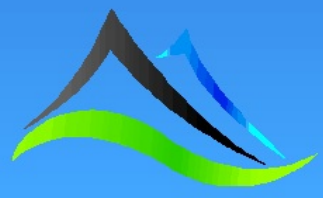
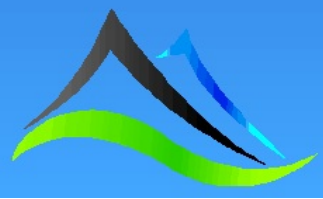
A robust filter for noisy data using stacked splines

ExploreGeo Technical Note 9



Problem

Field data, particularly from ground surveys, can contain a lot of noise. This can take the form of low amplitude chatter through to large amplitude spikes which may occur in clusters. The different types of noise impose challenges for common filters and often a combination of filters are required to smooth the data.



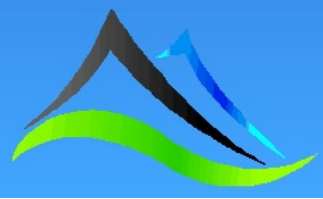
An example - Ground magnetic data containing both spikes and chatter

Problem

Existing filters can do a good job of removing the spikes (e.g median, non-linear) or smoothing the chatter (e.g. butterworth, moving average) but it is difficult to do both well and thus a combination of filters are often used to smooth data such as these.

Filters always destroy some data so we need to use them as lightly as possible.

What if we could filter both the chatter and the spikes using a single filter and still retain most of the data?

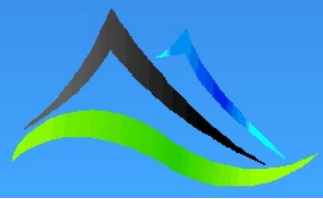


A solution

We create an array of datasets, each a subsampled version of the original, each with a similar sample interval but each offset from the other. The sample interval would ideally be the same but field data is rarely collected at equal increments.

For example we might create five datasets (D#) from the original (O).

D1=O(1), O(6), O(11)...
D2=O(2), O(7), O(12)...
D3=O(3), O(8), O(13)...
D4=O(4), O(9), O(15)...
D5=O(5), O(10), O(16)...

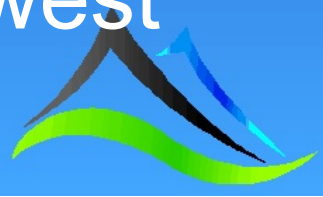


The Solution .. cont

Each subset is then splined back to the same spacing as the original data set to produce a series of stacked splines aliasing the data.

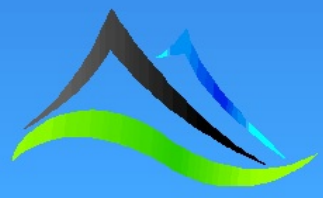
A new filtered data set can be generated from these splines by combining them. This could be done with a sorted weighted average, median or using statistical weights to exclude outliers.

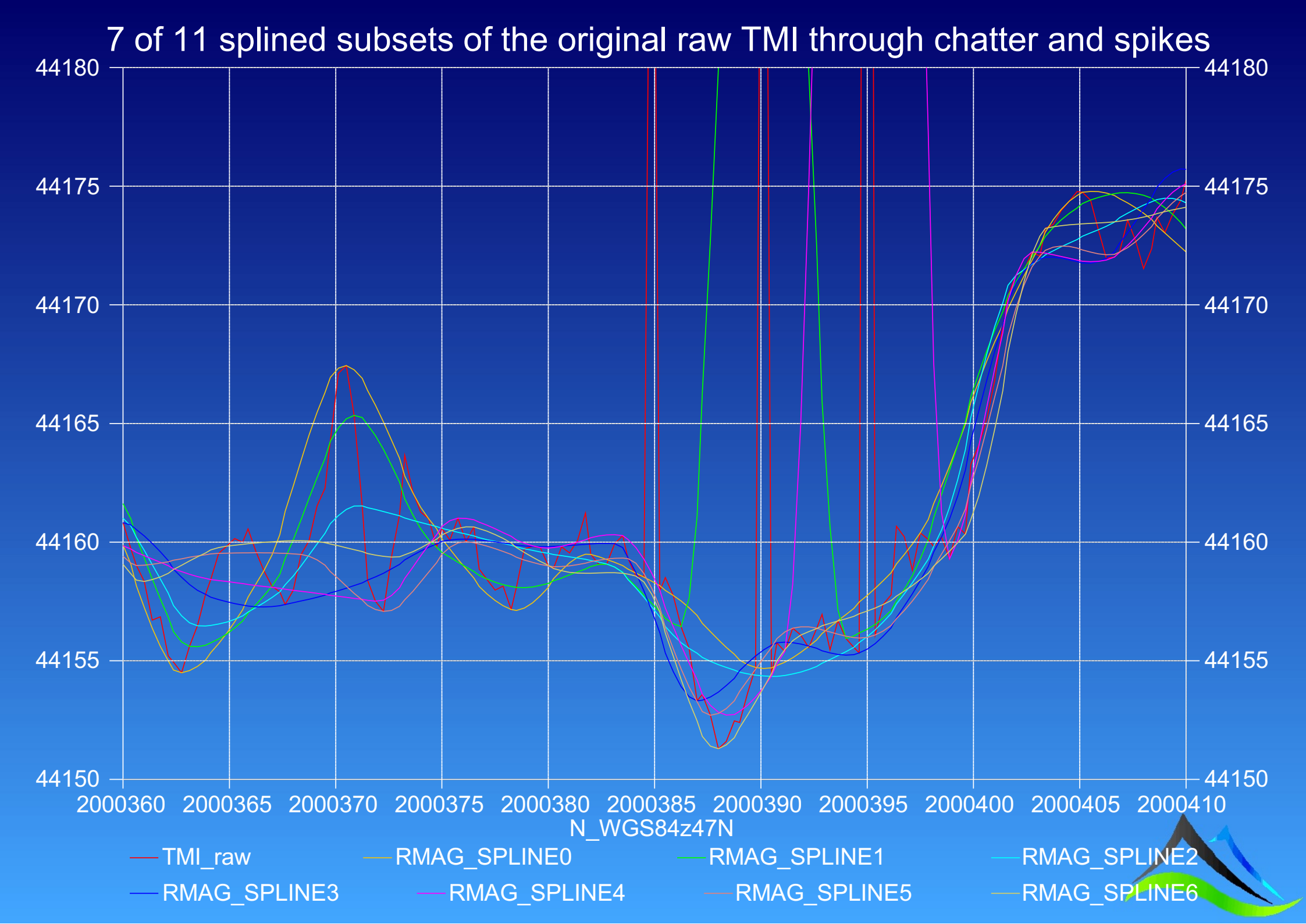
In the implementation described here the splined values for each point are sorted and an average is taken which excludes the highest and lowest value.

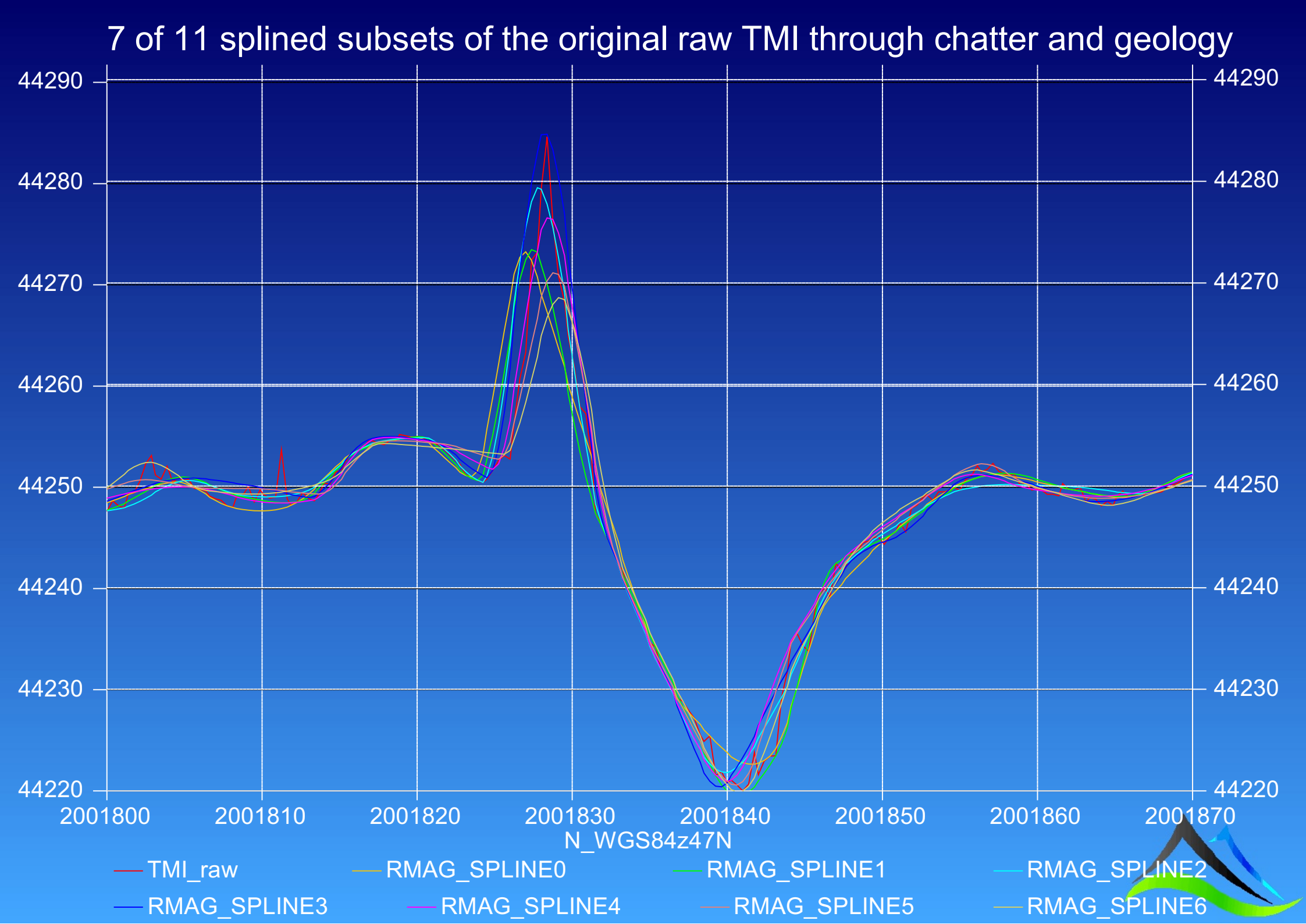


The stacked spline filter has been implemented as a spatial filter in DBaseO and the following slides show comparisons between it and other spatial filters in DBaseO.

Using low pass filters in the frequency domain should produce a similar result to the Butterworth filter



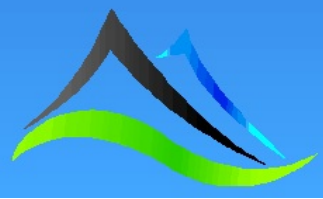

7 of 11 splined subsets of the original raw TMI through chatter and spikes
44180
44175
44170
44165
44160
44155
44150
2000360
2000365
2000370
2000375
2000380
2000385
2000390
2000395
2000400
2000405
2000410
N_WGS84z47N
TMI_raw
RMAG_SPLINE0
RMAG_SPLINE1
RMAG_SPLINE2
RMAG_SPLINE3
RMAG_SPLINE4
RMAG_SPLINE5
RMAG_SPLINE6


7 of 11 splined subsets of the original raw TMI through chatter and geology
44290
44280
44270
44260
44250
44240
44230
44220
2001800
2001810
2001820
2001830
2001840
2001850
2001860
2001870
N_WGS84z47N
TMI_raw
RMAG_SPLINE0
RMAG_SPLINE1
RMAG_SPLINE2
RMAG_SPLINE3
RMAG_SPLINE4
RMAG_SPLINE5
RMAG_SPLINE6

Comparison with existing filters

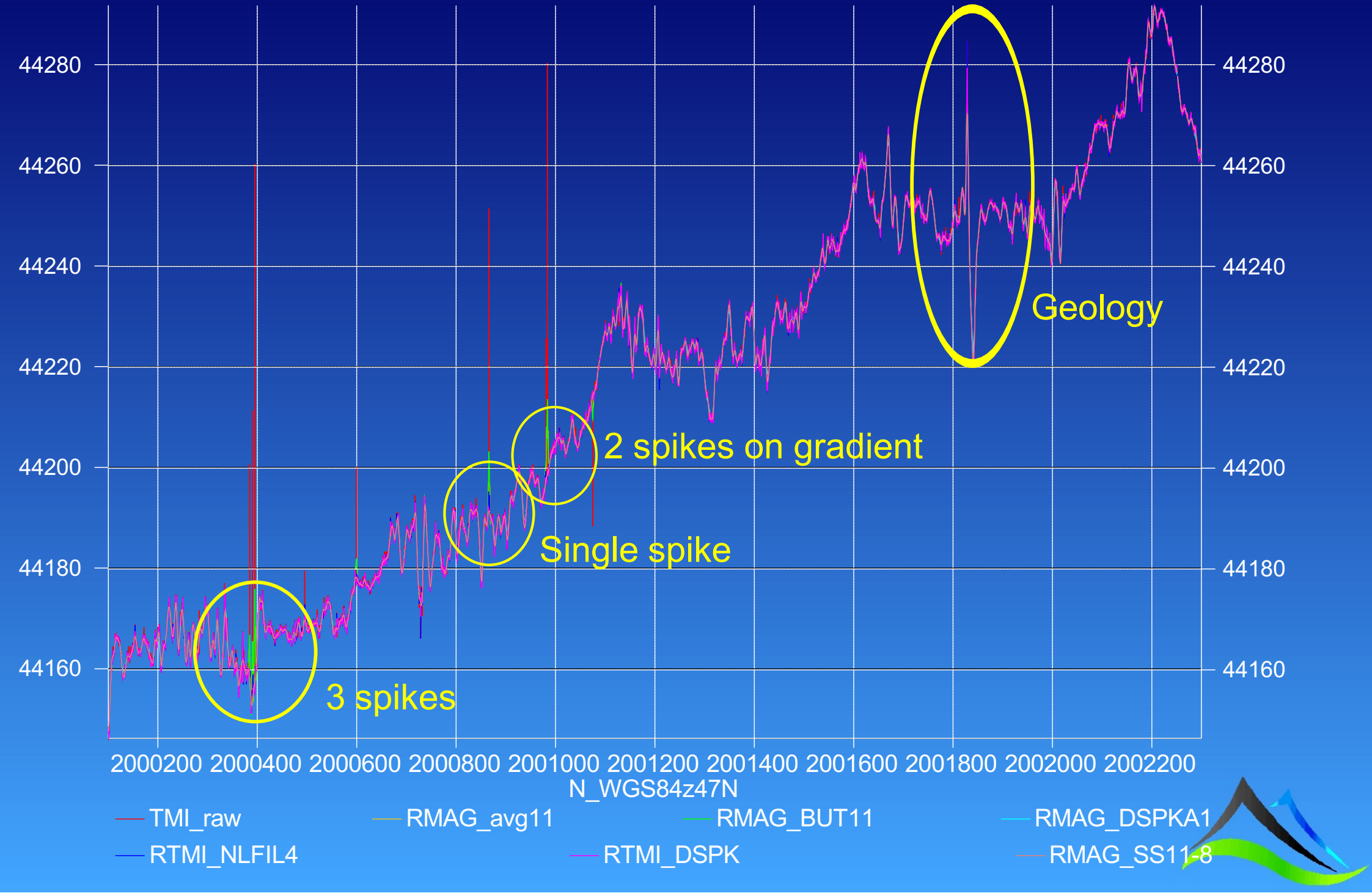
All filters us the same kernel size of 3.85m or 11 samples

- Avg11 - unweighted moving average
- BUT11 - Butterworth filter
- DSPKA1 - Median followed by an unweighted moving average
- NLFILT4 - Non-linear filter (Naudy & Dreyer, 1968)
- DSPK - Iterative smoothing with a cubic spline. (Guerra and Tapia 1974)
- SS11-8 - Stacked spline filter



All filters compared - all have reduced the intensity of the spikes.

Zoom in to the areas circled below




Filter Comparison in area of 3 spikes
44180
44175
44170
44165
44160
44155
44150
2000360
2000365
2000370
2000375
2000380
2000385
2000390
2000395
2000400
2000405
2000410
N_WGS84z47N
TMI_raw
RMAG_avg11
RMAG_BUT11
RMAG_DSPKA1
RTMI_NLFIL4
RTMI_DSPK
RMAG_SS11-8

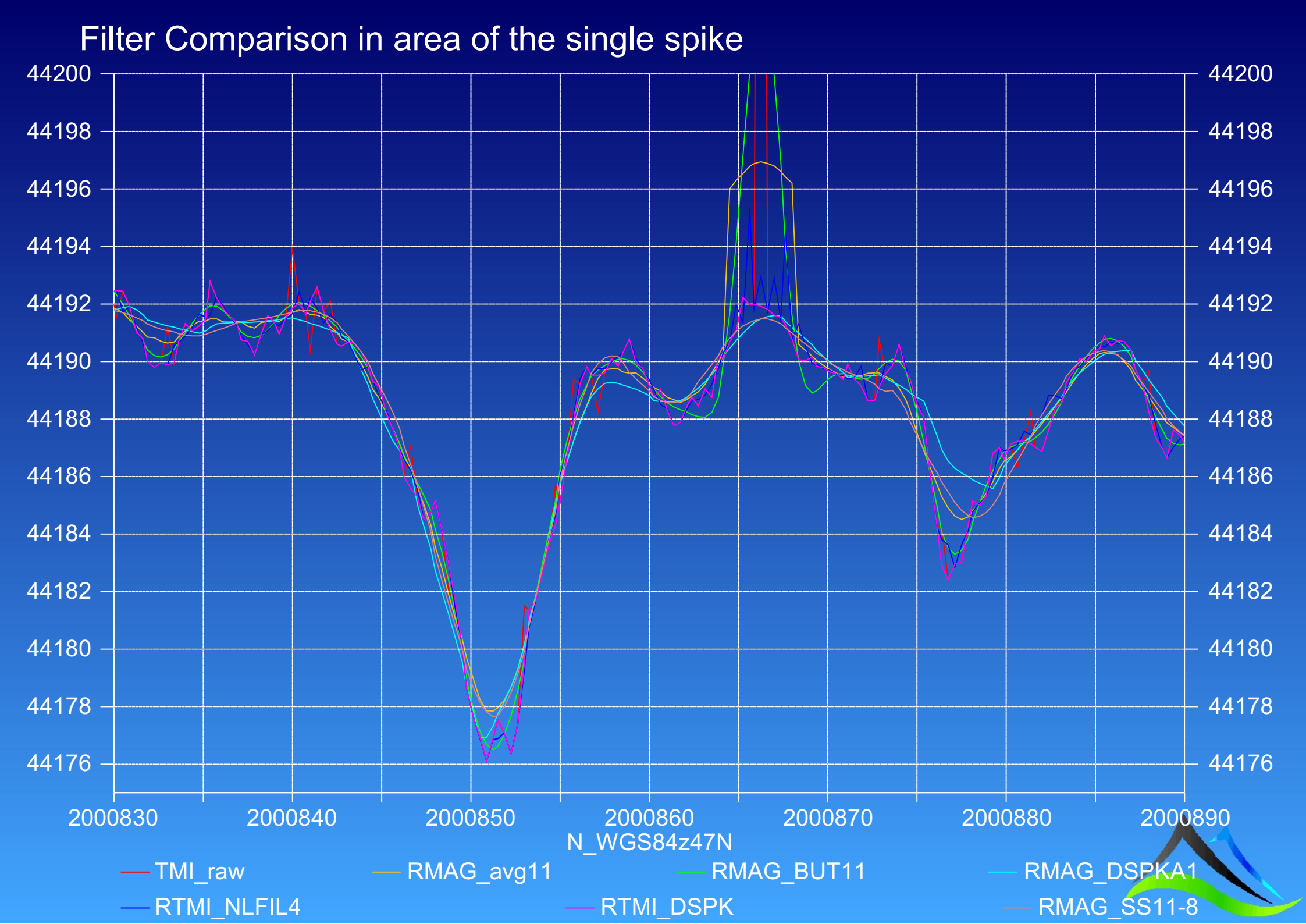
Observations

3 spikes

The Butterworth and moving average filters do a poor job of removing the high amplitude spikes. Weighting of the moving average would taper the plateaus around the spikes but still not remove them.

The central spike is leaking into the non-linear filter introducing low amplitude noise (chatter).

The stacked spline and the median followed by moving average filters appear to be doing the best job although the latter is not smooth at this scale.


Filter Comparison in area of the single spike
44200
44198
44196
44194
44192
44190
44188
44186
44184
44182
44180
44178
44176
2000830
2000840
2000850
2000860
2000870
2000880
2000890
N_WGS84z47N
TMI_raw
RMAG_avg11
RMAG_BUT11
RMAG_DSPKA1
RTMI_NLFIL4
RTMI_DSPK
RMAG_SS11-8

Observations

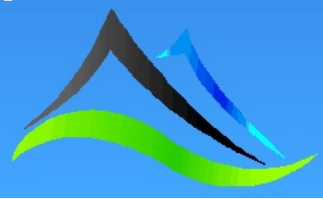
Single spike

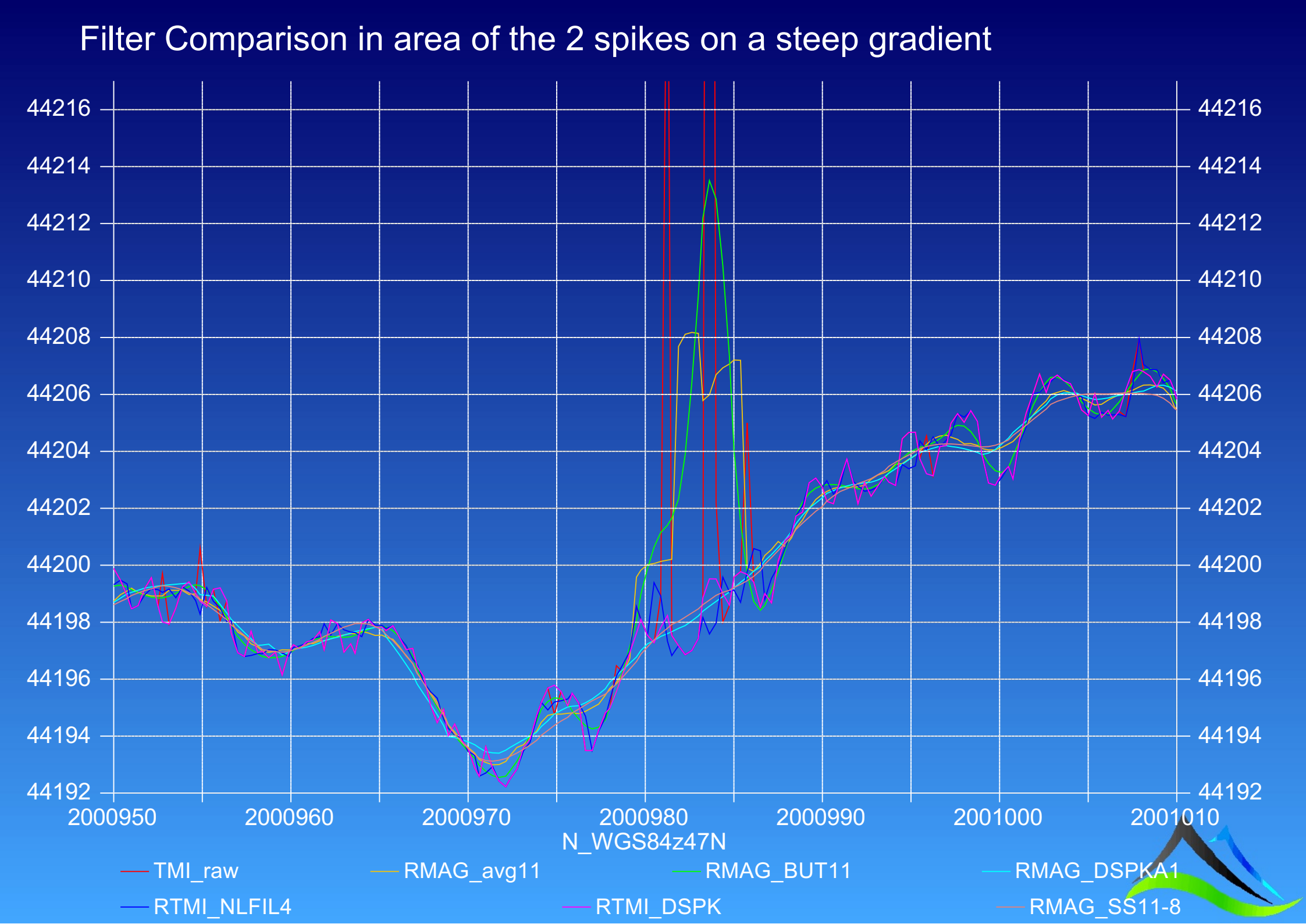
The Butterworth and moving average filters do a poor job of removing the high amplitude spike.

The spike is leaking into the non-linear filter introducing low amplitude noise (chatter).

The median followed by moving average filter is introducing significant phase shift in places.

The stacked spline filter appears to be doing the best job although there is some small (~1m) phase shift in places.




Filter Comparison in area of the 2 spikes on a steep gradient
44216
44214
44212
44210
44208
44206
44204
44202
44200
44198
44196
44194
44192
2000950
2000960
2000970
2000980
2000990
2001000
2001010
N_WGS84z47N
TMI_raw
RMAG_avg11
RMAG_BUT11
RMAG_DSPKA1
RTMI_NLFIL4
RTMI_DSPK
RMAG_SS11-8

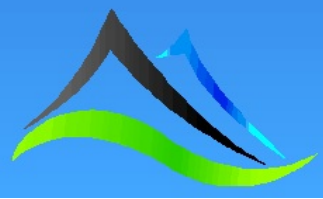
Observations

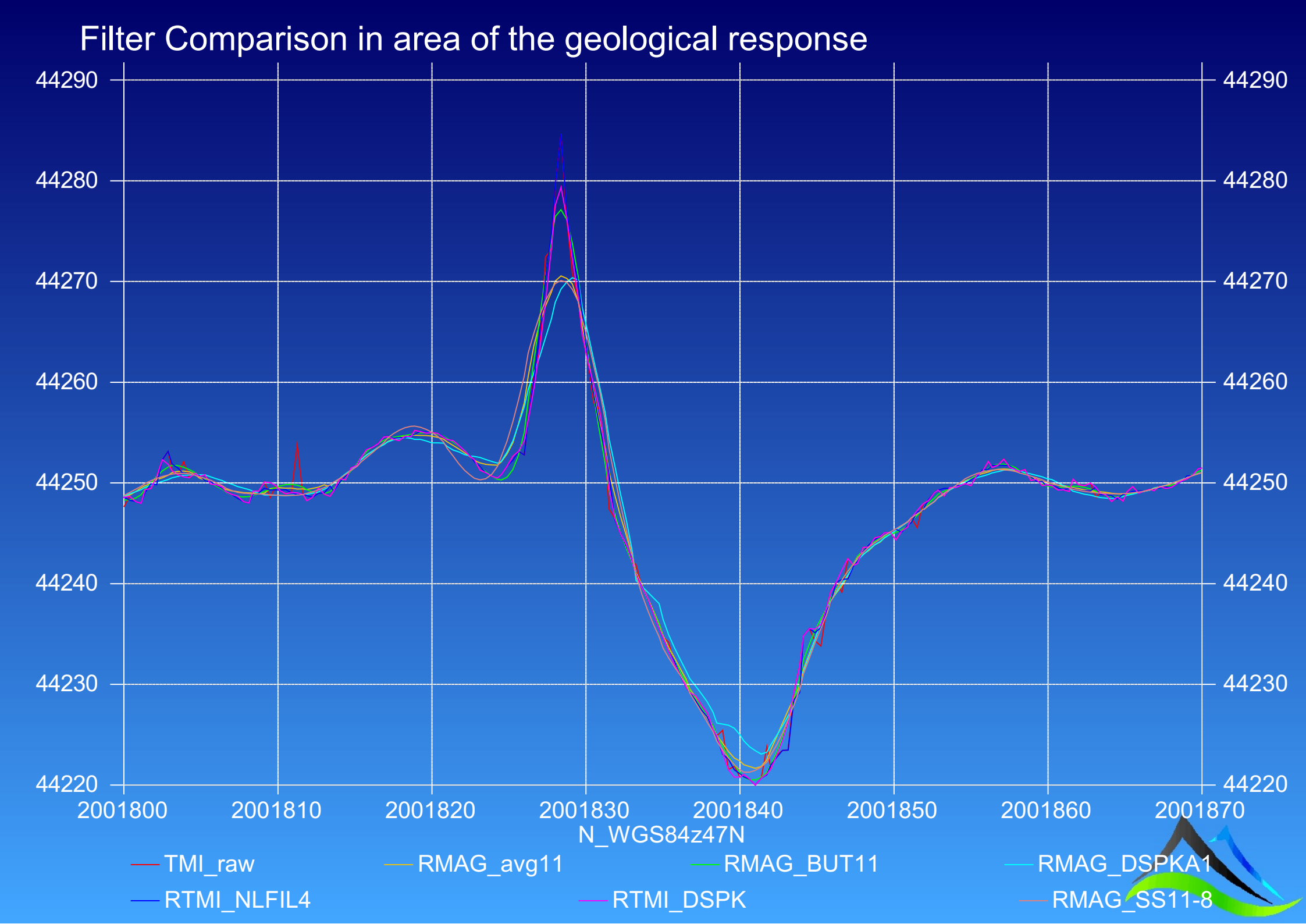
2 spikes on a steep gradient

The Butterworth and moving average filters do a poor job of removing the high amplitude spike.

The spike is leaking into the non-linear filter introducing low amplitude noise (chatter).

The stacked spline and median followed by moving average filters appears to be doing the best job.




Filter Comparison in area of the geological response
44290
44280
44270
44260
44250
44240
44230
44220
2001800
2001810
2001820
2001830
2001840
2001850
2001860
2001870
N_WGS84z47N
TMI_raw
RMAG_avg11
RMAG_BUT11
RMAG_DSPKA1
RTMI_NLFIL4
RTMI_DSPK
RMAG_SS11-8

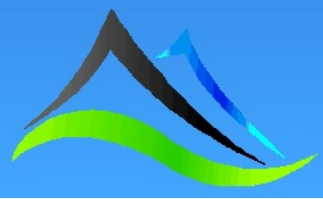
Observations

Geological response with minor chatter

All filters do a good job of smoothing the data.

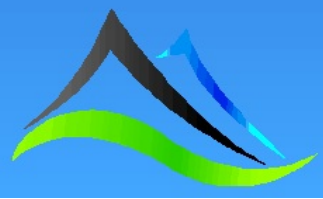
The moving average, stacked spline and median followed by moving average filters under estimate the amplitude of the anomaly maximum.

The stacked spline filter has broadened the anomaly peak by a small amount.



The previous slides had a lot of information which may not always have been easy to resolve.

The following slides just compare the stacked spline with a Butterworth filter. A second Butterworth filter at twice the wavelength has been computed and added to some of the plots for comparison.




Filter Comparison in area of 3 spikes
44180
44175
44170
44165
44160
44155
44150
2000360
2000365
2000370
2000375
2000380
2000385
2000390
2000395
2000400
2000405
2000410
N_WGS84z47N
TMI_raw
RMAG_BUT11
RMAG_SS11-8
RMAG_BUT21

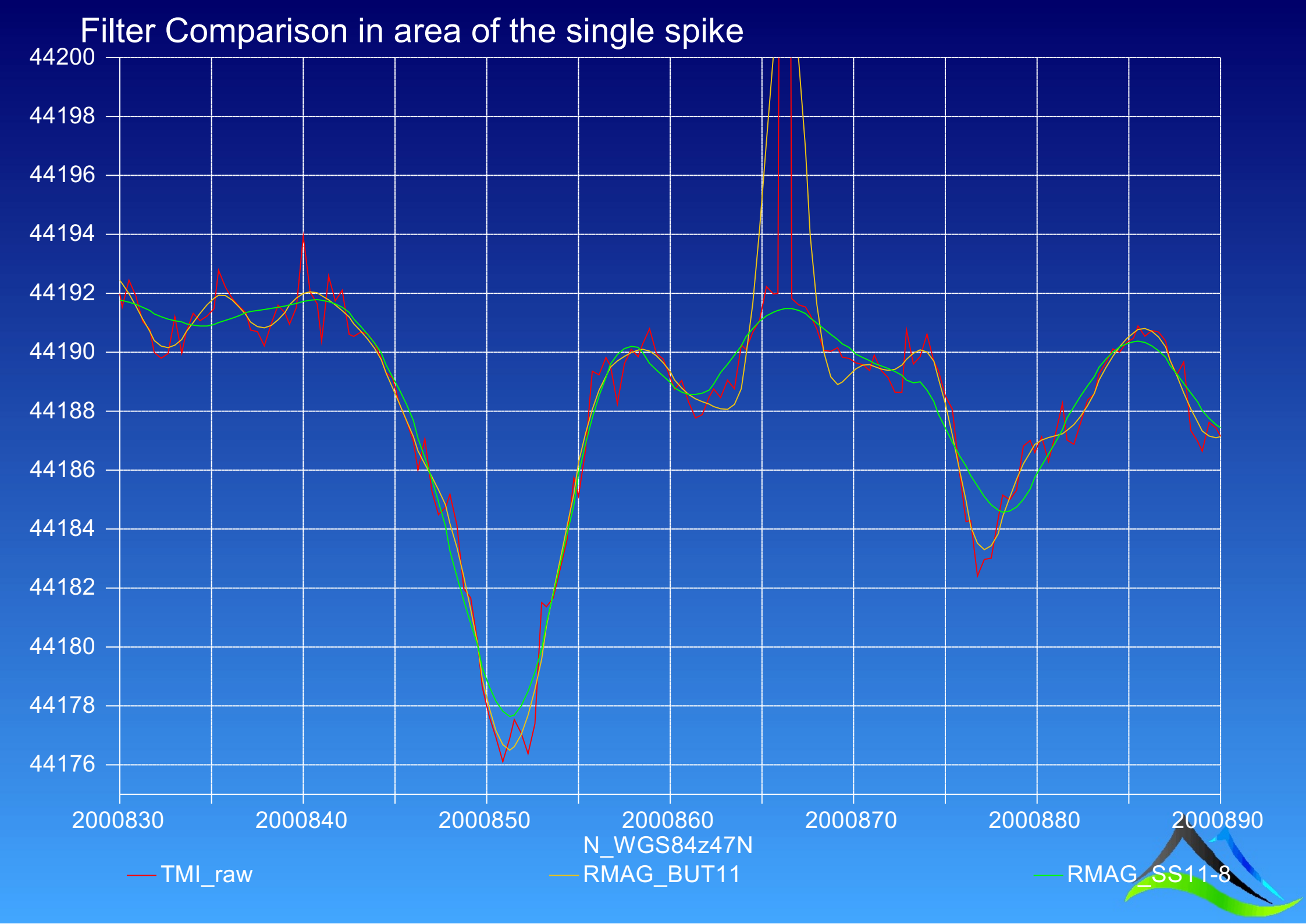
Observations

3 spikes

Increasing the wavelength of the Butterworth filter improves the suppression of the spikes but they are still evident.

Away from the spikes the longer wavelength Butterworth is very similar to the stacked spline filter.

The stacked spline filter is doing a better job.

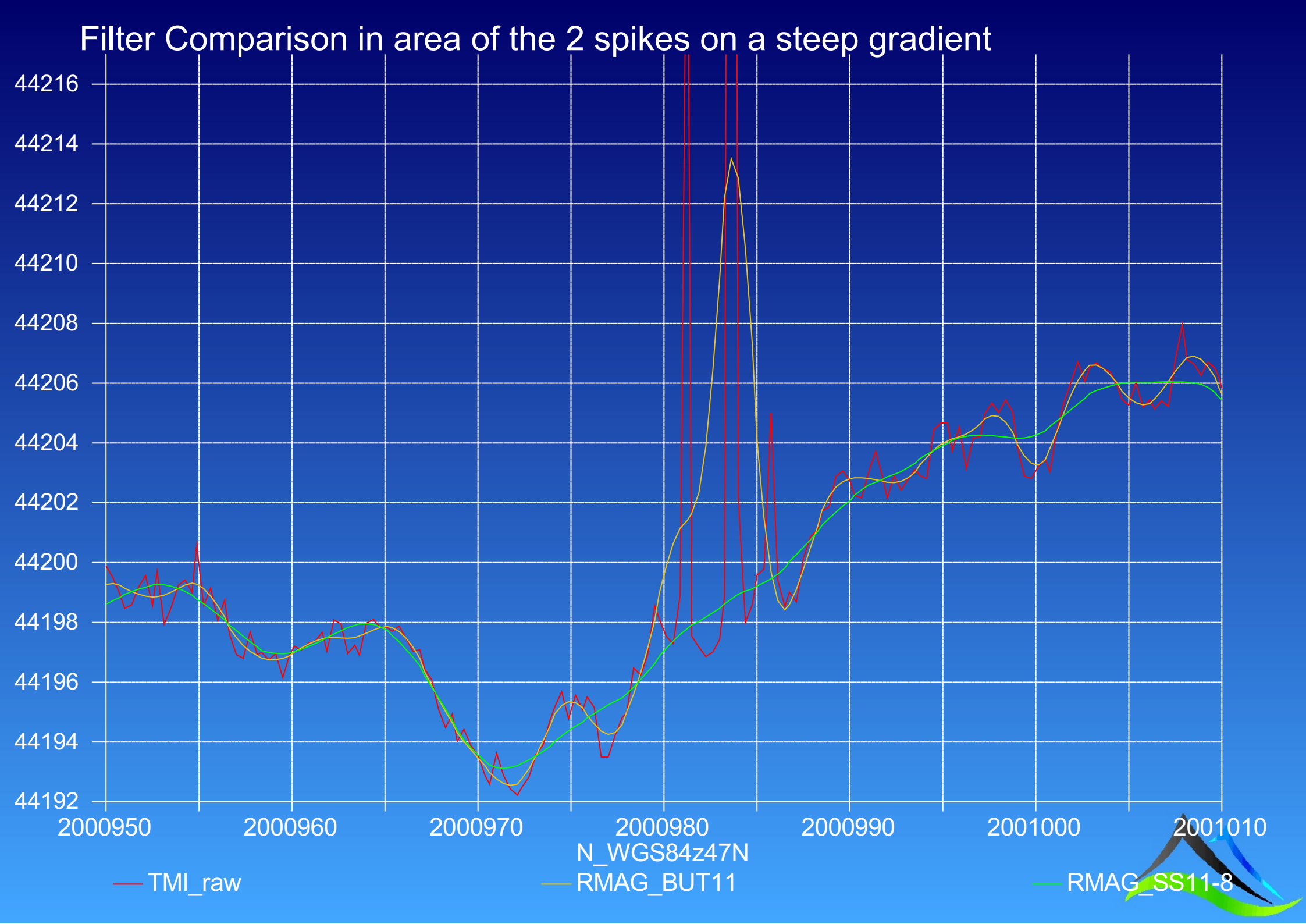

Filter Comparison in area of the single spike
44200
44198
44196
44194
44192
44190
44188
44186
44184
44182
44180
44178
44176
2000830
2000840
2000850
2000860
2000870
2000880
2000890
N_WGS84z47N
TMI_raw
RMAG_BUT11
RMAG_SS11-8

Observations

Single spike

The Butterworth filter is doing a poor job of suppressing the spike.

The stacked spline filter is doing a better job although it has introduced a small phase shift in the trough at ~2000877.


Filter Comparison in area of the 2 spikes on a steep gradient
44216
44214
44212
44210
44208
44206
44204
44202
44200
44198
44196
44194
44192
2000950
2000960
2000970
2000980
2000990
2001000
2001010
N_WGS84z47N
TMI_raw
RMAG_BUT11
RMAG_SS11-8

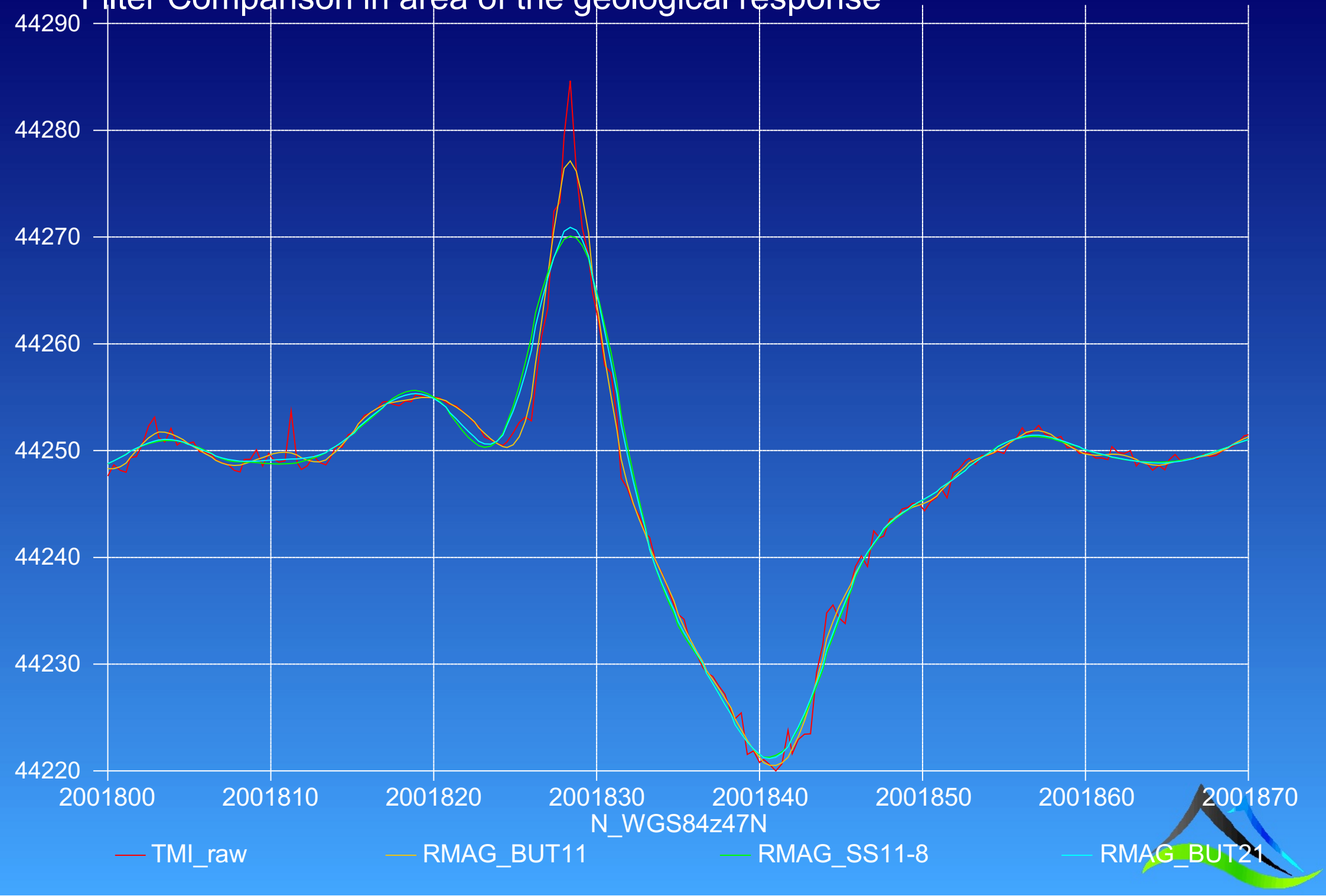
Observations

2 spikes on a gradient

The Butterworth filter is doing a poor job of suppressing the spike.

The stacked spline filter is doing a good job.

Filter Comparison in area of the geological response



Observations

Geological response with minor chatter

Both filters do a good job of suppressing chatter while honouring the geological response.

The longer wavelength Butterworth is very similar to the stacked spline filter.