image with the reference dataset. 102 GCPs were used to help correct discrepancies from previous iterations. The issue with this transformation was the significant warping of the polygons in the georeferenced image.
8. For the last iteration, it was decided that the Polynomial 1 transformation warped the internal polygons of Mozambique's Southern Region the least but more GCPs would help with the discrepancies between the reference dataset and the tif. 55 GCPs, with residual pixels lower than 10, were used for the georeferencing and ended up with the best result. The image below shows the georeferencer with the GCPs on the map and the GCP table showing the residuals being below 10 (Ground Control Point 40 has the highest Residual Pixels at 8.976885).



GCP table

Enabled	ID	Source X	Source Y	Dest. X	Dest. Y	dX (pixels)	dY (pixels)	Residual (pixels)
✓	40	3404.4798	-9675.4694	-340498.81	-2927800.27	-4.014147	-8.029389	8.976885
✓	17	6906.1683	-2897.3234	-41672.4568	-2356368.14	0.116355	8.424800	8.425604
✓	34	3896.8232	-2313.5882	-293794.20	-2305850.64	8.366036	-0.336645	8.372807
✓	51	3678.2719	-2403.9012	-312844.42	-2314139.59	2.027779	7.843456	8.101339
✓	53	5061.1685	-3403.7870	-197141.33	-2397530.52	-0.894672	-7.987702	8.037651
✓	21	4087.8013	-370.605440	-277025.46	-2142459.29	4.559890	6.362147	7.827484
✓	32	4133.6304	-1901.2683	-274415.18	-2270501.16	-1.528750	-7.345832	7.503221
✓	24	4085.8149	-906.649603	-277716.81	-2186538.57	1.289620	-7.315953	7.428747
✓	46	7368.9492	-3536.4571	-2434.9155	-2409427.04	7.296693	0.533485	7.316169
✓	30	4188.2564	-1730.4380	-269633.43	-2257315.54	-0.503798	7.060268	7.078220
✓	2	6463.8137	-1404.5816	-78501.7271	-2229802.49	-6.820701	-1.690515	7.027077
✓	8	7446.1916	-4342.9709	2571.9518	-2477655.90	-6.444464	2.091488	6.775355

The image below shows the GCPs for Mozambique's Southern Region relative to the GAUL reference dataset.



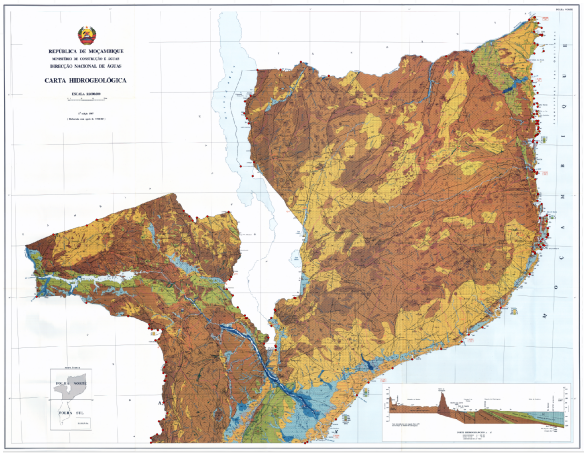
The GCPs for Mozambique's Southern Region can be found here. There were still small discrepancies between the georeferenced image and the reference dataset but to correct the discrepancies would warp the polygons too much to be a viable image.

1.1.2 Georeferencing Mozambique Northern Region

Note:

Fewer iterations were required for the Northern Region of Mozambique as it was done using knowledge gained from georeferencing the Southern Region. The first 2 iterations were also done before the final iteration of georeferencing the Zouth Region of Mozambique so were done before it was decided that using the Thin Plate Spline transformation was providing false results.

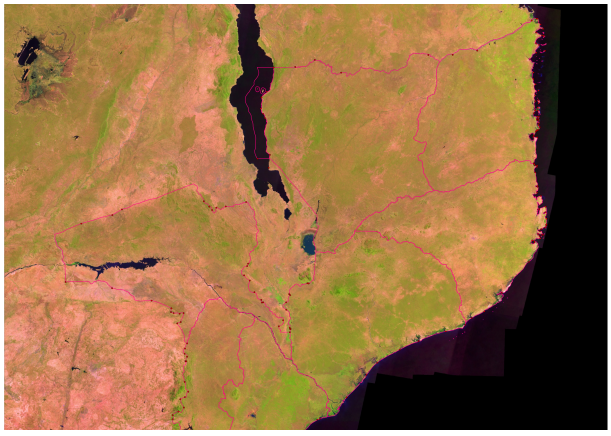
1. The first iteration was done using the Thin Plate Spline transformation where all the residual pixels were a false zero. 10 GCPs were used and resulted in good approximation that had some discrepancies between the georeferenced image and the reference dataset.
2. The second iteration also used the Thin Plate Spline transformation. All the Residual Pixels were zero. Points were added to the previous attempt's GCPs to total 201 GCPs. The Northern Region of Mozambique had many small islands and outlying points that had to be 'forced' into the correct place.
3. The last iteration of georeferencing Mozambique's Northern Region was done using a Polynomial 3 transformation with 149 GCPs. All the Residual Pixels for the GCPs were under 10. A polynomial 3 transformation was chosen as it warped the image the least but had the fewest discrepancies between the reference dataset and the georeferenced image. The image below shows the georeferencer with the GCPs on the map and the GCP table showing the residuals being below 10 (Ground Control Point 124 has the highest Residual Pixels at 9.948533).



GCP table

Enabled	ID	Source X	Source Y	Dest. X	Dest. Y	dX (pixels)	dY (pixels)	Residual (pixels)
✓	124	13820.6497	-2489.8600	547430.16	-1375803.10	4.648851	-8.795538	9.948533
✓	55	13902.4872	-2614.4079	554374.37	-1386635.54	7.011045	-7.002337	9.908960
✓	85	7611.4868	-7658.0422	18807.9399	-1811587.78	-9.338396	3.236964	9.883500
✓	54	13915.1731	-2704.7623	555187.68	-1394226.42	5.342640	-8.032720	9.647196
✓	123	6381.9529	-3007.9004	-83781.5031	-1418550.11	-4.841487	8.267735	9.580993
✓	145	6828.4184	-8847.6888	-47377.2778	-1911697.51	-9.570307	-0.265958	9.574002
✓	139	6019.9561	-7595.5836	-114811.07	-1805572.58	-4.284493	-8.541520	9.555859
✓	16	12954.8048	-7914.9026	469780.73	-1834016.94	4.532569	-8.280012	9.439427
✓	18	13848.2085	-6512.0380	545787.72	-1716617.08	6.378409	-6.937360	9.423962
✓	61	13829.4150	-1322.9850	549843.13	-1276887.15	5.233735	-7.697691	9.308407
✓	99	2224.1852	-7786.5180	-434096.45	-1823274.76	-5.687354	7.252747	9.216742
✓	14	12131.0444	-8611.3013	400394.53	-1892206.91	5.893561	-6.913481	9.084618

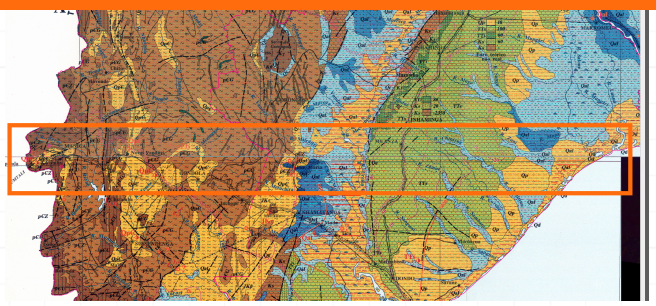
The image below shows the GCPs for Mozambique's Northern Region relative to the GAUL reference dataset.



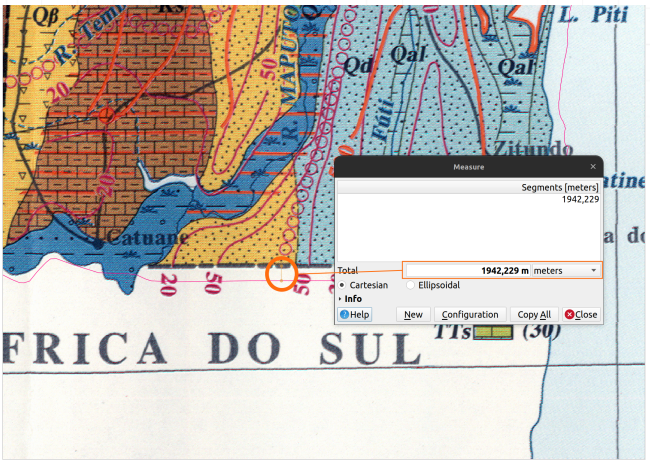
The GCPs for Mozambique's Northern Region can be found here.

1.1.3 Checking Alignment of georeferenced Mozambique Images

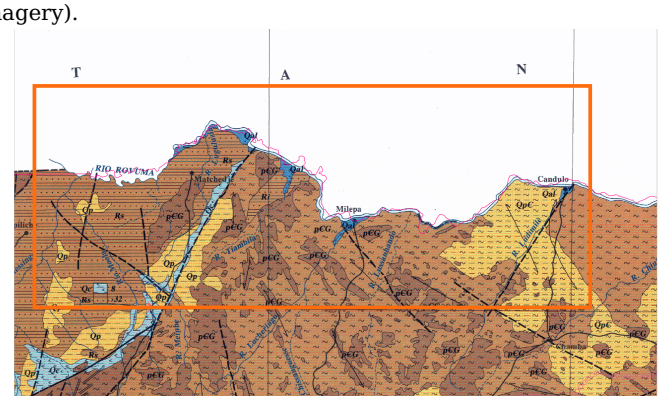
Due to the nature of the scanned maps, there are slight discrepancies where the sheets meet eachother but they are minimal.



The main discrepancy for Mozambique's Southern Region between the southernmost border of the **Mozambique Hidrogeological Map_South Region** tif and the GAUL reference dataset.



There are multiple discrepancies between the Northern Border of **Mozambique Hidrogeological Map_North Region** tif and the GAUL reference dataset. There are discrepancies where the administration boundary on the map and the GAUL boundray differ, and where the Ruvuma (formerly Rovuma) River's path on the **Mozambique Hidrogeological Map_North Region** tif differs from its path on the GAUL dataset(the path on the GAUL dataset is ratified by GLAD Landsat imagery).



A closer look at one of the worse discrepancies:

1.2 Zimbabwe

1. Create a QGIS project for Zimbabwe and import the four Zimbabwe sheets into the project.
2. Attempt to find any info about Zim projection, datum, etc. through research. Finding the memoir for the sheets would be the ideal situation. The maps being from 1986 mean that their projection system was most likely based on the Arc 1950 datum (which is based on the Clarke 1880 ellipsoid).
3. Research did not yield the projection used for creating the sheets. The projection is required to make custom projected CRS for better georeferencing. The closest option found during research was a report from UNESCO from 1995 (Hydrogeological Maps A Guide and Standard Legend. Vol. 17., by Struckmeier, Wilhelm F, and Jean Margat) referencing the Zimbabwe Hydrogeological maps stating that "preferably UTM grid" was used for map locations. Using this information, a custom tmerc (Transverse mercator) projection system was made using the proj string:

```
Custom CRS: +proj=tmerc +lat_0=-19 +lon_0=30 +k=1 +x_0=0 +y_0=0 +a=6378249.145 +rf=2
```

The Central Meridian (lon_0) and the Latitude of Origin (lat_0) were taken from the Zimbabwe sheets and the **+a** and **+rf** values are for the Arc 1950 datum.

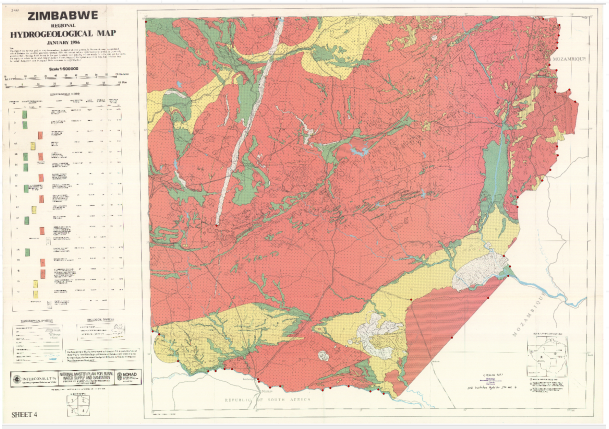
1.2.4 Georeferencing Zimbabwe Sheets 1 through 4

1. The first iteration of georeferencing Zimbabwe Sheet 1 was done using a Thin Plate Spline transformation using 11 GCPS. There were multiple discrepancies between the georeferenced image and the reference image.
2. The first iteration of georeferencing Zimbabwe Sheet 2 was done using a Thin Plate Spline transformation using 12 GCPS. There were multiple discrepancies between the georeferenced image and the reference image.
3. The first iteration of georeferencing Zimbabwe Sheet 3 was done using a Thin Plate Spline transformation using 8 GCPS. There were multiple discrepancies between the georeferenced image and the reference image.
4. The first iteration of georeferencing Zimbabwe Sheet 4 was done using a Thin Plate Spline transformation using 12 GCPS. There were multiple discrepancies between the georeferenced image and the reference image. It was the best result of the initial georeferencing attempts so will be refined first and then used as a basis for the other images.

1.2.5 Refining Zimbabwe Sheet 4

1. The initial refinement of Zimbabwe sheet 4 was done using a Thin Plate Spline transformation with 46 GCPS. There were still discrepancies between the GAUL reference dataset and the georeferenced sheet's boundaries.
2. The second iteration of refining Zimbabwe sheet 4 was done using a Thin Plate Spline transformation with 75 GCPS. There were still major discrepancies between the GAUL reference dataset and georeferenced sheet's boundaries. This transformation type also warped the internal polygons of the georeferenced image.

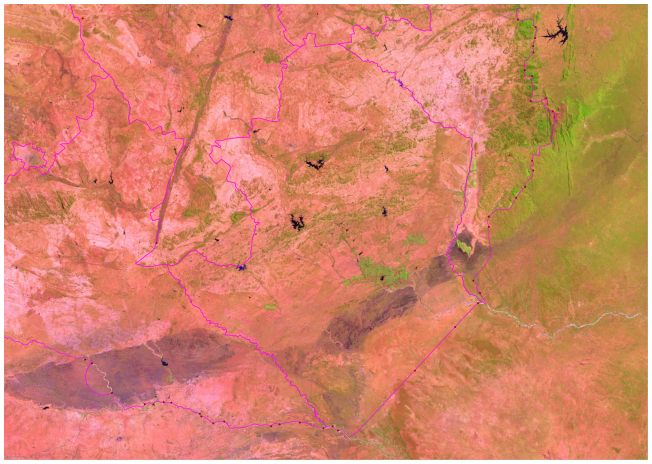
3. After a discussion, it was decided that the Residual Pixels from a Thin plate Spline transformation being zero was a false reading and so a Polynomial 3 transformation was chosen as it gave the best results. 42 GCPs were used with all the Residual Pixels being less than 10 (Point 16 was the GCP with the highest Residual Pixel value of 9.592993). The image below shows the georeferencer and part of the associated GCP table:



GCP table

Enabled	ID	Source X	Source Y	Dest. X	Dest. Y	dX (pixels)	dY (pixels)	Residual (pixels)
✓	16	12059.0472	-4535.4099	259773.52	-181120.12	-6.326829	7.210877	9.592993
✓	3	11733.3660	-5908.9306	245964.80	-238435.51	-0.595306	-9.522927	9.541516
✓	11	7495.1895	-9050.6090	64933.2807	-368828.63	0.831252	-9.378813	9.415579
✓	22	12979.4776	-1048.5164	299817.04	-33958.3403	6.803286	6.294343	9.268412
✓	13	9566.1123	-6930.7185	154118.78	-281464.46	-1.052394	9.043016	9.104047
✓	5	13265.3508	-2020.4474	311072.64	-75148.3183	-5.565829	7.097632	9.019692
✓	17	12633.2722	-4145.0603	284349.58	-164179.58	0.932807	-8.728199	8.777903
✓	6	4818.4554	-8601.7675	-48089.7288	-349074.12	-7.541414	-4.098827	8.583315
✓	36	12978.3858	-3347.1085	298619.68	-130956.94	-8.278182	1.820911	8.476085
✓	15	12053.5574	-4764.4148	259969.93	-190228.73	4.621455	-7.103343	8.474393
✓	9	6003.6210	-8847.8507	2427.3734	-360102.22	8.281035	-1.662210	8.446211
✓	4	12653.6130	-367.840350	286076.29	-4596.7382	-0.304856	-8.247223	8.252856

The image below shows the GCPs for Zimbabwe Sheet 4 relative to the GAUL reference dataset over the GLAD Landsat dataset.



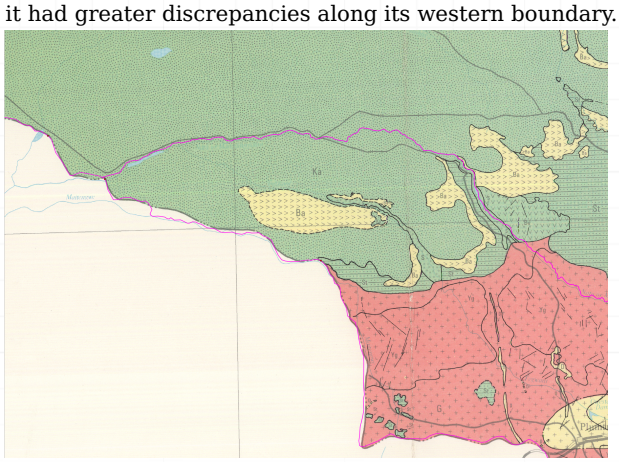
The GCPs can be found here.

> 👩‍🔧 **Note:**
>
> There was not a lot of spatial information in the North Western section (Central region of Zimbabwe) of the sheet and the roads that are represented on the scanned map do not intersect with each other as they would in the real world.

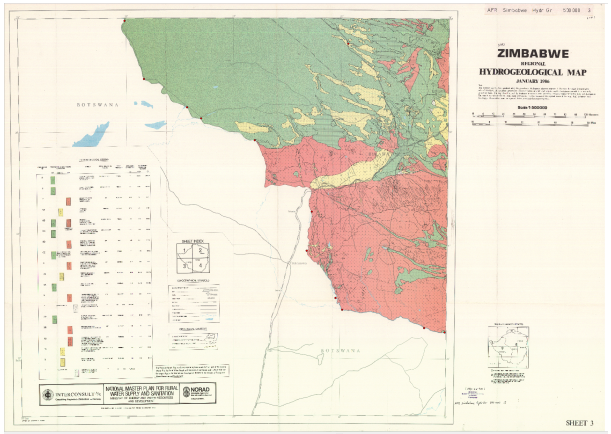
1.2.6 Refining Zimbabwe Sheet 3

1. The initial refinement of Zimbabwe Sheet 3 was done using a Thin Plate Spline transformation with 29 GCPS. There were significant discrepancies between the GAUL reference dataset and georeferenced sheet's boundaries.

2. The second iteration of refining Zimbabwe Sheet 3 was done using a Polynomial 1 transformation with 11 GCPs. All the Residual Pixels for the GCPs were lower than 10. Despite Sheet 3 having all if its GCPs be lower values than Sheet 4's GCPs it had greater discrepancies along its western boundary.



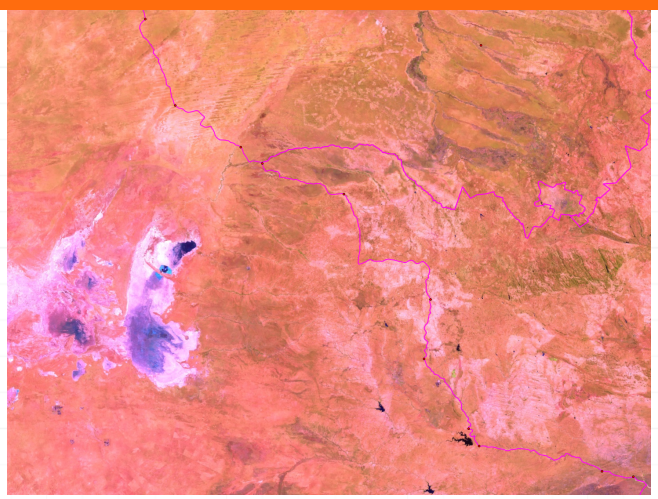
The image below shows the georeferencer and part of the associated GCP table (Point 7 was the GCP with the highest Residual Pixel value of 1.211239):



GCP table

Enabled	ID	Source X	Source Y	Dest. X	Dest. Y	dX (pixels)	dY (pixels)	Residual (pixels)
✓	7	4229.5936	-2390.3705	-359970.85	-92100.4220	-1.033199	-0.632139	1.211239
✓	4	4553.7313	-2645.9286	-345973.13	-102731.61	0.939693	0.613419	1.122188
✓	1	7659.1106	-6797.0951	-211586.25	-275514.75	-0.721744	0.058489	0.724110
✓	2	7813.4818	-7069.6268	-204756.36	-286810.34	0.516433	-0.030211	0.517316
✓	9	6994.1599	-5704.2594	-240618.99	-230022.46	0.231755	-0.052456	0.237617
✓	6	3232.7300	-1737.1763	-402663.85	-64939.2455	0.125027	0.072250	0.144402
✓	8	7112.6177	-4801.7050	-236379.00	-191344.14	-0.056514	-0.003300	0.056610
✓	5	5794.3286	-3150.5829	-292822.56	-122734.22	-0.039641	-0.037575	0.054620
✓	0	2782.8002	-392.543470	-422174.59	-8667.3601	0.026736	0.010513	0.028729
✓	3	10180.5060	-7593.7264	-104343.81	-306840.31	0.005428	0.006413	0.008402
✓	10	9708.6167	-7493.0649	-124735.05	-303286.80	0.005596	-0.005583	0.007905
✓	11	7956.7356	-920.668256	-203544.52	-25754.9338	0.000430	0.000179	0.000466

The image below shows the GCPs for Zimbabwe Sheet 4 relative to the GAUL reference dataset over the GLAD Landsat dataset.

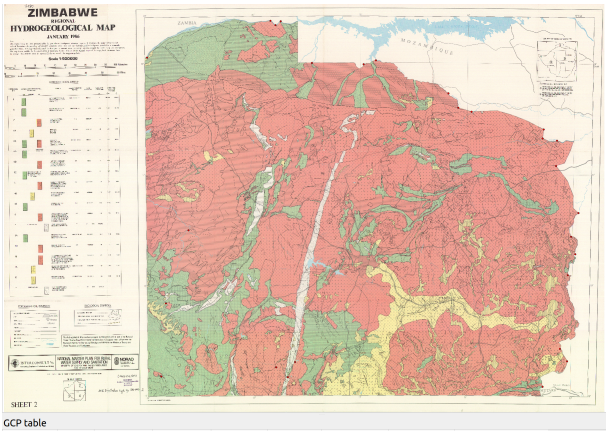


The GCPs can be found here. A Polynomial 1 transformation was used as it gave the best result out of the transformation types.

> 👷 **Note:**
>
> There was not a lot of spatial information in the North Eastern section (Central region of Zimbabwe) of the sheet and the roads that are represented on the scanned map do not intersect with each other as they would in the real world.

1.2.7 Refining Zimbabwe Sheet 2

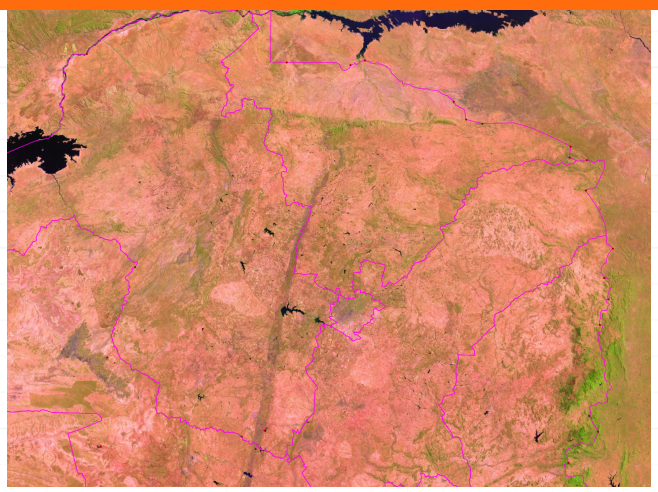
1. The initial refinement of Zimbabwe sheet 2 was done using a Thin Plate Spline transformation with 53 GCPS. There were significant discrepancies between the GAUL reference dataset and georeferenced sheet's boundaries.
2. The second iteration of refining Zimbabwe Sheet 2 was done using a Polynomial 3 transformation with 27 GCPs. All the Residual Pixels for the GCPs were lower than 10. The image below shows the georeferencer and part of the associated GCP table (Point 18 was the GCP with the highest Residual Pixel value of 9.794608):



GCP table

Enabled	ID	Source X	Source Y	Dest. X	Dest. Y	dX (pixels)	dY (pixels)	Residual (pixels)
✓	18	6395.6281	-502.735365	23140.0702	369468.74	9.231988	3.271808	9.794608
✓	8	13372.1546	-5498.0493	316784.98	156469.74	7.543473	5.523552	9.349525
✓	0	13017.4568	-3827.1430	301899.16	227662.35	3.488188	-7.616324	8.377103
✓	13	10709.5120	-2451.5250	204372.01	285404.22	1.603954	8.113146	8.270176
✓	12	11022.5811	-2546.1306	217632.16	281982.24	1.783547	-8.039807	8.235261
✓	2	12698.9310	-9038.2895	287748.14	5991.7156	7.277971	-0.252950	7.282365
✓	15	8766.0521	-1333.4240	122032.78	333590.91	-6.840831	-1.995618	7.125971
✓	16	8475.5004	-1425.3317	110242.89	329728.71	5.950349	2.904295	6.621298
✓	7	13471.0730	-5910.4260	320354.04	139087.57	-5.321654	2.557996	5.904519
✓	4	13109.5092	-8142.3982	304764.94	44218.4188	-4.945320	-2.789172	5.677646
✓	14	10477.9190	-2112.0170	194347.43	300012.62	-5.474623	0.912162	5.550093
✓	6	13294.9153	-6404.9639	313161.42	118354.28	1.604331	-5.105242	5.351390

The image below shows the GCPs for Zimbabwe Sheet 2 relative to the GAUL reference dataset over the GLAD Landsat dataset.

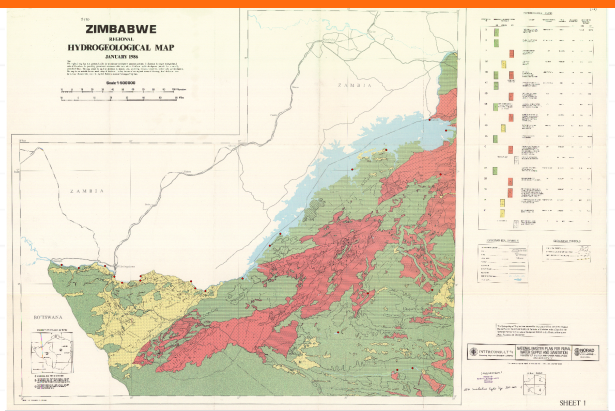


The GCPs can be found here. A Polynomial 3 transformation was used as it gave the best result out of the transformation types.

> **Note:**
>
> There was not a lot of spatial information in the South Western section (Central region of Zimbabwe) of the sheet and the roads that are represented on the scanned map do not intersect with each other as they would in the real world. The Great Dyke was used where it was clear that a GCP could be placed.

1.2.8 Refining Zimbabwe Sheet 1

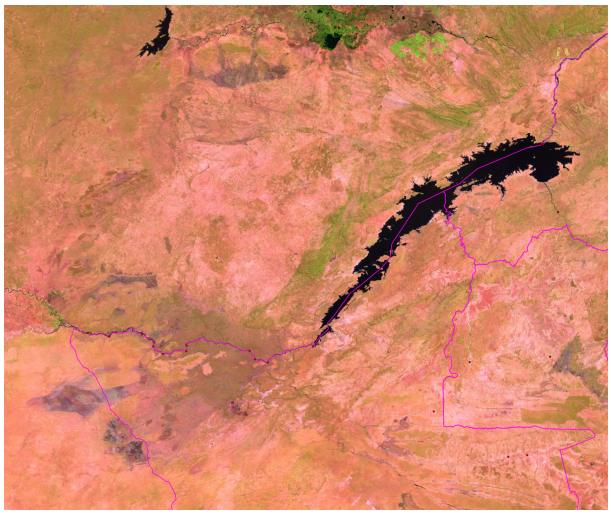
1. The initial refinement of Zimbabwe sheet 1 was done using a Thin Plate Spline transformation with 53 GCPS. There were significant discrepancies between the GAUL reference dataset and georeferenced sheet's boundaries.
2. The second iteration of refining Zimbabwe Sheet 1 was done using a Polynomial 1 transformation with 23 GCPs. A Polynomial 1 transformation was used as it gave the best result out of the transformation types. Lake Kariba would have been ideal for reference points, however in the years since the map was published Lake Kariba's water level has lowered significantly and the shoreline has changed. All the Residual Pixels for the GCPs were lower than 10 (Point 17 was the GCP with the highest Residual Pixel value of 7.111510).



GCP table

Enabled	ID	Source X	Source Y	Dest. X	Dest. Y	dX (pixels)	dY (pixels)	Residual (pixels)
✓	17	8820.9022	-3657.9982	-174828.08	241075.15	3.302497	-6.298181	7.111510
✓	16	8028.6930	-4112.7822	-208041.10	220723.10	-1.406674	6.865745	7.008365
✓	0	9076.7494	-8778.4376	-160468.49	24868.7076	-1.327313	-6.863767	6.990927
✓	1	9430.7815	-8736.7930	-145596.03	26235.7324	1.860861	6.443499	6.706824
✓	9	2673.8510	-6748.6568	-432582.52	105910.77	-5.763949	-2.236409	6.182607
✓	14	6111.1510	-5885.7171	-287939.63	144990.53	-3.704304	-4.054105	5.491597
✓	8	2493.2100	-6607.0927	-440150.05	111777.51	-4.494837	-3.058648	5.436809
✓	6	1666.6038	-6406.1983	-474815.42	119478.80	4.975378	1.436778	5.178679
✓	10	3094.9913	-6698.7002	-414294.34	108051.08	2.737396	4.308065	5.104190
✓	13	5475.8157	-6650.4754	-314101.09	111930.39	-1.758300	3.646664	4.048429
✓	12	4376.7787	-6879.7114	-359964.87	101591.93	3.897226	-0.443525	3.922382
✓	3	4611.7686	-6941.0578	-350020.27	99194.2403	3.659892	-0.674746	3.721571

The image below shows the GCPs for Zimbabwe Sheet 1 relative to the GAUL reference dataset over the GLAD Landsat dataset.

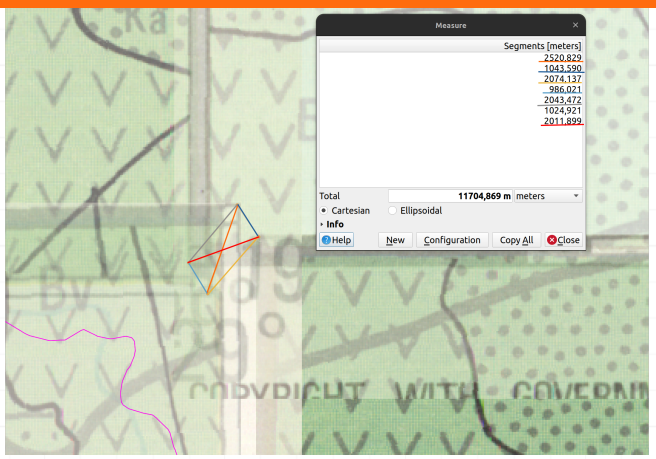


The GCPs can be found here.

> 👩‍🔧 **Note:**
>
> There was some clear spatial information in the South Eastern section (Central region of Zimbabwe) of the sheet and the roads that are represented on the scanned map do not intersect with each other as they would in the real world. The Shangani River was used where it was clear that GCPs could be placed.

1.2.9 Checking Alignment of Zimbabwe sheets

Due to the nature of the scanned maps, there are discrepancies where the sheets interact with eachother and there was a lack of reference information to be used as GCPs. The main discrepncies are where the four sheets meet in the central area of Zimbabwe. The discrepancies and their respective measurements are highlighted using the same coloured lines in the image below. The discrepancies between sheets edges get smaller closer to the country boundary of Zimbabwe due to there being multiple GCPs along the administration boundary.


Measure
Segments [meters]
2520,829
1043,590
2074,137
986,021
2043,472
1024,921
2011,899
Total
11704,869 m
meters
Cartesian
Ellipsoidal
Info
Help
New
Configuration
Copy All
Close