Virtual Rasters

tiny files, big impact

Justin McAllister

CUGOS Fall Fling 2019

About Me

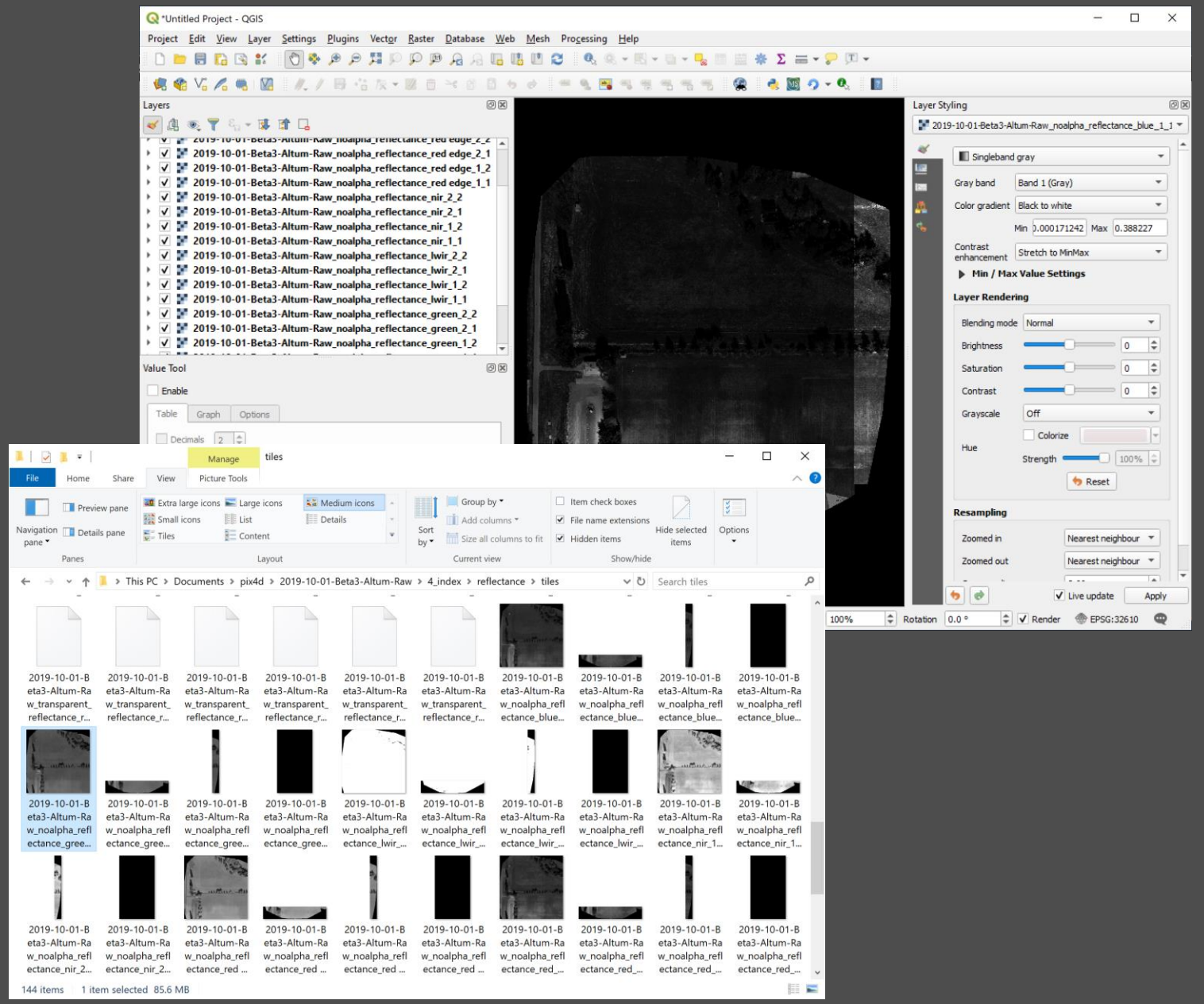
- BSEE (University of Texas-Dallas)
- CTO @micasense designing multispectral sensors for drones
- Enjoy family, skiing, hiking, biking, ham radio

MicaSense

Rasters are a pain

- Too many tiles
- Too many layers
- Raster Calculator!
  - Not programmatic
  - Output files are huge

- Is there a better way?



Virtual Rasters

- XML Files (really small text files)
- Original use was to combine tiled, geolocated imagery
- Contain reference to other files
- GDAL treats them as first class citizens – just another input type

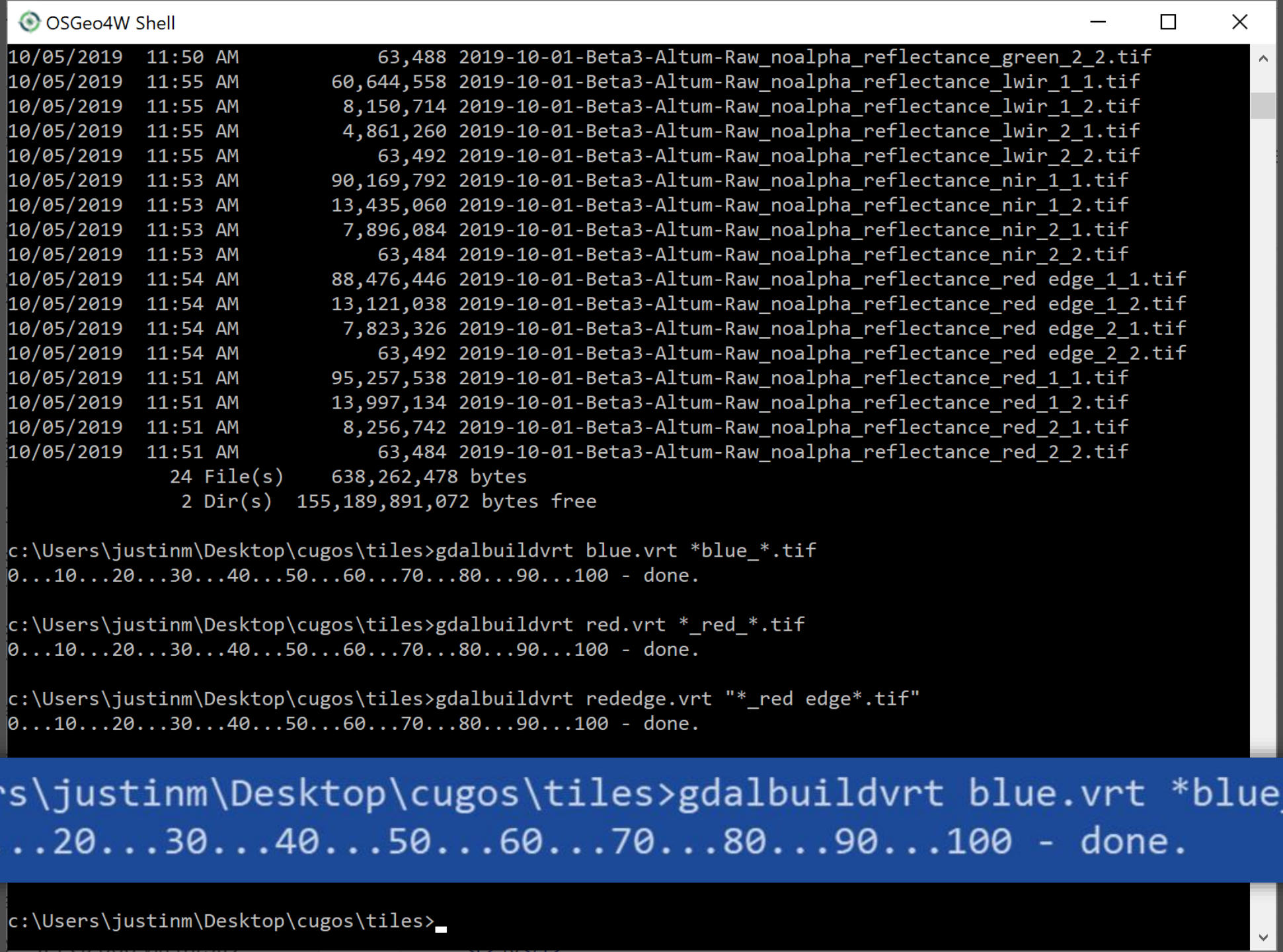
- https://gdal.org/drivers/raster/vrt.html

Basic VRT Creation

- gdalbuildvrt output.vrt list_of_inputs*
  - all files into the same layer

- gdalbuildvrt output.vrt list_of_inputs --separate
  - Each file into a different layer in the order specified
  - Files needs to have the same reference system and type (e.g. all UINT16)

- Super useful for combining datasates that come in tiles or individual layers


OSGeo4W Shell
10/05/2019 11:50 AM 63,488 2019-10-01-Beta3-Altum-Raw_noalpha_reflectance_green_2_2.tif
10/05/2019 11:55 AM 60,644,558 2019-10-01-Beta3-Altum-Raw_noalpha_reflectance_lwir_1_1.tif
10/05/2019 11:55 AM 8,150,714 2019-10-01-Beta3-Altum-Raw_noalpha_reflectance_lwir_1_2.tif
10/05/2019 11:55 AM 4,861,260 2019-10-01-Beta3-Altum-Raw_noalpha_reflectance_lwir_2_1.tif
10/05/2019 11:55 AM 63,492 2019-10-01-Beta3-Altum-Raw_noalpha_reflectance_lwir_2_2.tif
10/05/2019 11:53 AM 90,169,792 2019-10-01-Beta3-Altum-Raw_noalpha_reflectance_nir_1_1.tif
10/05/2019 11:53 AM 13,435,060 2019-10-01-Beta3-Altum-Raw_noalpha_reflectance_nir_1_2.tif
10/05/2019 11:53 AM 7,896,084 2019-10-01-Beta3-Altum-Raw_noalpha_reflectance_nir_2_1.tif
10/05/2019 11:53 AM 63,484 2019-10-01-Beta3-Altum-Raw_noalpha_reflectance_nir_2_2.tif
10/05/2019 11:54 AM 88,476,446 2019-10-01-Beta3-Altum-Raw_noalpha_reflectance_red edge_1_1.tif
10/05/2019 11:54 AM 13,121,038 2019-10-01-Beta3-Altum-Raw_noalpha_reflectance_red edge_1_2.tif
10/05/2019 11:54 AM 7,823,326 2019-10-01-Beta3-Altum-Raw_noalpha_reflectance_red edge_2_1.tif
10/05/2019 11:54 AM 63,492 2019-10-01-Beta3-Altum-Raw_noalpha_reflectance_red edge_2_2.tif
10/05/2019 11:51 AM 95,257,538 2019-10-01-Beta3-Altum-Raw_noalpha_reflectance_red_1_1.tif
10/05/2019 11:51 AM 13,997,134 2019-10-01-Beta3-Altum-Raw_noalpha_reflectance_red_1_2.tif
10/05/2019 11:51 AM 8,256,742 2019-10-01-Beta3-Altum-Raw_noalpha_reflectance_red_2_1.tif
10/05/2019 11:51 AM 63,484 2019-10-01-Beta3-Altum-Raw_noalpha_reflectance_red_2_2.tif
24 File(s) 638,262,478 bytes
2 Dir(s) 155,189,891,072 bytes free
c:\Users\justinm\Desktop\cugos\tiles>gdalbuildvrt blue.vrt *blue_*.tif
0...10...20...30...40...50...60...70...80...90...100 - done.
c:\Users\justinm\Desktop\cugos\tiles>gdalbuildvrt red.vrt *_red_*.tif
0...10...20...30...40...50...60...70...80...90...100 - done.
c:\Users\justinm\Desktop\cugos\tiles>gdalbuildvrt rededge.vrt "*_red edge*.tif"
0...10...20...30...40...50...60...70...80...90...100 - done.
c:\Users\justinm\Desktop\cugos\tiles>gdalbuildvrt blue.vrt *blue_*.tif
0...10...20...30...40...50...60...70...80...90...100 - done.
c:\Users\justinm\Desktop\cugos\tiles>

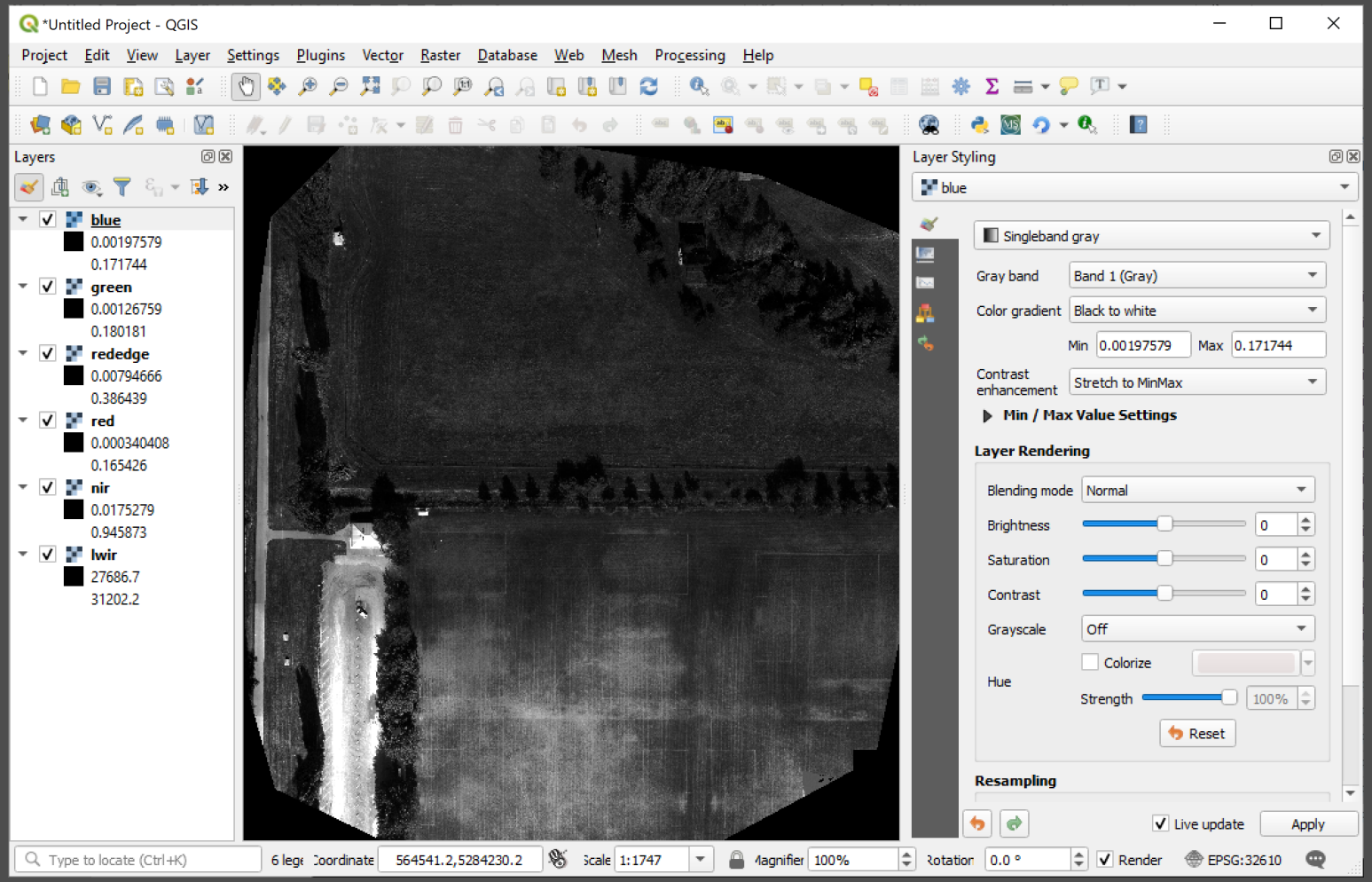
blue.xml

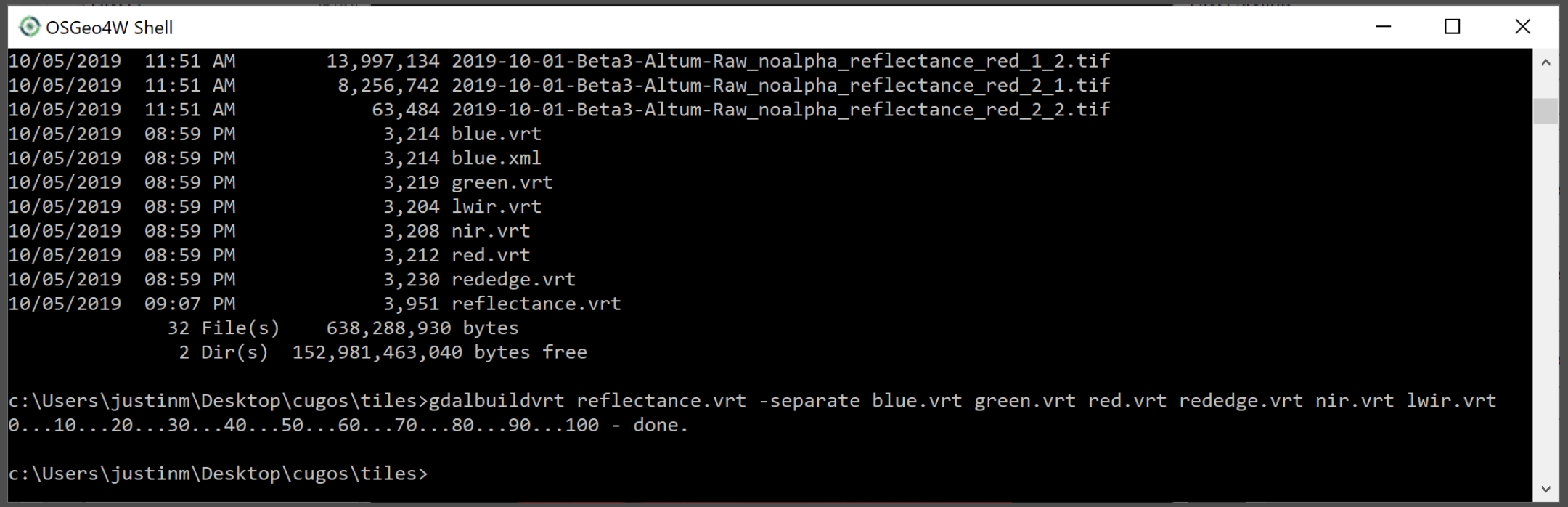
C: > Users > justinm > Desktop > cugos > tiles > blue.xml

```
<VRTDataset rasterXSize="5552" rasterYSize="5986">
  <SRS>PROJCS["WGS 84 / UTM zone 10N",GEOGCS["WGS 84",DATUM["WGS_1984",SPHEROID["WGS 84",6378137,298.257223563,AUTHORITY["EP
  <GeoTransform>  5.6430981462000008e+05,  4.5220000000000003e-02,  0.0000000000000000e+00,  5.2843024623000007e+06,  0.0000
  <VRTRasterBand dataType="Float32" band="1">
    <Metadata>…
    </Metadata>
    <NoDataValue>-10000</NoDataValue>
    <ColorInterp>Gray</ColorInterp>
    <ComplexSource>
      <SourceFilename relativeToVRT="1">2019-10-01-Beta3-Altum-Raw_noalpha_reflectance_blue_1_1.tif</SourceFilename>
      <SourceBand>1</SourceBand>
      <SourceProperties RasterXSize="5000" RasterYSize="5000" DataType="Float32" BlockXSize="5000" BlockYSize="1" />
      <SrcRect xOff="0" yOff="0" xSize="5000" ySize="5000" />
      <DstRect xOff="0" yOff="0" xSize="5000" ySize="5000" />
      <NODATA>-10000</NODATA>
    </ComplexSource>
    <ComplexSource>…
    </ComplexSource>
    <ComplexSource>…
    </ComplexSource>
    <ComplexSource>…
    </ComplexSource>
  </VRTRasterBand>
</VRTDataset>
```

How do we manage separate layers?

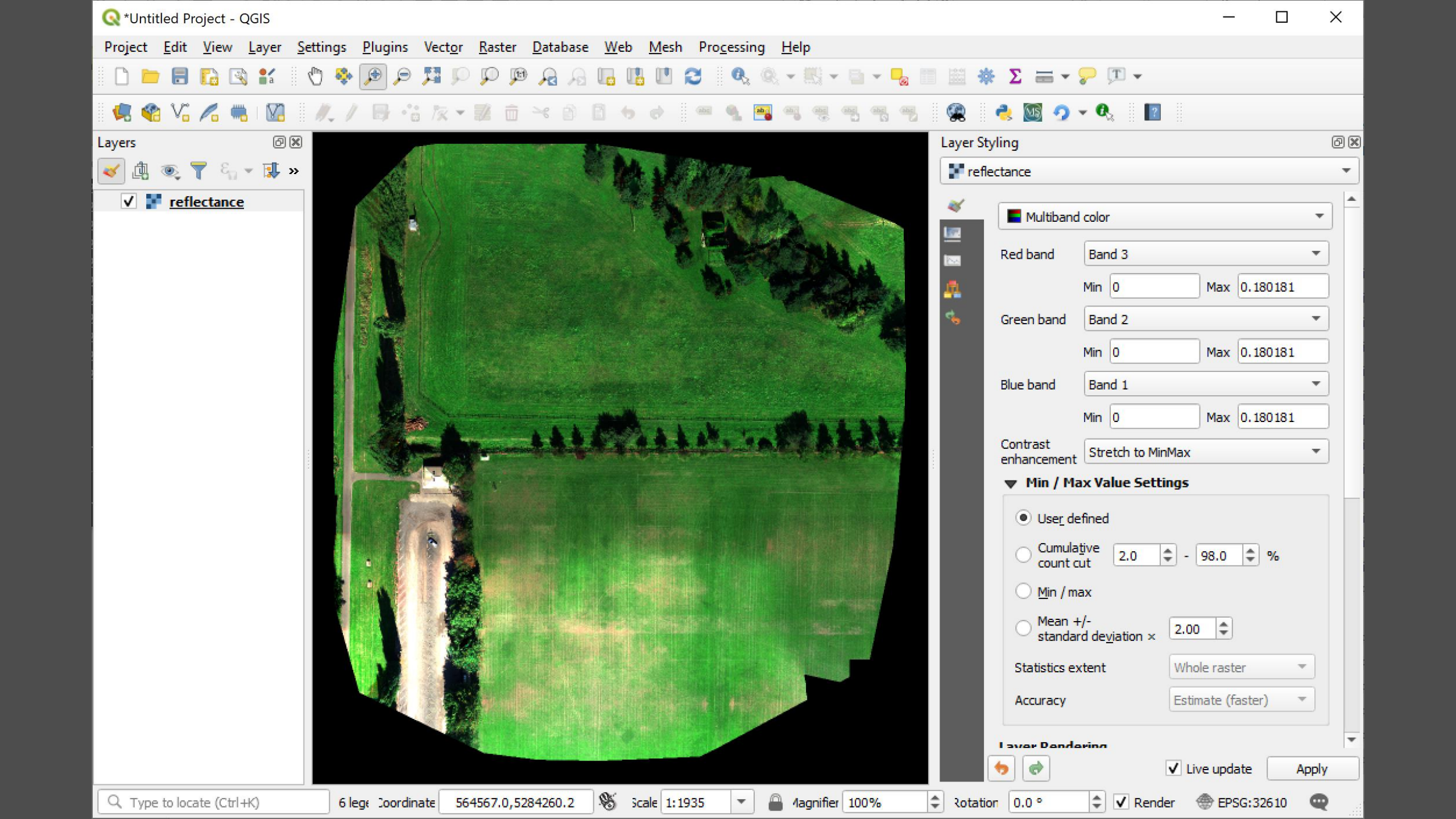
- Create composites?
- Perform math?



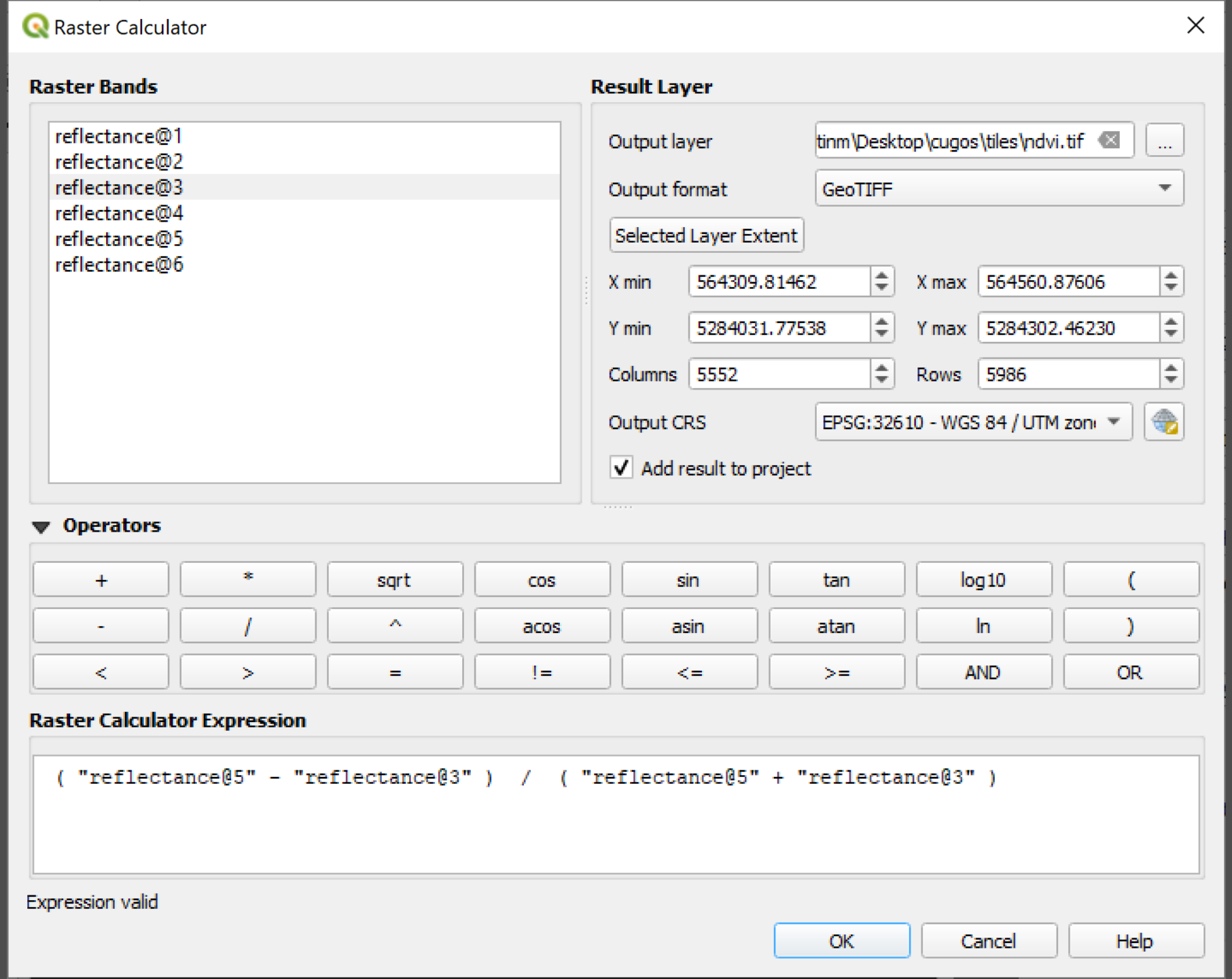


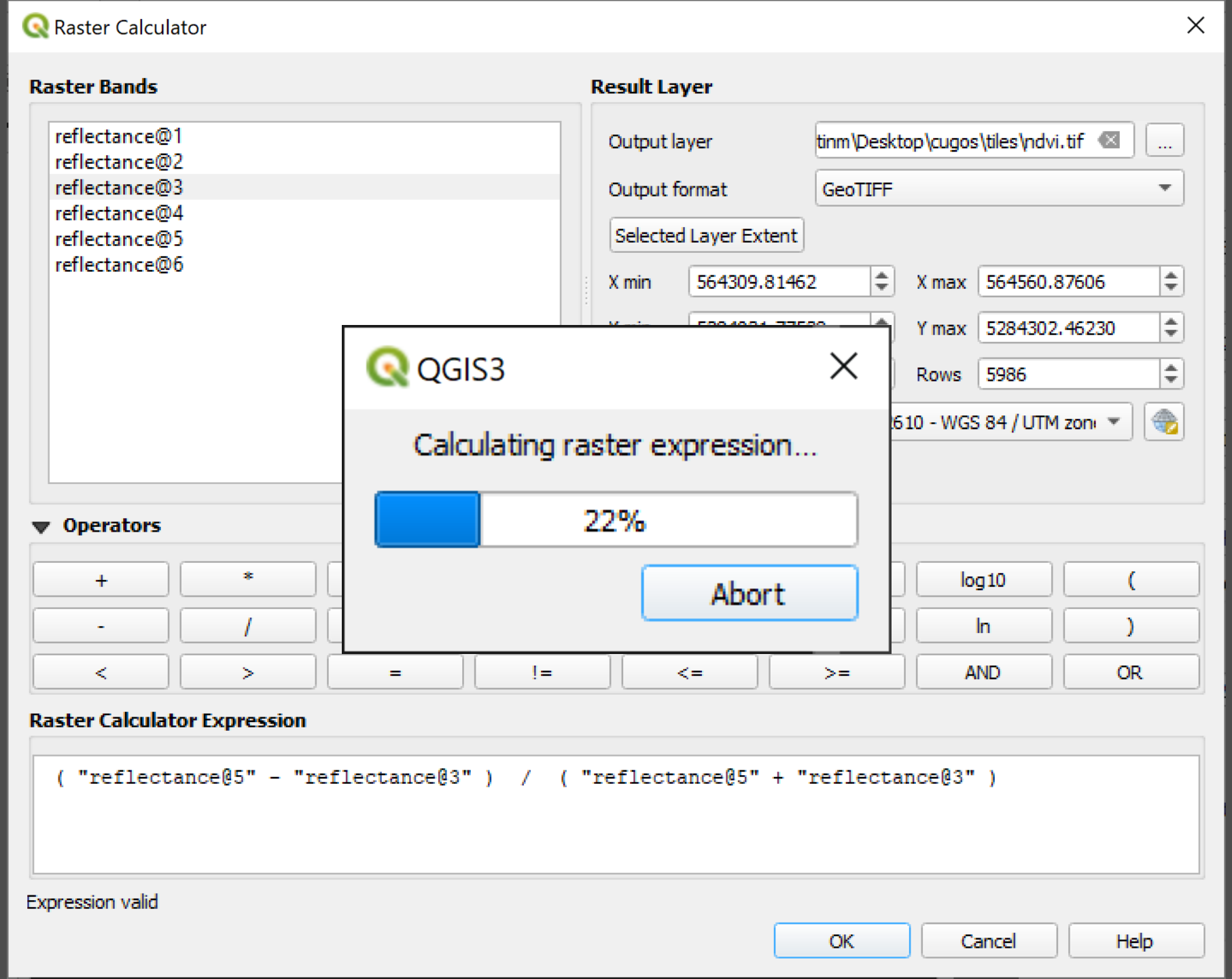

```
>gdalbuildvrt reflectance.vrt -separate
```

```
blue.vrt green.vrt red.vrt rededge.vrt nir.vrt lwir.vrt
```

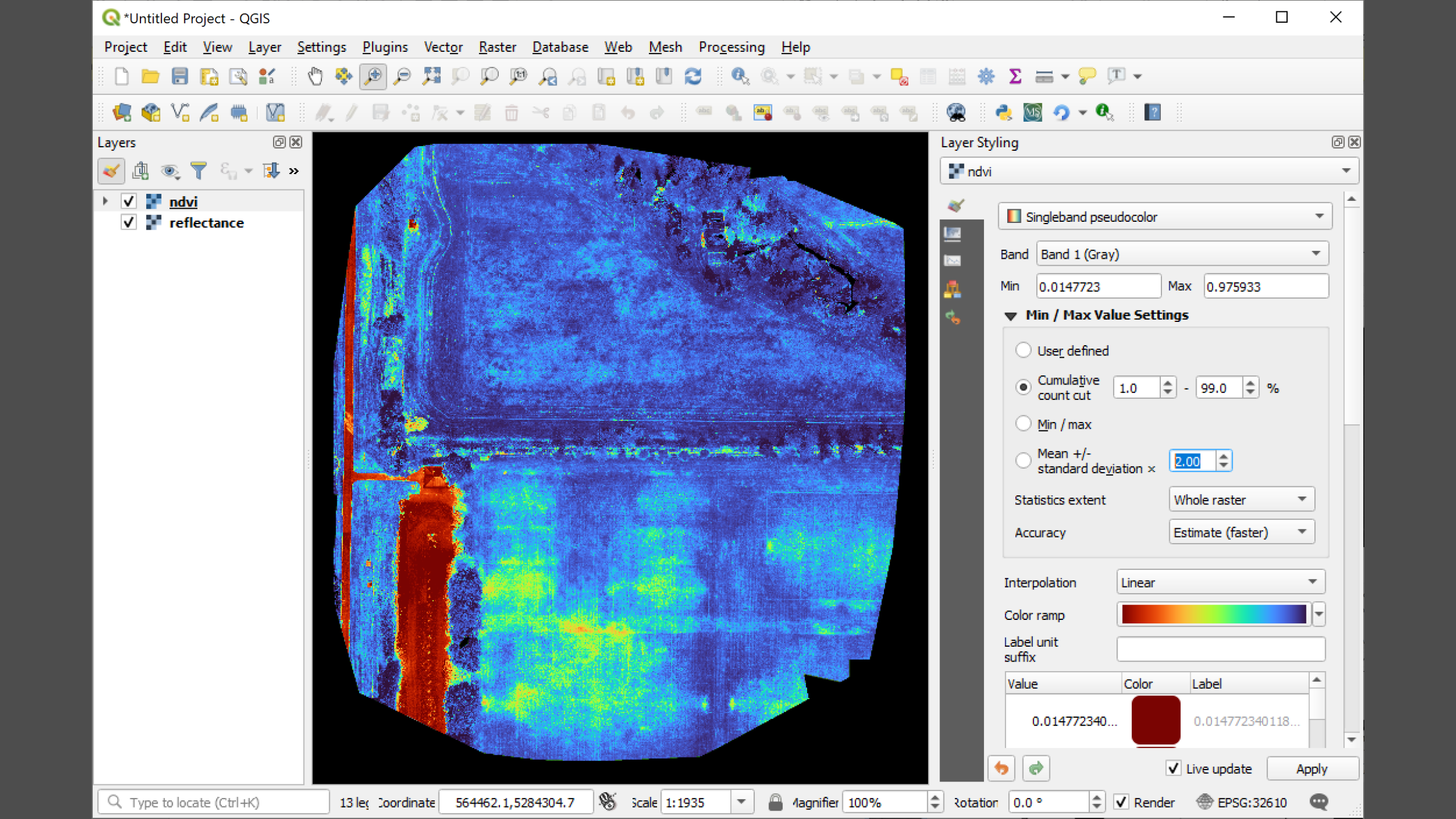

*Untitled Project - QGIS
Project Edit View Layer Settings Plugins Vector Raster Database Web Mesh Processing Help
Layers
reflectance
Layer Styling
reflectance
Multiband color
Red band
Band 3
Min 0
Max 0.180181
Green band
Band 2
Min 0
Max 0.180181
Blue band
Band 1
Min 0
Max 0.180181
Contrast enhancement
Stretch to MinMax
Min / Max Value Settings
User defined
Cumulative count cut
2.0
98.0
%
Min / max
Mean +/- standard deviation ×
2.00
Statistics extent
Whole raster
Accuracy
Estimate (faster)
Live update
Apply
Type to locate (Ctrl+K)
Coordinate
564567.0,5284260.2
Scale
1:1935
Magnifier
100%
Rotation
0.0 °
Render
EPSG:32610

Conventional Raster Calculator


Raster Calculator
Raster Bands
reflectance@1
reflectance@2
reflectance@3
reflectance@4
reflectance@5
reflectance@6
Result Layer
Output layer
tinm\Desktop\cugos\tiles\ndvi.tif
...
Output format
GeoTIFF
Selected Layer Extent
X min
564309.81462
X max
564560.87606
Y min
5284031.77538
Y max
5284302.46230
Columns
5552
Rows
5986
Output CRS
EPSG:32610 - WGS 84 / UTM zon
Add result to project
Operators
+
*
sqrt
cos
sin
tan
log10
(
-
/
^
acos
asin
atan
ln
)
<
>
=
!=
<=
>=
AND
OR
Raster Calculator Expression
("reflectance@5" - "reflectance@3") / ("reflectance@5" + "reflectance@3")
Expression valid
OK
Cancel
Help


Raster Calculator
Raster Bands
reflectance@1
reflectance@2
reflectance@3
reflectance@4
reflectance@5
reflectance@6
Result Layer
Output layer
tinm\Desktop\cugos\tiles\ndvi.tif
Output format
GeoTIFF
Selected Layer Extent
X min
564309.81462
X max
564560.87606
Y max
5284302.46230
Rows
5986
610 - WGS 84 / UTM zon
QGIS3
Calculating raster expression...
22%
Abort
Operators
+
*
log10
(
-
/
ln
)
<
>
=
!=
<=
>=
AND
OR
Raster Calculator Expression
("reflectance@5" - "reflectance@3") / ("reflectance@5" + "reflectance@3")
Expression valid
OK
Cancel
Help

ndvi.tif result is 128MB!


*Untitled Project - QGIS
Project Edit View Layer Settings Plugins Vector Raster Database Web Mesh Processing Help
Layers
ndvi
reflectance
Layer Styling
ndvi
Singleband pseudocolor
Band Band 1 (Gray)
Min 0.0147723 Max 0.975933
Min / Max Value Settings
User defined
Cumulative count cut 1.0 - 99.0 %
Min / max
Mean +/- standard deviation × 2.00
Statistics extent Whole raster
Accuracy Estimate (faster)
Interpolation Linear
Color ramp
Label unit suffix
Value Color Label
0.014772340...
0.014772340118...
Live update
Apply
Type to locate (Ctrl+K)
Coordinate 564462.1,5284304.7
Scale 1:1935
Magnifier 100%
Rotation 0.0 °
Render
EPSG:32610

VrtDerivedRasterBand

- Call code from a VRT that performs some transform on data
- To GDAL, the VRT just acts like another raster containing the result
- No hard drive space is harmed in the making of this VRT

- C/C++ (voodoo)
- Inline Python
- Library python

Mechanics of a Derived VRT

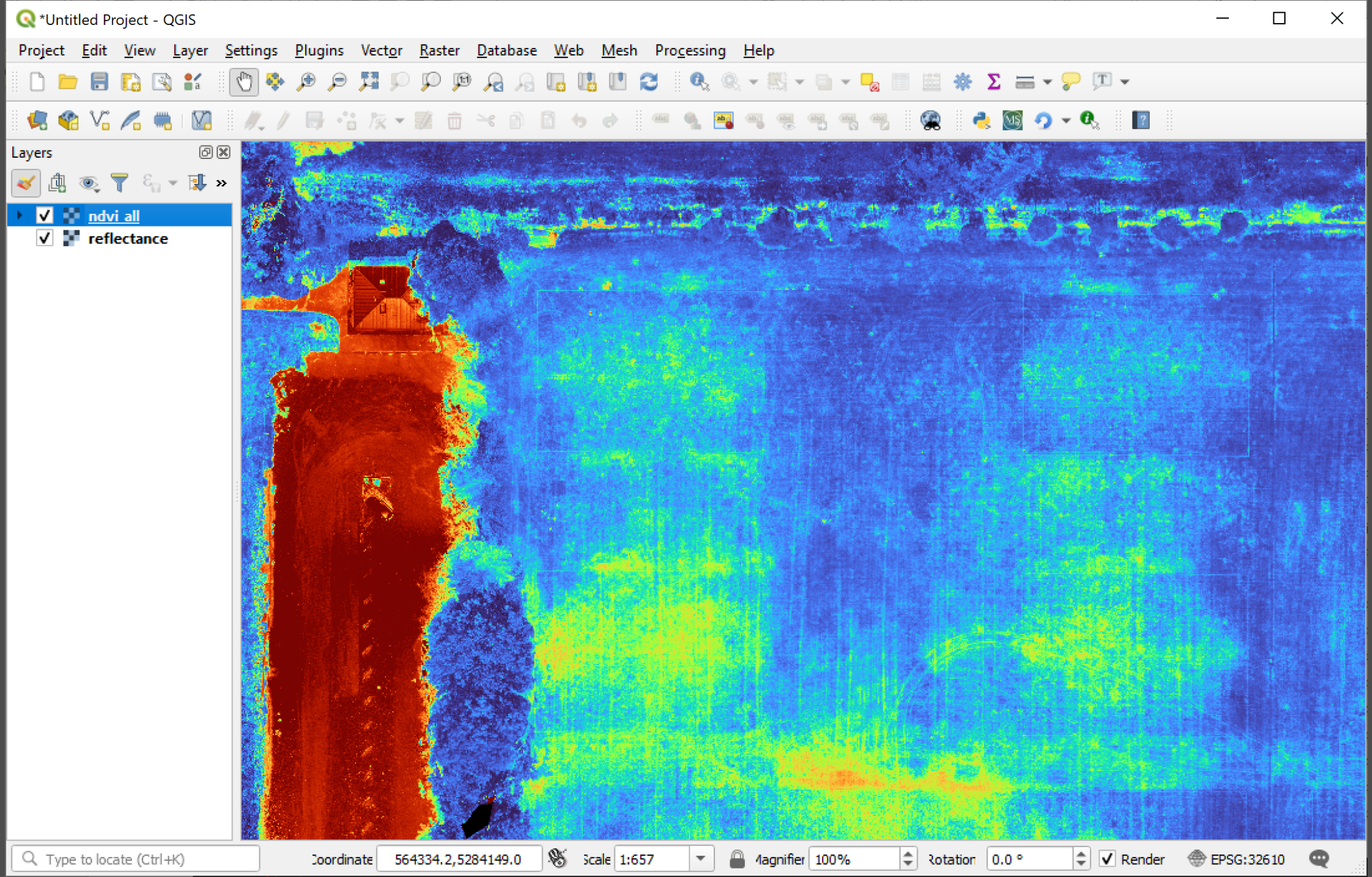
```
<VRTDataset rasterXSize="5552" rasterYSize="5986">
  <SRS>PROJCS["WGS 84 / UTM zone 10N",GEOGCS["WGS 84",DATUM["WGS_1984",SPHEROID["WGS 84",6378137,298.257223563,AUTHORITY
  <GeoTransform>  5.6430981462000008e+05,  4.5220000000000003e-02,  0.0000000000000000e+00,  5.2843024623000007e+06,  0.(
  <VRTRasterBand dataType="Float32" band="1" subClass="VRTDerivedRasterBand">
    <ComplexSource>
      <SourceFilename relativeToVRT="1" shared="0">reflectance.vrt</SourceFilename>
      <SourceBand>5</SourceBand>
      <SourceProperties RasterXSize="5552" RasterYSize="5986" DataType="Float32" BlockXSize="128" BlockYSize="128" />
      <SrcRect xOff="0" yOff="0" xSize="5552" ySize="5986" />
      <DstRect xOff="0" yOff="0" xSize="5552" ySize="5986" />
      <NODATA>-10000</NODATA>
    </ComplexSource>
    <ComplexSource>…
    </ComplexSource>
    <NoDataValue>-2</NoDataValue>
    <PixelFunctionLanguage>Python</PixelFunctionLanguage>
    <PixelFunctionType>ndvi_all</PixelFunctionType>
    <PixelFunctionCode>
      <![CDATA[
import numpy as np
def ndvi_all(in_ar, out_ar, xoff, yoff, xsize, ysize, raster_xsize, raster_ysize, buf_radius, gt, **kwargs):
    np.seterr(divide='ignore', invalid='ignore')
    num = np.subtract(in_ar[0].astype(float),in_ar[1].astype(float), dtype = float)
    den = np.add(in_ar[0].astype(float),in_ar[1].astype(float), dtype = float)
    np.divide( num, den, dtype = float, out = out_ar )
    out_ar[den == 0] = -2.0]]>
    </PixelFunctionCode>
    <SourceTransferType>Float32</SourceTransferType>
  </VRTRasterBand>
</VRTDataset>
```

```
<VRTDataset rasterXSize="5552" rasterYSize="5986">
  <SRS>PROJCS["WGS 84 / UTM zone 10N",GEOGCS["WGS 84",DATUM["WGS_1984",SPHEROID["WGS 84",6378137,298.257223563,AUTHORITY
  <GeoTransform>  5.6430981462000008e+05,  4.5220000000000003e-02,  0.0000000000000000e+00,  5.2843024623000007e+06,  0.(
  <VRTRasterBand dataType="Float32" band="1" subClass="VRTDerivedRasterBand">
    <ComplexSource>
      <SourceFilename relativeToVRT="1" shared="0">reflectance.vrt</SourceFilename>
      <SourceBand>5</SourceBand>
      <SourceProperties RasterXSize="5552" RasterYSize="5986" DataType="Float32" BlockXSize="128" BlockYSize="128" />
      <SrcRect xOff="0" yOff="0" xSize="5552" ySize="5986" />
      <DstRect xOff="0" yOff="0" xSize="5552" ySize="5986" />
      <NODATA>-10000</NODATA>
    </ComplexSource>
    <ComplexSource>…
    </ComplexSource>
    <NoDataValue>-2</NoDataValue>
    <PixelFunctionLanguage>Python</PixelFunctionLanguage>
    <PixelFunctionType>ndvi_all</PixelFunctionType>
    <PixelFunctionCode>
      <![CDATA[
import numpy as np
def ndvi_all(in_ar, out_ar, xoff, yoff, xsize, ysize, raster_xsize, raster_ysize, buf_radius, gt, **kwargs):
    np.seterr(divide='ignore', invalid='ignore')
    num = np.subtract(in_ar[0].astype(float),in_ar[1].astype(float), dtype = float)
    den = np.add(in_ar[0].astype(float),in_ar[1].astype(float), dtype = float)
    np.divide( num, den, dtype = float, out = out_ar )
    out_ar[den == 0] = -2.0]]>
    </PixelFunctionCode>
    <SourceTransferType>Float32</SourceTransferType>
  </VRTRasterBand>
</VRTDataset>
```

```
<VRTDataset rasterXSize="5552" rasterYSize="5986">
  <SRS>PROJCS["WGS 84 / UTM zone 10N",GEOGCS["WGS 84",DATUM["WGS_1984",SPHEROID["WGS 84",6378137,298.257223563,AUTHORITY
  <GeoTransform>  5.6430981462000008e+05,  4.5220000000000003e-02,  0.0000000000000000e+00,  5.2843024623000007e+06,  0.(
  <VRTRasterBand dataType="Float32" band="1" subClass="VRTDerivedRasterBand">
    <ComplexSource>
      <SourceFilename relativeToVRT="1" shared="0">reflectance.vrt</SourceFilename>
      <SourceBand>5</SourceBand>
      <SourceProperties RasterXSize="5552" RasterYSize="5986" DataType="Float32" BlockXSize="128" BlockYSize="128" />
      <SrcRect xOff="0" yOff="0" xSize="5552" ySize="5986" />
      <DstRect xOff="0" yOff="0" xSize="5552" ySize="5986" />
      <NODATA>-10000</NODATA>
    </ComplexSource>
    <ComplexSource>…
    </ComplexSource>
    <NoDataValue>-2</NoDataValue>
    <PixelFunctionLanguage>Python</PixelFunctionLanguage>
    <PixelFunctionType>ndvi_all</PixelFunctionType>
    <PixelFunctionCode>
      <![CDATA[
import numpy as np
def ndvi_all(in_ar, out_ar, xoff, yoff, xsize, ysize, raster_xsize, raster_ysize, buf_radius, gt, **kwargs):
    np.seterr(divide='ignore', invalid='ignore')
    num = np.subtract(in_ar[0].astype(float),in_ar[1].astype(float), dtype = float)
    den = np.add(in_ar[0].astype(float),in_ar[1].astype(float), dtype = float)
    np.divide( num, den, dtype = float, out = out_ar )
    out_ar[den == 0] = -2.0]]>
    </PixelFunctionCode>
    <SourceTransferType>Float32</SourceTransferType>
  </VRTRasterBand>
</VRTDataset>
```

```
<VRTDataset rasterXSize="5552" rasterYSize="5986">
  <SRS>PROJCS["WGS 84 / UTM zone 10N",GEOGCS["WGS 84",DATUM["WGS_1984",SPHEROID["WGS 84",6378137,298.257223563,AUTHORITY
  <GeoTransform>  5.6430981462000008e+05,  4.5220000000000003e-02,  0.0000000000000000e+00,  5.2843024623000007e+06,  0.(
  <VRTRasterBand dataType="Float32" band="1" subClass="VRTDerivedRasterBand">
    <ComplexSource>
      <SourceFilename relativeToVRT="1" shared="0">reflectance.vrt</SourceFilename>
      <SourceBand>5</SourceBand>
      <SourceProperties RasterXSize="5552" RasterYSize="5986" DataType="Float32" BlockXSize="128" BlockYSize="128" />
      <SrcRect xOff="0" yOff="0" xSize="5552" ySize="5986" />
      <DstRect xOff="0" yOff="0" xSize="5552" ySize="5986" />
      <NODATA>-10000</NODATA>
    </ComplexSource>
    <ComplexSource>…
    </ComplexSource>
    <NoDataValue>-2</NoDataValue>
    <PixelFunctionLanguage>Python</PixelFunctionLanguage>
    <PixelFunctionType>ndvi_all</PixelFunctionType>
    <PixelFunctionCode>
      <![CDATA[
import numpy as np
def ndvi_all(in_ar, out_ar, xoff, yoff, xsize, ysize, raster_xsize, raster_ysize, buf_radius, gt, **kwargs):
    np.seterr(divide='ignore', invalid='ignore')
    num = np.subtract(in_ar[0].astype(float),in_ar[1].astype(float), dtype = float)
    den = np.add(in_ar[0].astype(float),in_ar[1].astype(float), dtype = float)
    np.divide( num, den, dtype = float, out = out_ar )
    out_ar[den == 0] = -2.0]]>
    </PixelFunctionCode>
    <SourceTransferType>Float32</SourceTransferType>
  </VRTRasterBand>
</VRTDataset>
```

Results


*Untitled Project - QGIS
Project Edit View Layer Settings Plugins Vector Raster Database Web Mesh Processing Help
Layers
ndvi_all
reflectance
Type to locate (Ctrl+K)
Coordinate 564334.2,5284149.0
Scale 1:657
Magnifier 100%
Rotation 0.0 °
Render
EPSG:32610

Tips & Tricks

- Set GDAL_VRT_ENABLE_PYTHON = YES in your environment!
- Use np.add / np.subtract / etc.
  - seem faster efficient than carrying a lot of intermediates
  - Watch out for your data types if different than float
- For more complex work, call your own library functions
- Use gdalinfo to find errors - QGIS just doesn’t display buggy VRTs
- Chaining complex VRTs can make for easy debug

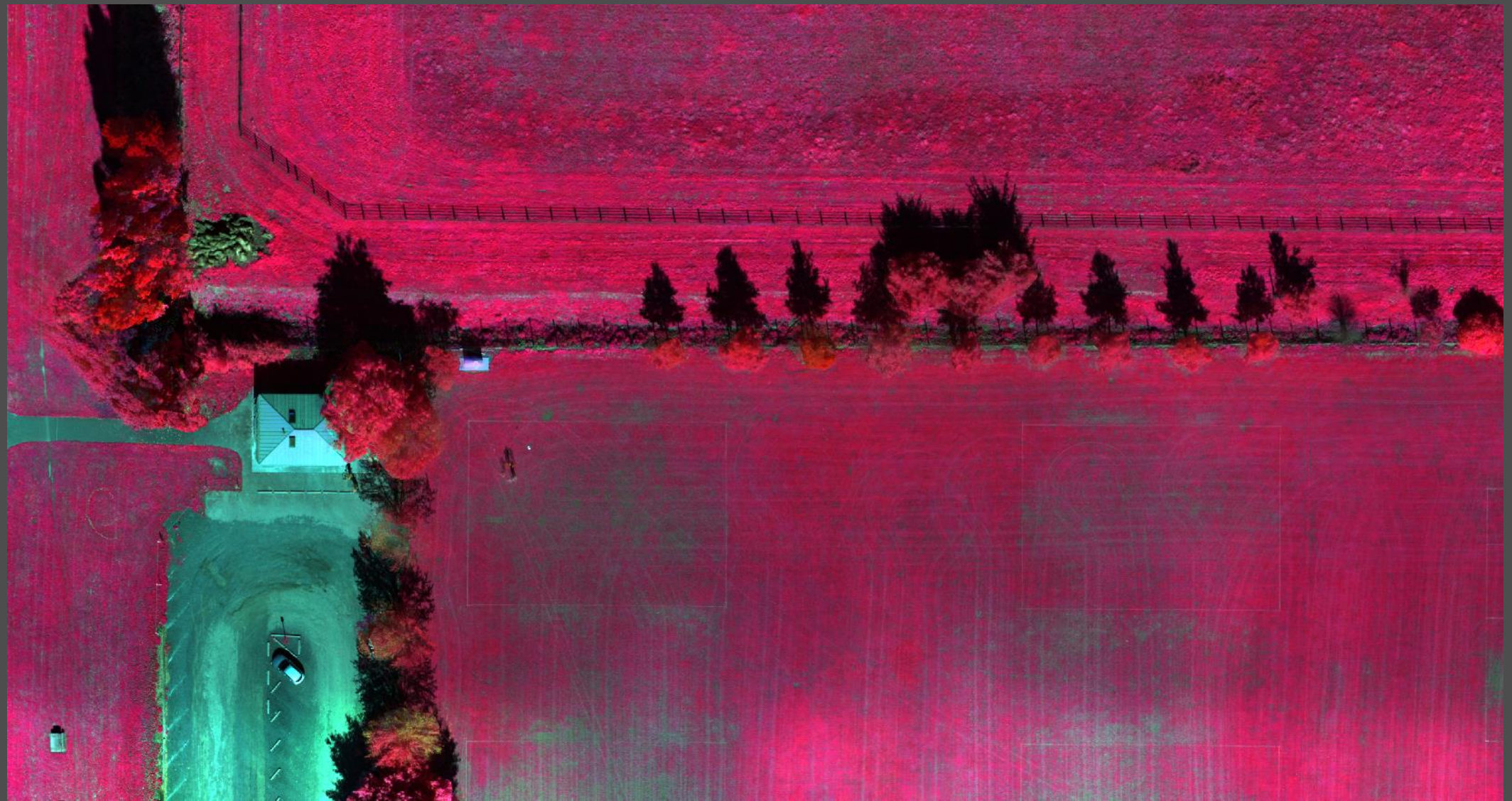
Examples

RGB Composite (scaling+gamma correction)

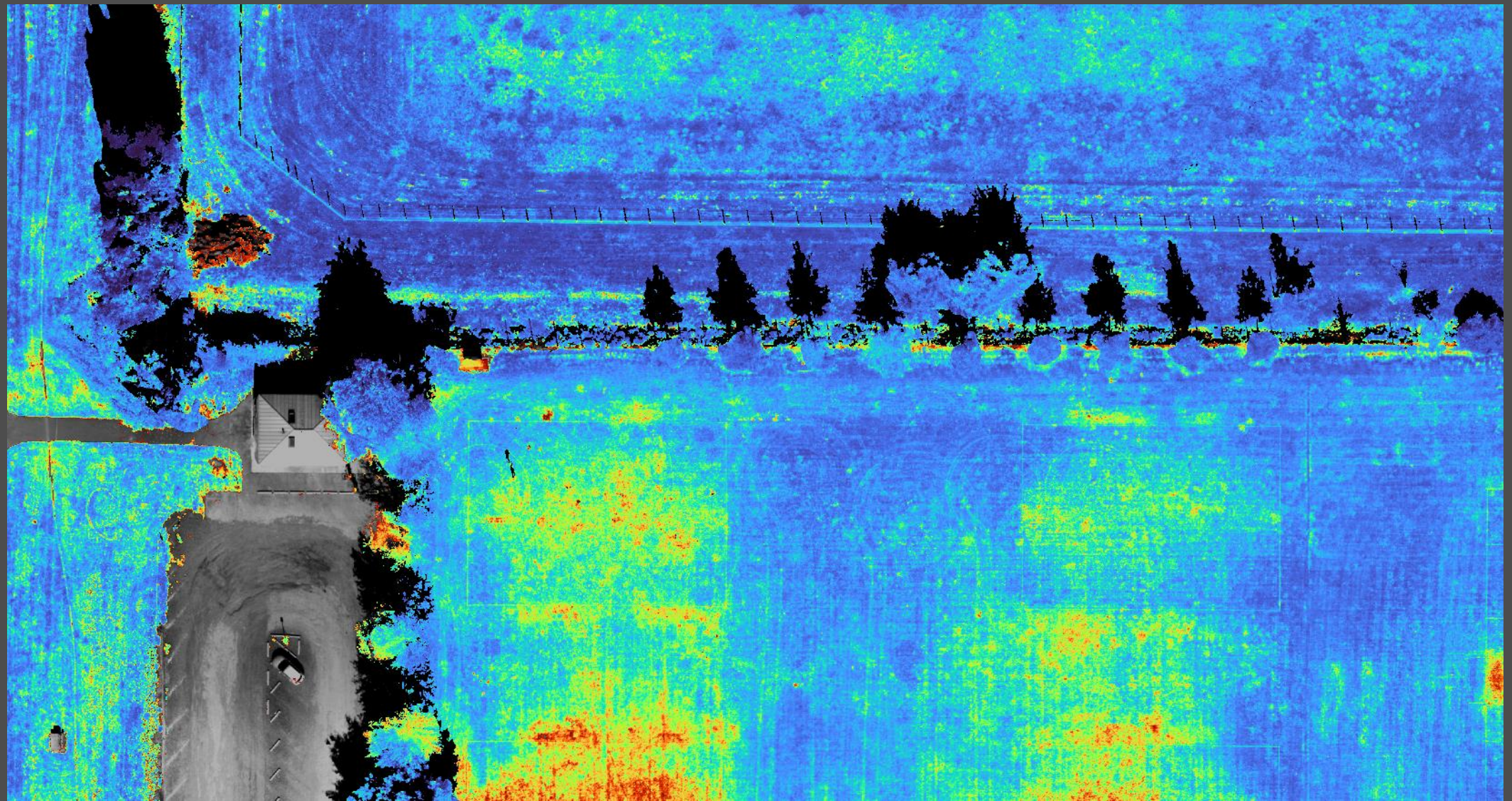
```
<VRTDataset rasterXSize="5552" rasterYSize="5986">
  <SRS>PROJCS["WGS 84 / UTM zone 10N",GEOGCS["WGS 84",DATUM["WGS_1984",SPHEROID["WGS 84",6378137,298.257223563,AUTHORIT
  <GeoTransform>  5.6430981462000008e+05,  4.5220000000000003e-02,  0.0000000000000000e+00,  5.2843024623000007e+06,  0
  <VRTRasterBand dataType="Byte" band="1">
    <NoDataValue>0</NoDataValue>
    <ComplexSource>
      <SourceFilename relativeToVRT="1" shared="0">reflectance.vrt</SourceFilename>
      <SourceBand>3</SourceBand>
      <SourceProperties RasterXSize="5552" RasterYSize="5986" DataType="Float32" BlockXSize="128" BlockYSize="128" />
      <SrcRect xOff="0" yOff="0" xSize="5552" ySize="5986" />
      <DstRect xOff="0" yOff="0" xSize="5552" ySize="5986" />
      <NODATA>-10000</NODATA>
      <Exponent>0.75</Exponent>
      <SrcMin>0</SrcMin>
      <SrcMax>0.35</SrcMax>
      <DstMin>0</DstMin>
      <DstMax>255</DstMax>
    </ComplexSource>
  </VRTRasterBand>
  <VRTRasterBand dataType="Byte" band="2">…
  </VRTRasterBand>
  <VRTRasterBand dataType="Byte" band="3">…
  </VRTRasterBand>
</VRTDataset>
```

RGB Composite (with gamma)

CIR Composite



NDVI filtered to sunlit plants only



Summary

- VRTs are small, portable files
- See docs for other types of VRT I didn’t disucss
- Check out compiled pixel functions for efficiency

- Documentation
  - https://gdal.org/drivers/raster/vrt.html
  - That seems to be all. ☹

justinm@justinm.com

Twitter @k5em

Questions?