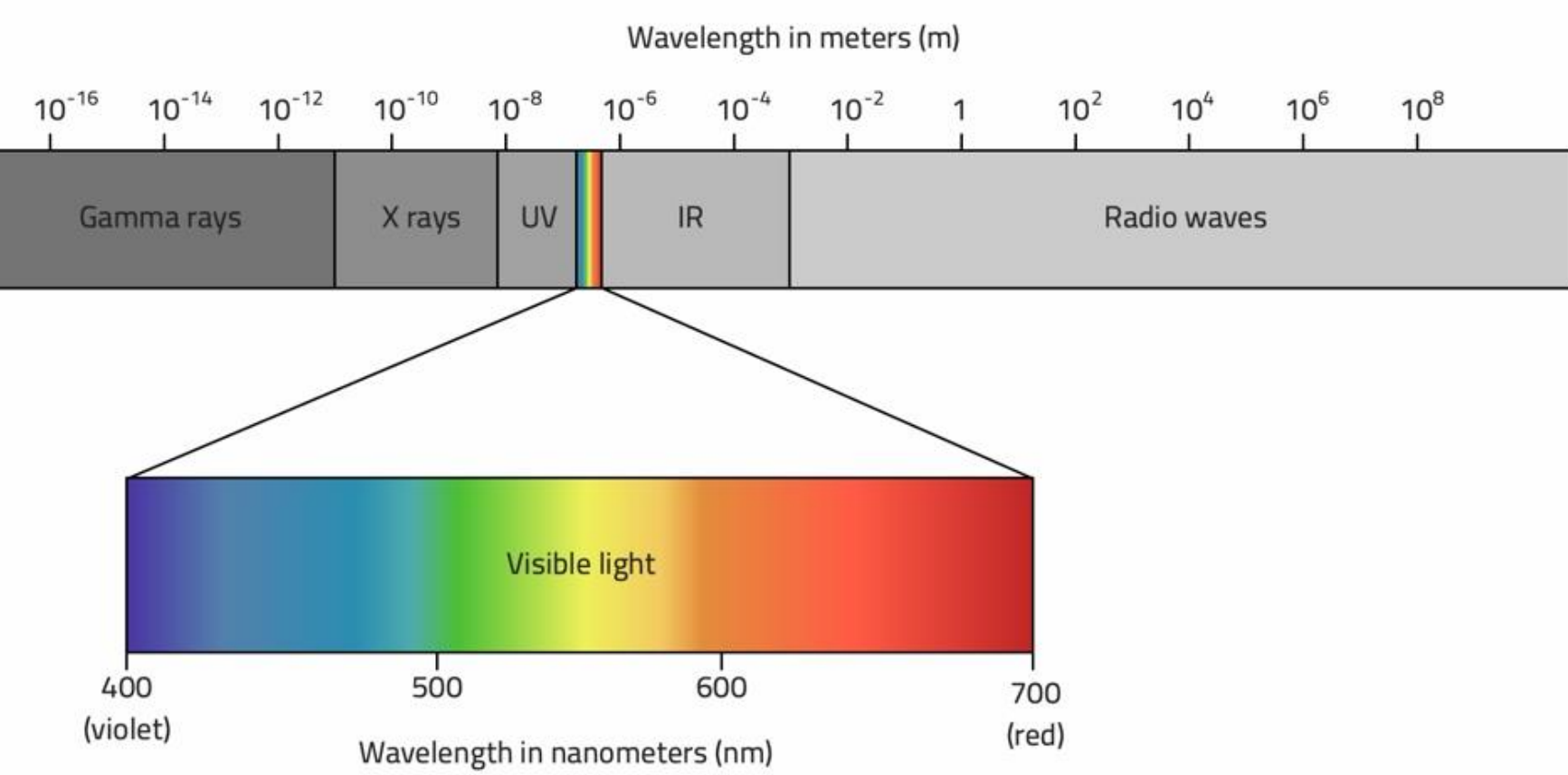


# Seminar Advanced Business Analytics

## Spectral Remote Sensing und Image Segmentation mit K-Means

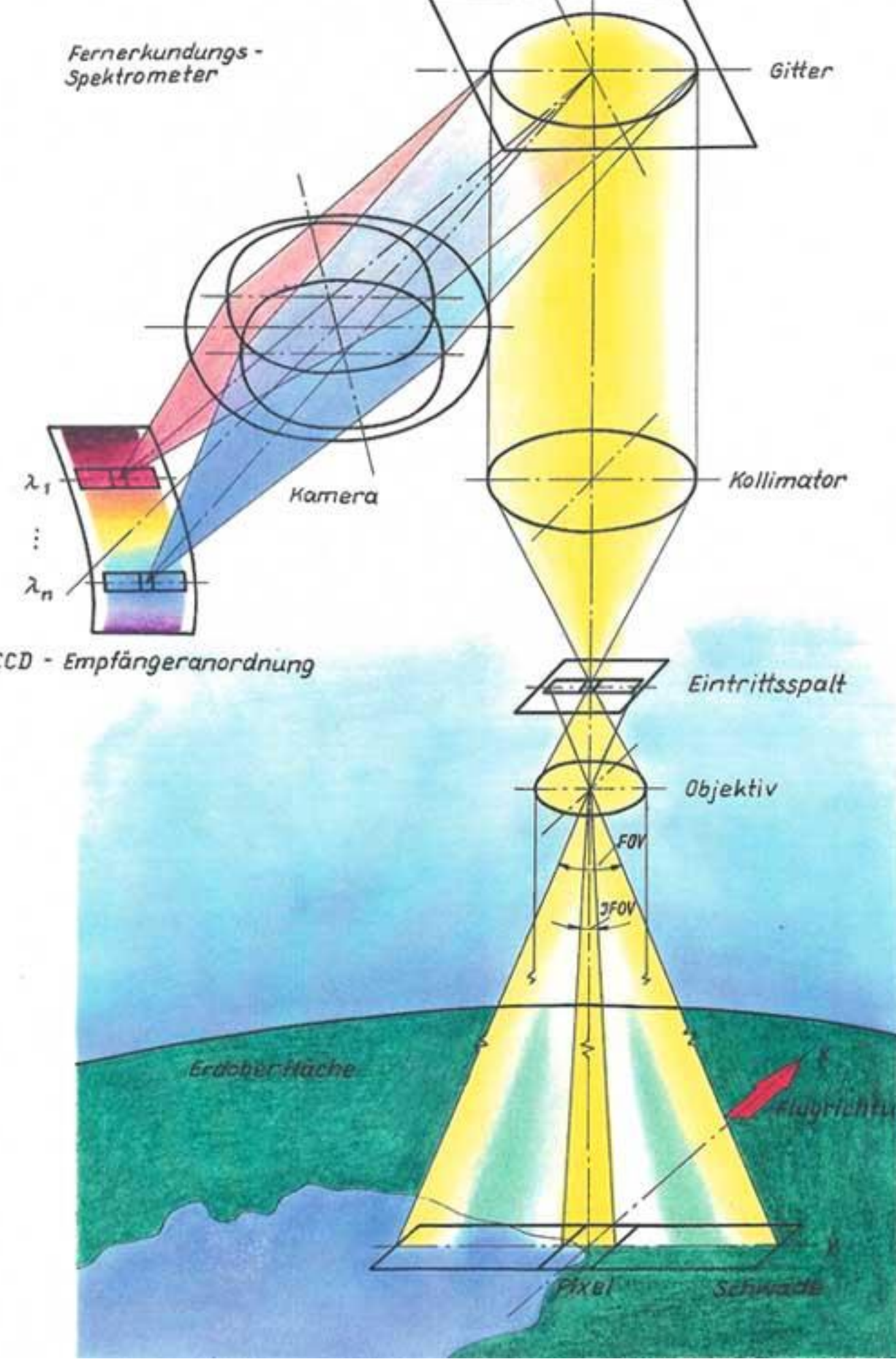
Alexander Hilberer  
Hochschule Trier

### Elektromagnetic Spectrum



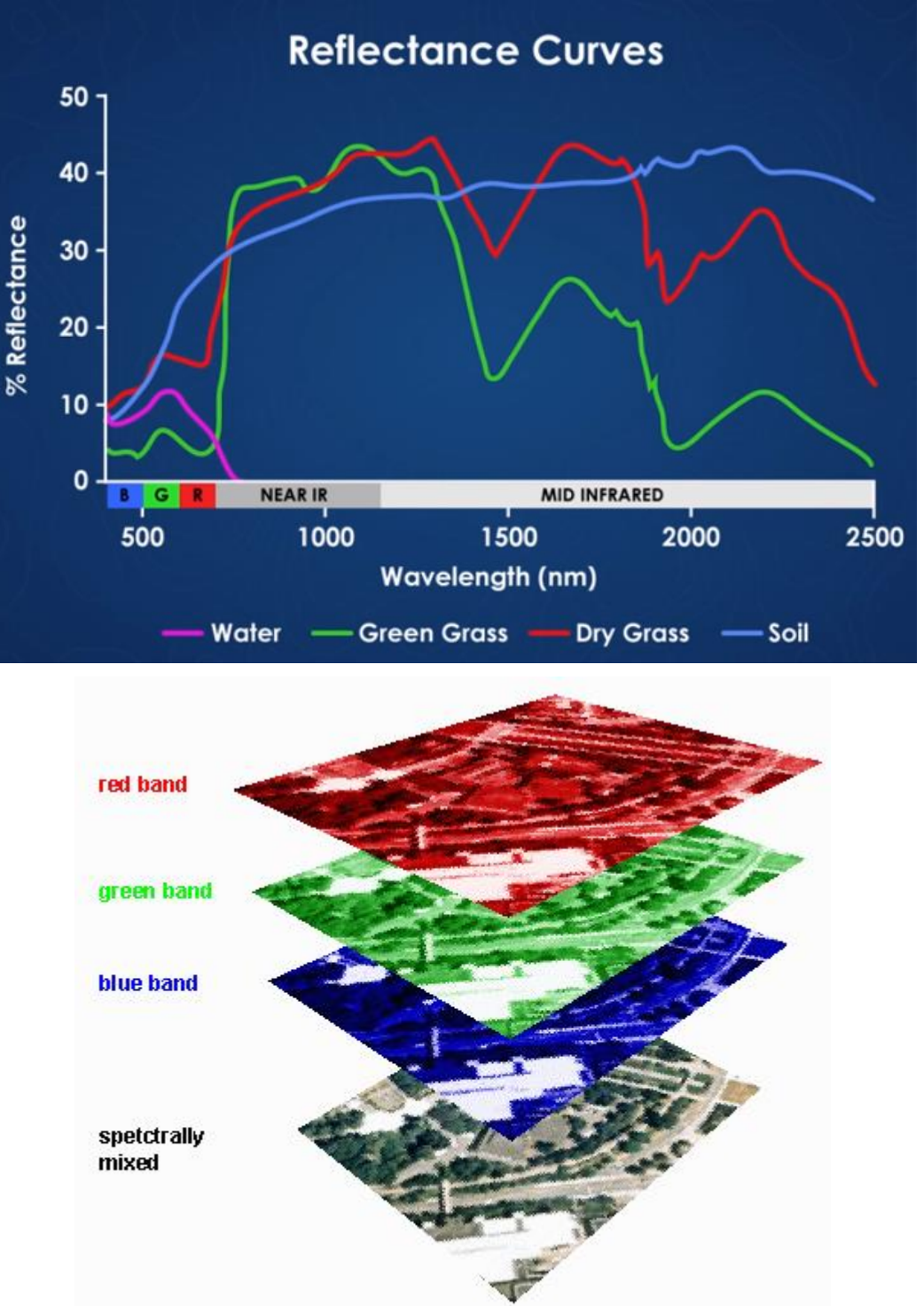
Das menschliche Auge sieht nur in Blau, Grün und Rot-tönen. Diese Lichtwellen haben eine Wellenlänge von 400 – 700nm. Alle Lichtwellen darüber hinaus sind nicht erkennbar für uns, können mittels Image Spectrometers jedoch gemessen werden.

### Remote Sensing



- Unter Remote Sensing versteht man die Eigenschaften eines Objektes zu messen, ohne dieses direct zu berühren.
- Dazu nutzt man so genannte Image Spectrometers.
- Diese spezielle Kamera ist dazu in der Lage, für das menschliche Auge nicht sichtbare Lichtwellen zu messen.
- Wird an Sateliten und Flugzeugen angebracht.

### Reflektions Kurven



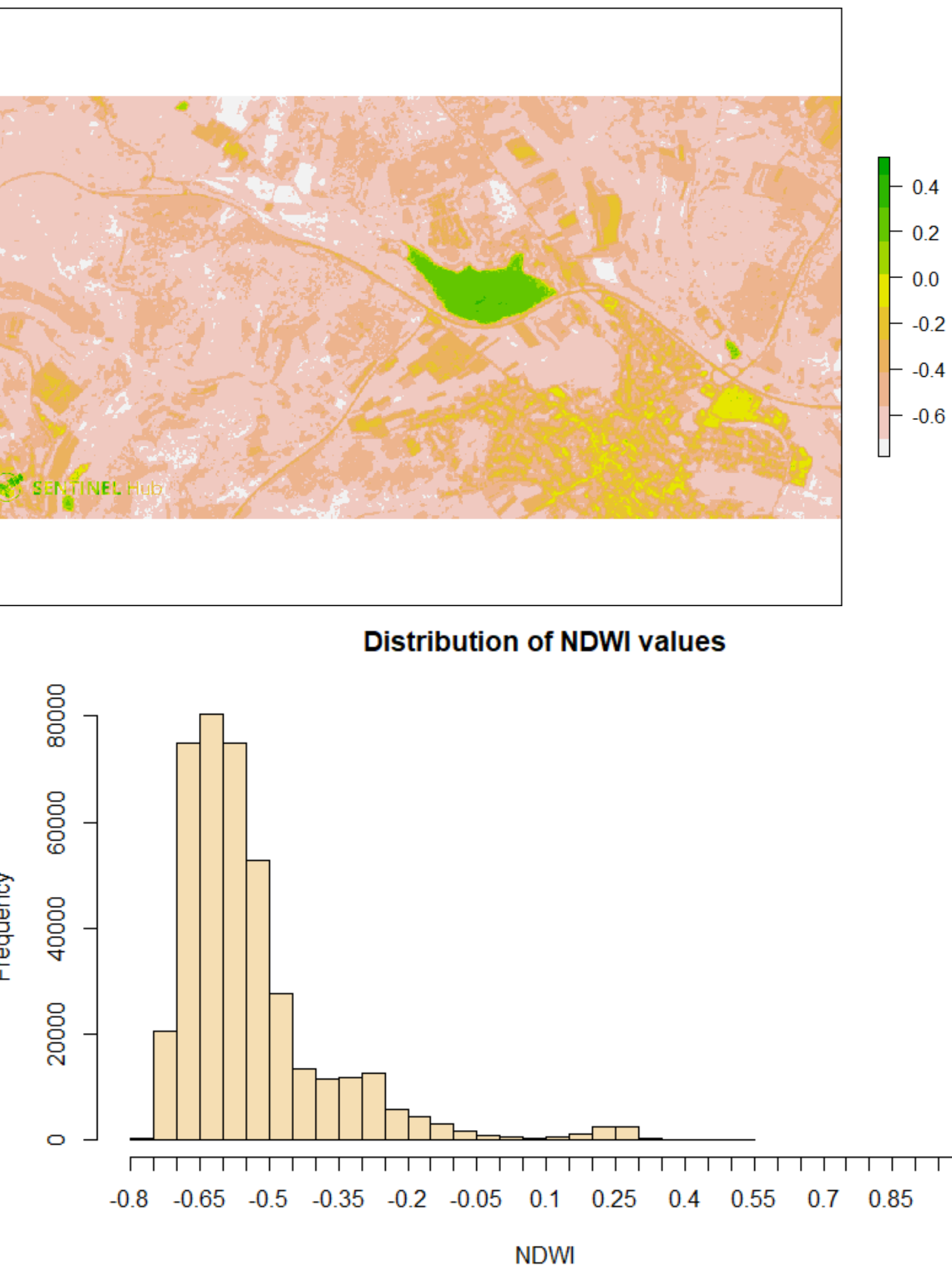
Lichtwellen werden unterschiedlich stark in verschiedenen Band Bereichen Reflektiert.

Ein Bild besteht aus mehreren Band-Layern, welche beliebig zusammengesetzt werden können, je mehr Bands ein Bild hat, desto größer ist die Resolution dieses Bildes.

### NDWI (Normalized Difference Water Index)

$$NDWI = \frac{(X_{green} - X_{nir})}{(X_{green} + X_{nir})}$$

Der NDWI subtrahiert und und dadurch negiert alle Reflektionen, welche keine Wasserkörper sind. Die Subtraktion geteilt durch die addition, sorgt dafür dass das Ergebnis auf einen Wertebereich von -1 bis 1 normiert wird. Dadurch werden alle positiven werte als Wasser klassifiziert und alle negativen als Landmasse.



### Image Segmentation mit K-Means

Algorithms:

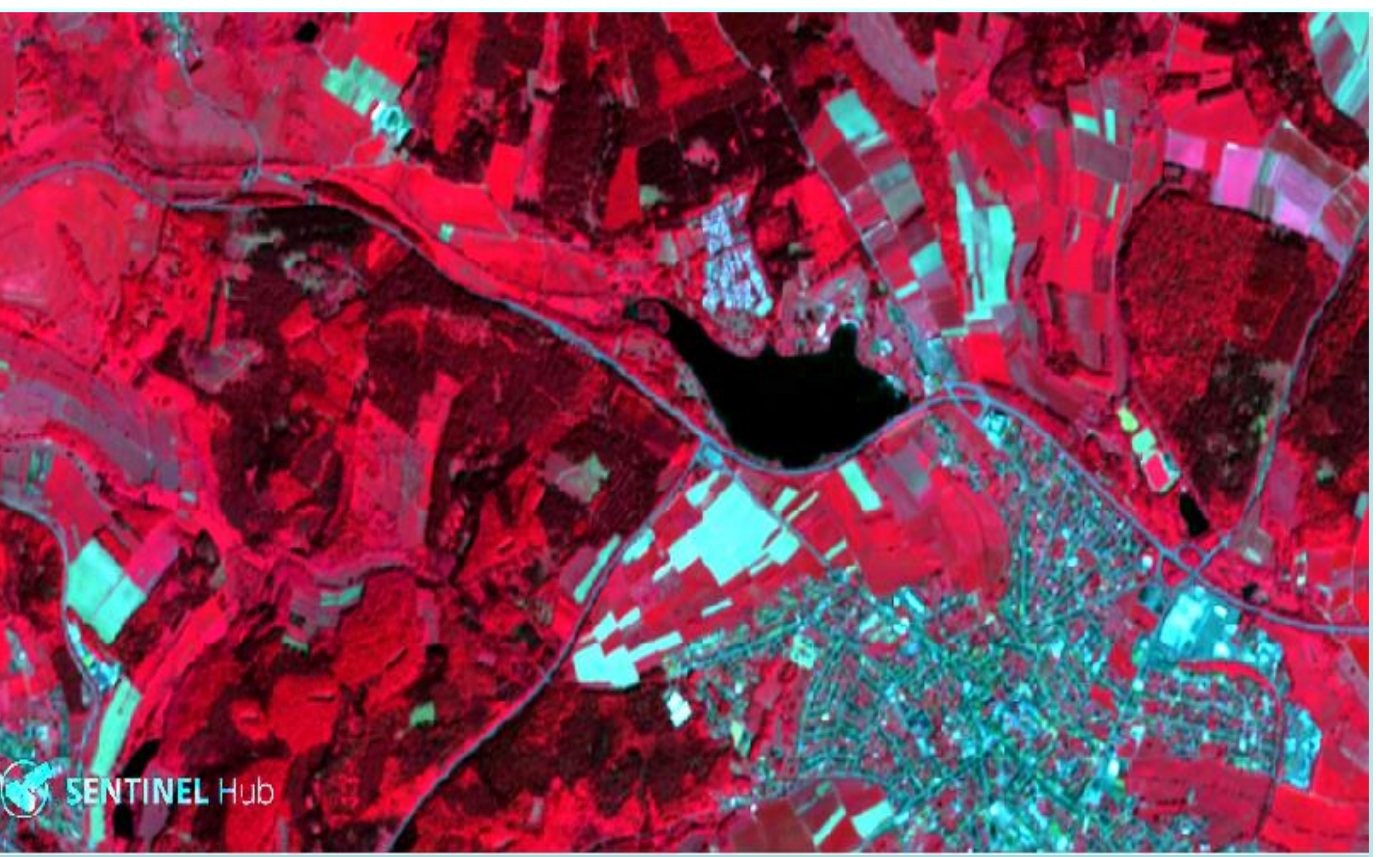
Zuerst wählt der Algorithmus k zufällige Clusterzentren zum Datensatz.

Dannach bildet der Algorithmus Cluster abhängig zum Geringsten Abstand zum Clusterzentrum..

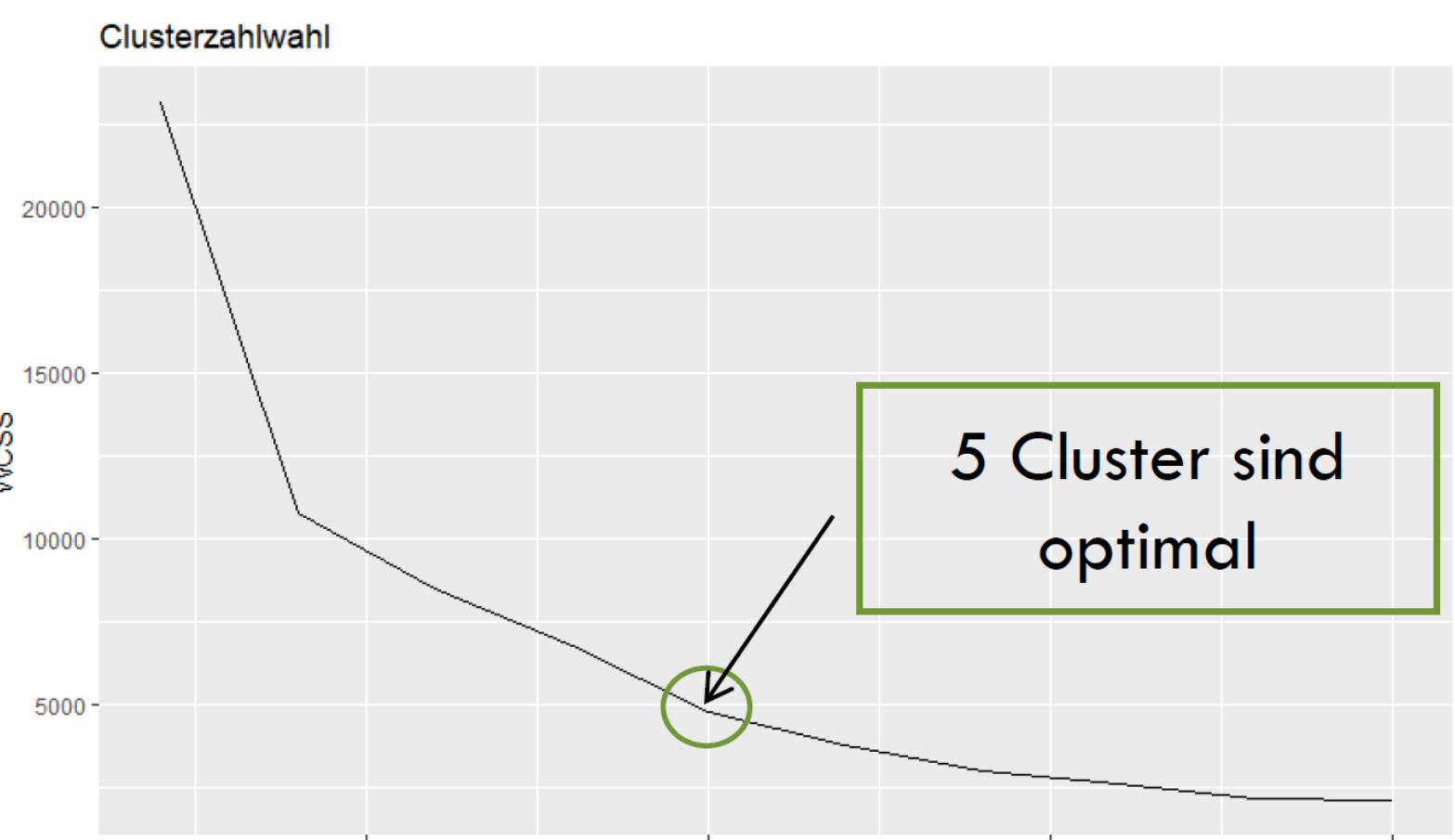
1st cluster attempt:

Zuletzt berechnet der Algorithmus die Variation zu jedem Cluster. Diese Prozedur wiederholt sich n mal (25 mal standartmäßig in R)

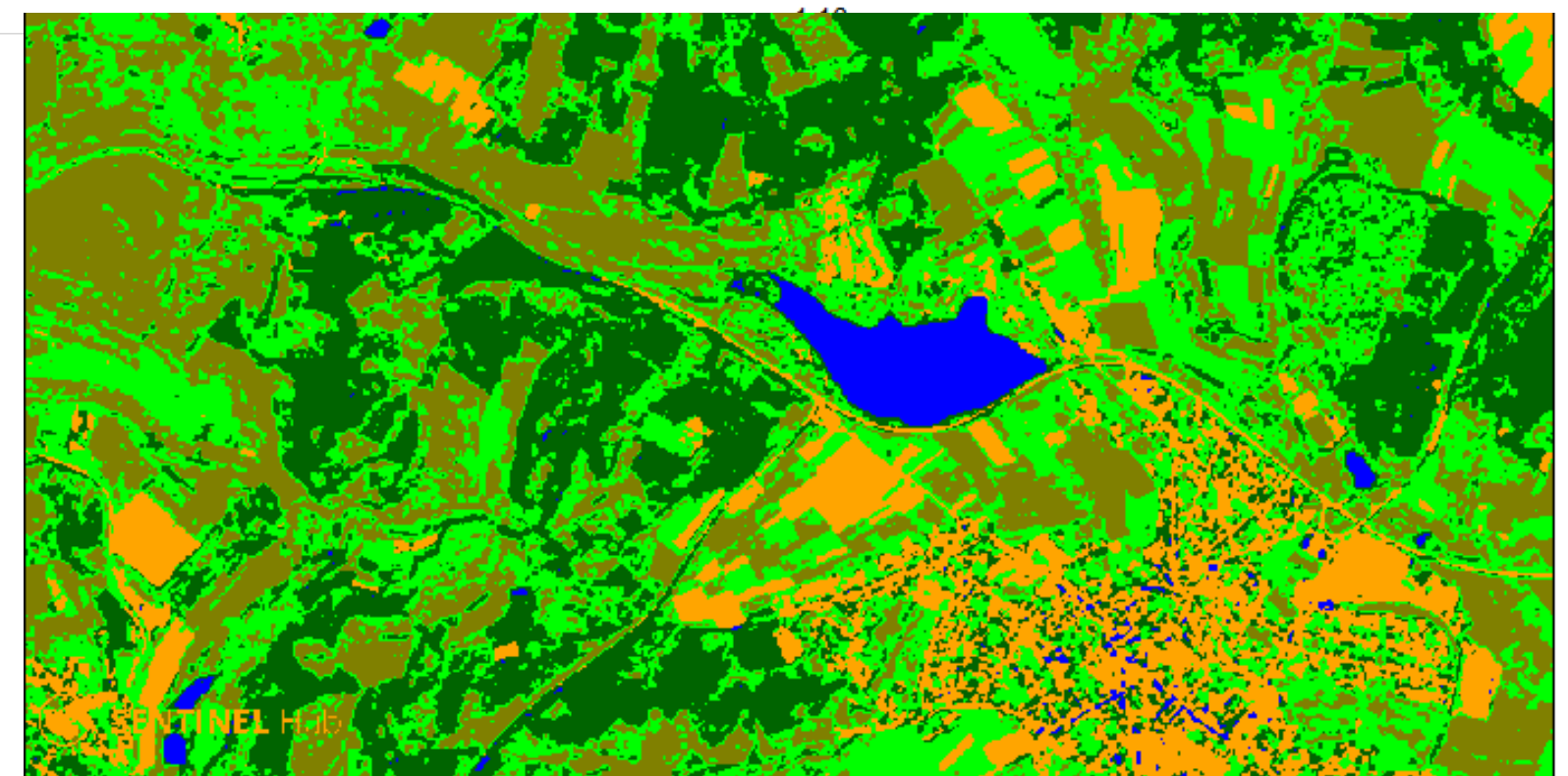
Zuletzt nimmt man die Iteration, die die geringste Variation zu seinen Clusterzentren aufweist.



Zur Klassifikation wird ein False Color Plot genommen, der Blue-Band wird hier mit dem NIR-Band getauscht, da der NIR-Band deutlich höhere Reflektionen aufweist und somit höhere Kontraste aufweist.

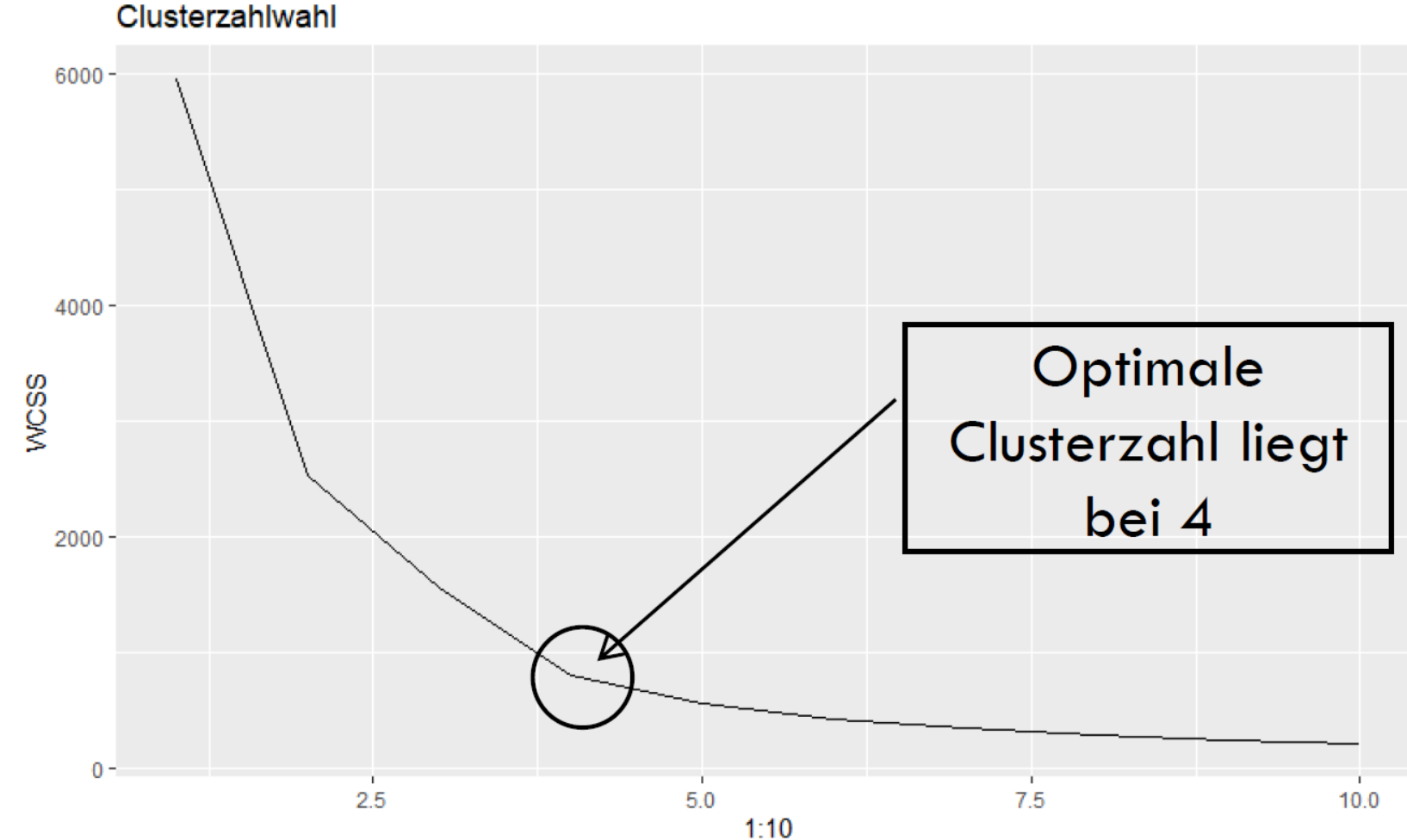


Mittels Elbogen-Methode wurde die optimale Clusterzahl auf 5 gesetzt.



Der K-Means – Algorithmus war dazu in der Lage, Wasser von anderen Landmassen zu Klassifizieren.

### Elbogenmethode zur Ermittlung von K



Die Ellbogenmethode summiert die gesammte Variation per Cluster auf und plotted diese. Das Ziel ist es die "geringste" Variation zu wählen. Jedoch nimmt die Variation mit zunehmender Clusterzahl immer weiter ab. Die Lösung hierzu ist es diejenige Clusterzahl zu wählen, ab der keine segnifikante Verbesserung in der Variation mehr festgestellt werden kann.

### Gegenüberstellung

NDWI:



K-Means:



Beide Verfahren sind dazu in der Lage Wasser zu klassifizieren. Jedoch weißt der NDWI eine deutlich geringere Anzahl an Missklassifikationen auf.

### Quellen

<https://www.youtube.com/user/joshstarmar>  
[https://www.youtube.com/watch?v=IRjcVYBUYVE&list=PLBvO9B D7ez\\_6lYVoZlRzVcOPIT5Lfjo0Y&index=13](https://www.youtube.com/watch?v=IRjcVYBUYVE&list=PLBvO9B D7ez_6lYVoZlRzVcOPIT5Lfjo0Y&index=13)  
<https://rsportal.org/raster/rs/index.html>  
<https://www.gis-blog.com/unsupervised-kmeans-classification-of-satellite-imagery-using-r/>  
<https://www.sentinel-hub.com>