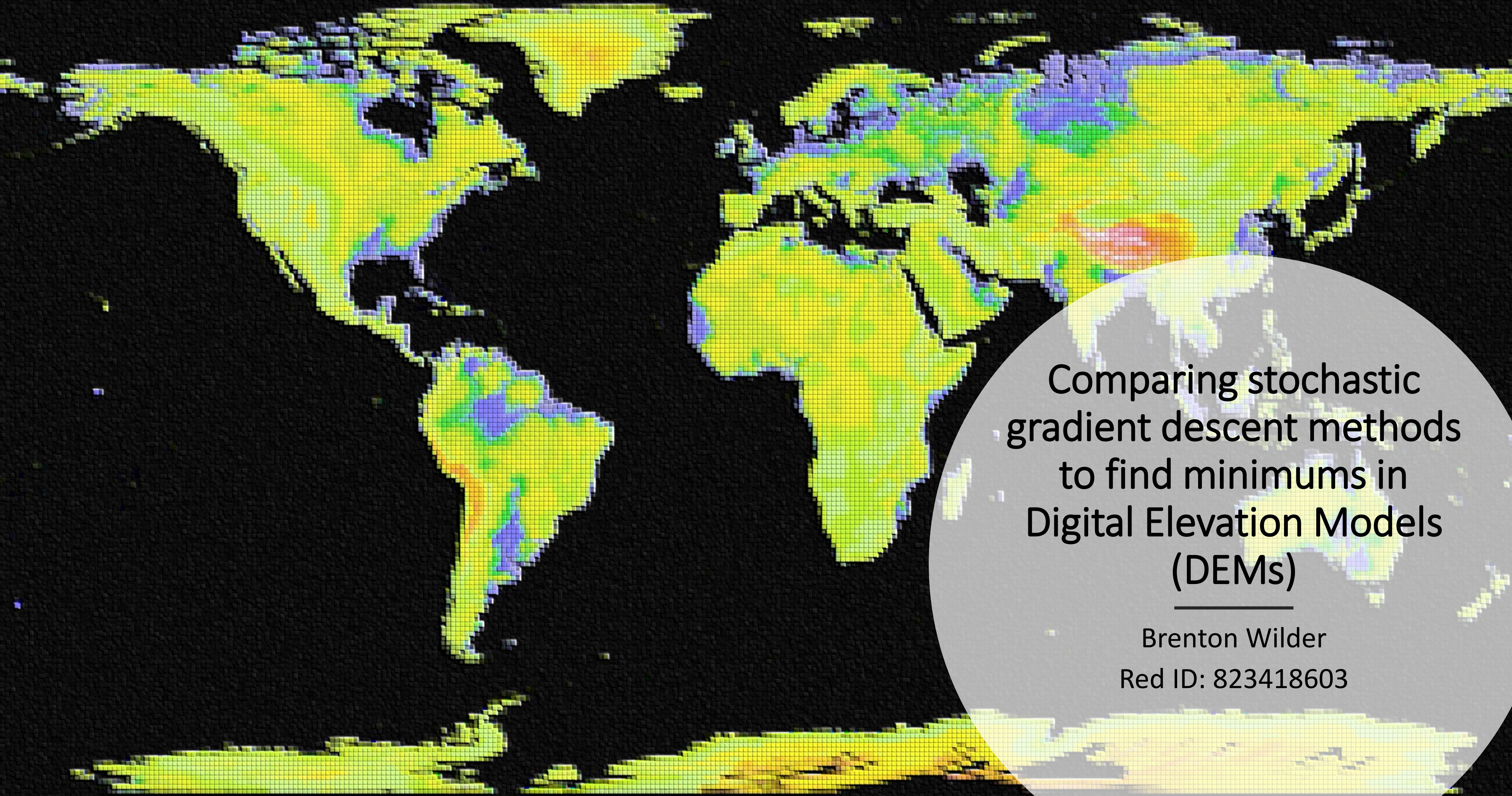
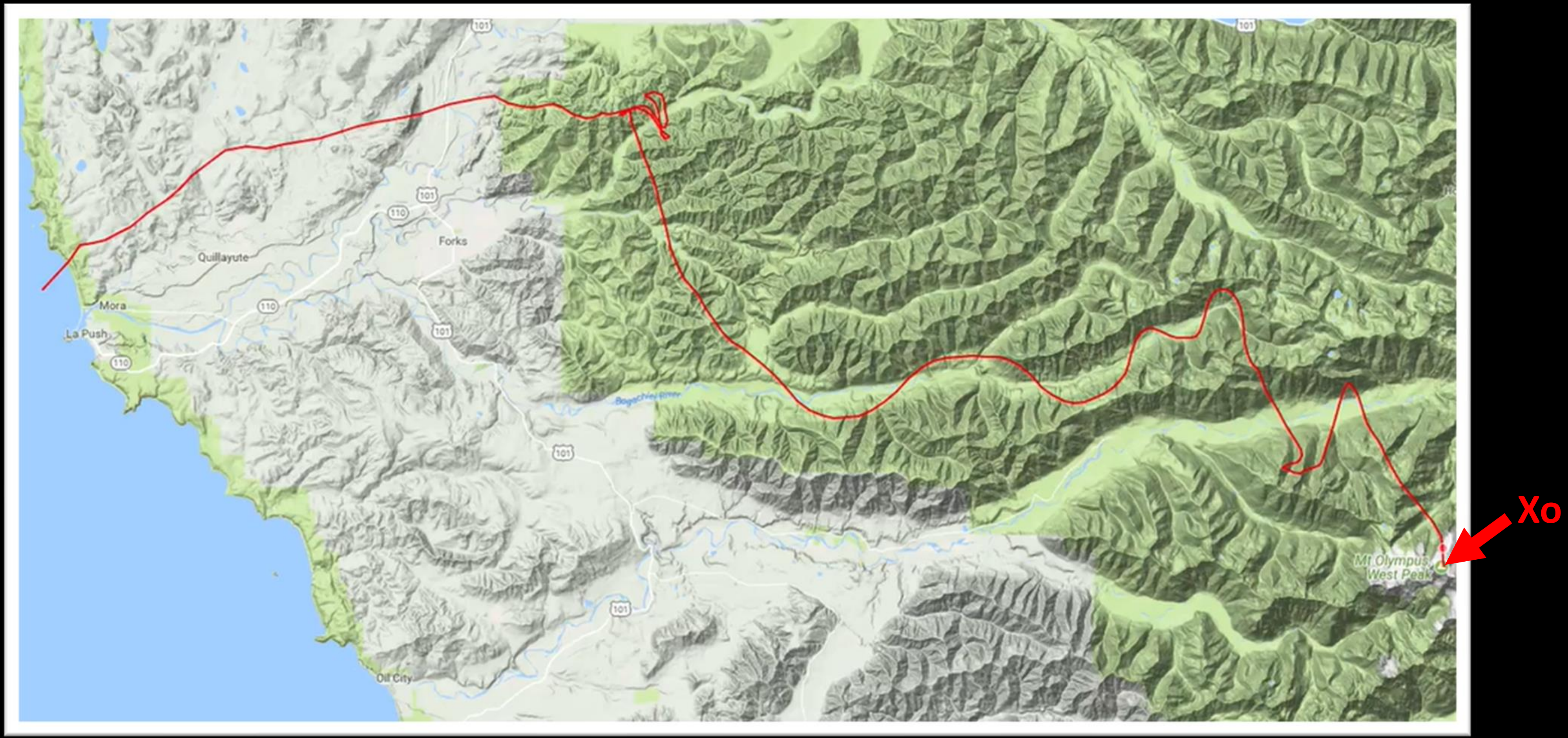


Comparing stochastic gradient descent methods to find minimums in Digital Elevation Models (DEMs)

Brenton Wilder

Red ID: 823418603

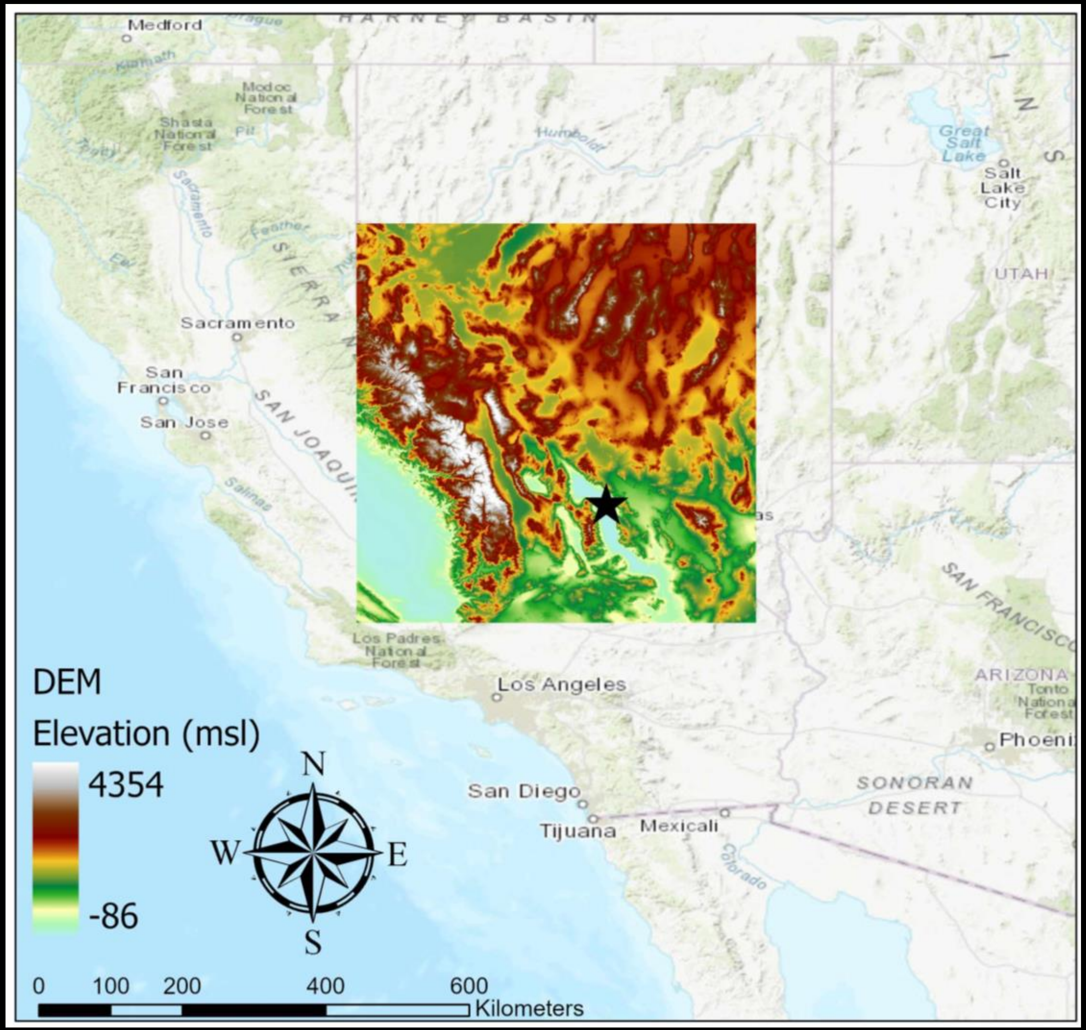
Introduction – Existing code using SGD and SGD with momentum to find ocean

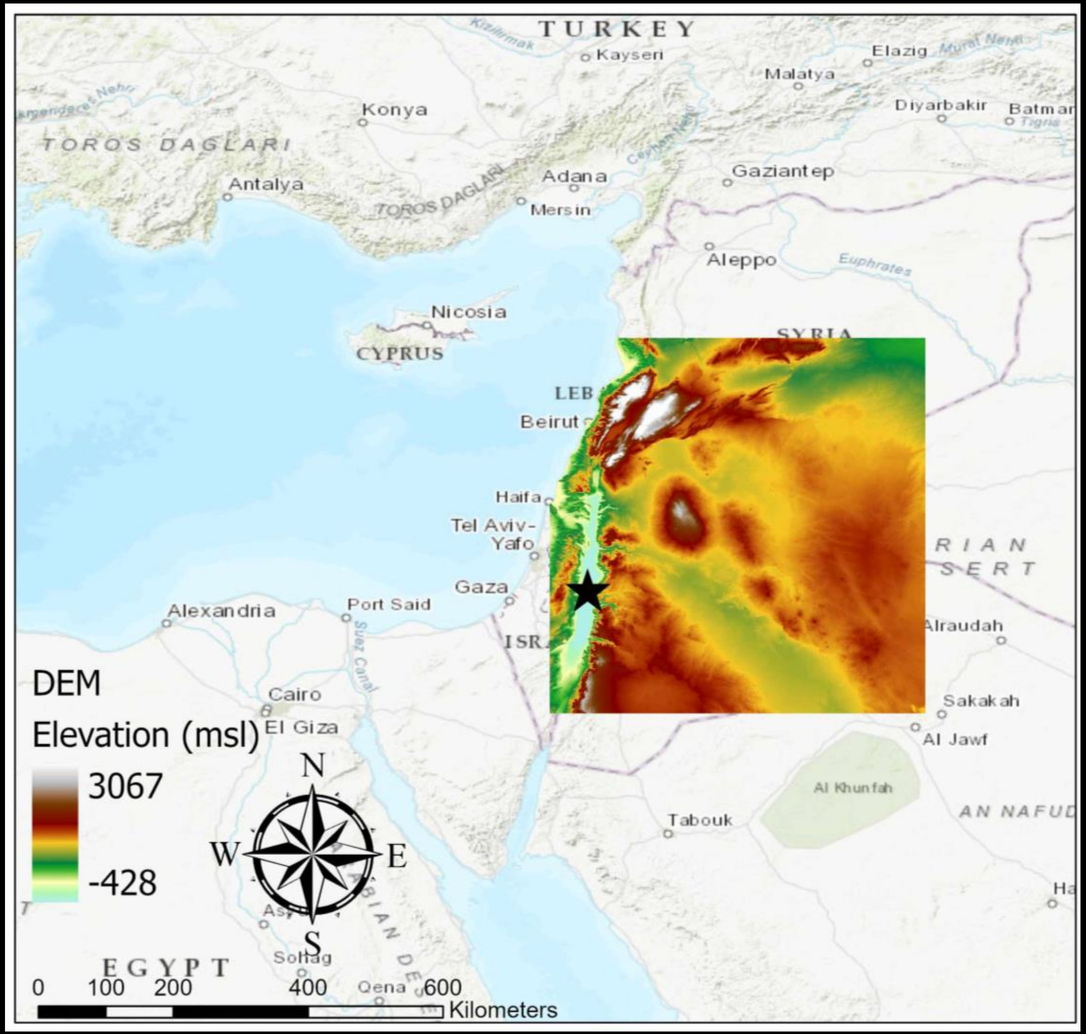


https://fosterelli.co/executing-gradient-descent-on-the-earth

Introduction – What about applying to more complicated task?

Materials and Methods – Two study sites





Materials and Methods – SGD on earth surface

$$\boldsymbol{\theta_{new} = \theta_{prev} - \alpha \, \nabla J(\theta_{prev})}$$

where:

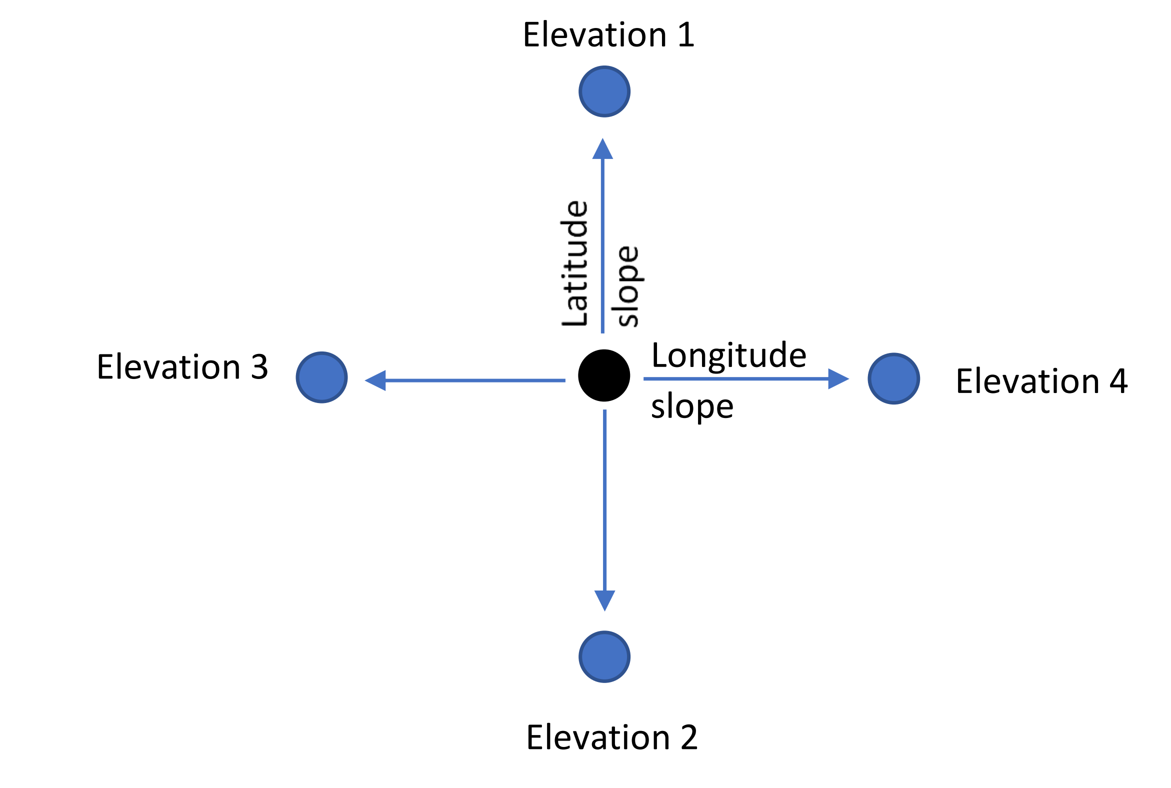
$$\alpha = \text{step size } (0.01)$$

$$\nabla J(\theta_{prev}) = \text{gradient of point}$$

$$\theta_{prev} = \textit{elevation of point}$$

$$\theta_{new} = \textit{new elevation}$$

```
# Fetch elevations at offsets in each dimension
elev1 = get_elevation(theta[0] + 0.001, theta[1])
elev2 = get_elevation(theta[0] - 0.001, theta[1])
elev3 = get_elevation(theta[0], theta[1] + 0.001)
elev4 = get_elevation(theta[0], theta[1] - 0.001)
```



SGD Methods

Stochastic Gradient Descent with Momentum

$V_t = \gamma V_{t-1} - \alpha \nabla J(\theta_{prev})$

$\theta_{new} = \theta_{prev} - V_t$

where:

α = step size (0.01)

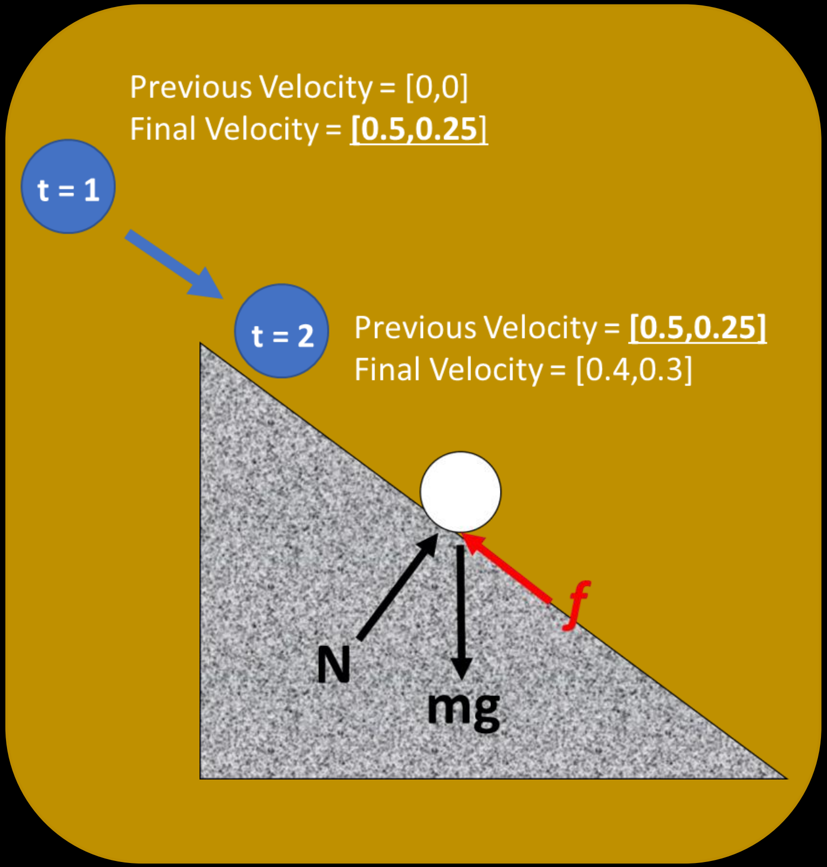
γ = momentum term (0.90)

V = "velocity" term

$\nabla J(\theta_{prev})$ = gradient of point

θ_{prev} = *elevation of point*

θ_{new} = *new elevation*



Adaptive Moment Estimation (Adam)

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1)\nabla J(\theta)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2)\nabla J(\theta)^2$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t}$$

$$\theta_{new} = \theta_{prev} - \frac{\alpha}{\sqrt{\hat{v}_t} + \epsilon}\hat{m}_t$$

where:

α = step size (0.01)

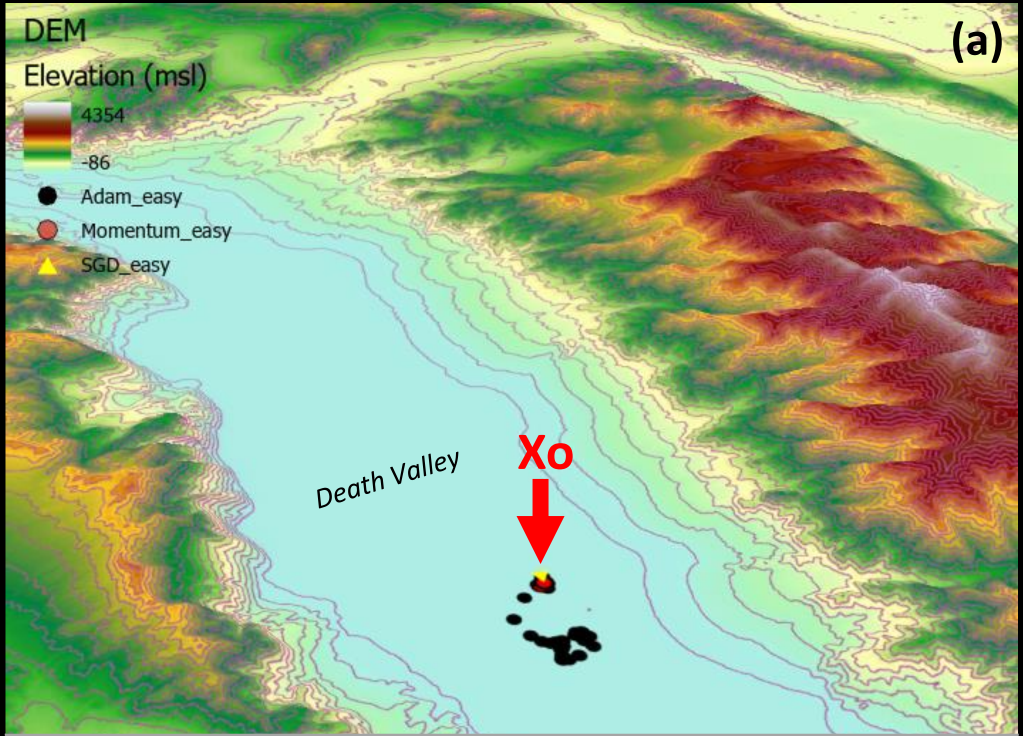
$\beta_1 = 0.90$

$\beta_2 = 0.999$

$\epsilon = 10e - 8$

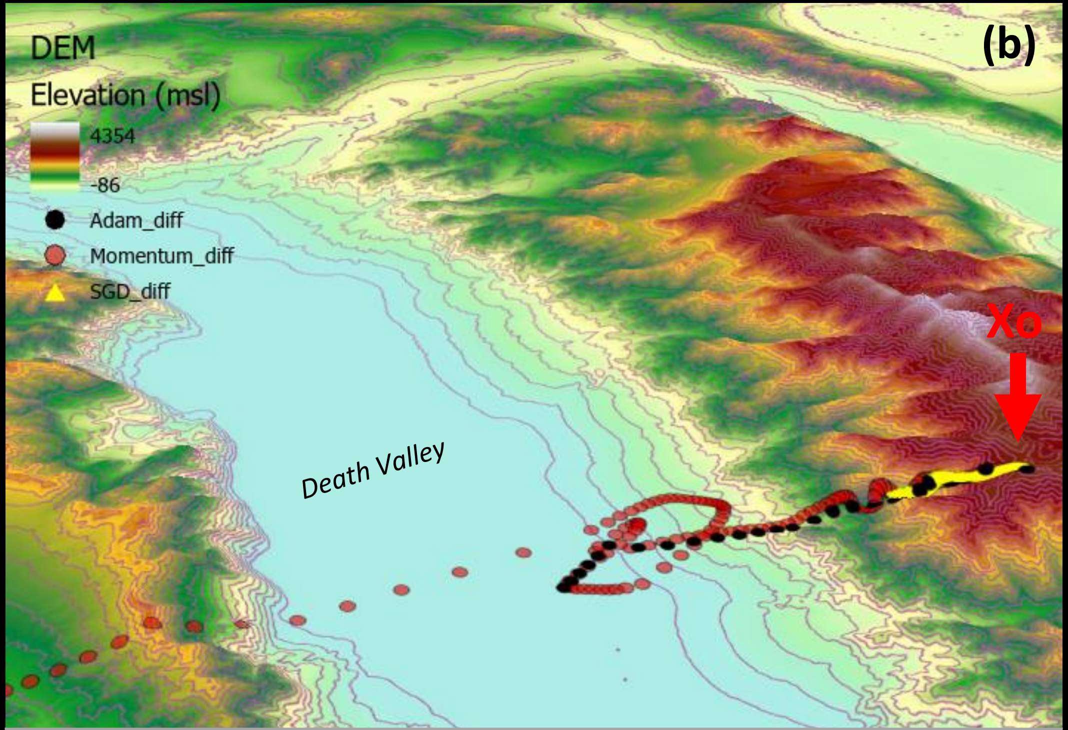
Results for site 1 – Death Valley

Xo = [36.2833290, -116.8716455] ; Min elevation = -86



Method	Iterations	Final Elevation (meters)
SGD	10,000+	-81
SGD with momentum	10,000+	-80
SGD with Adam	23	-86

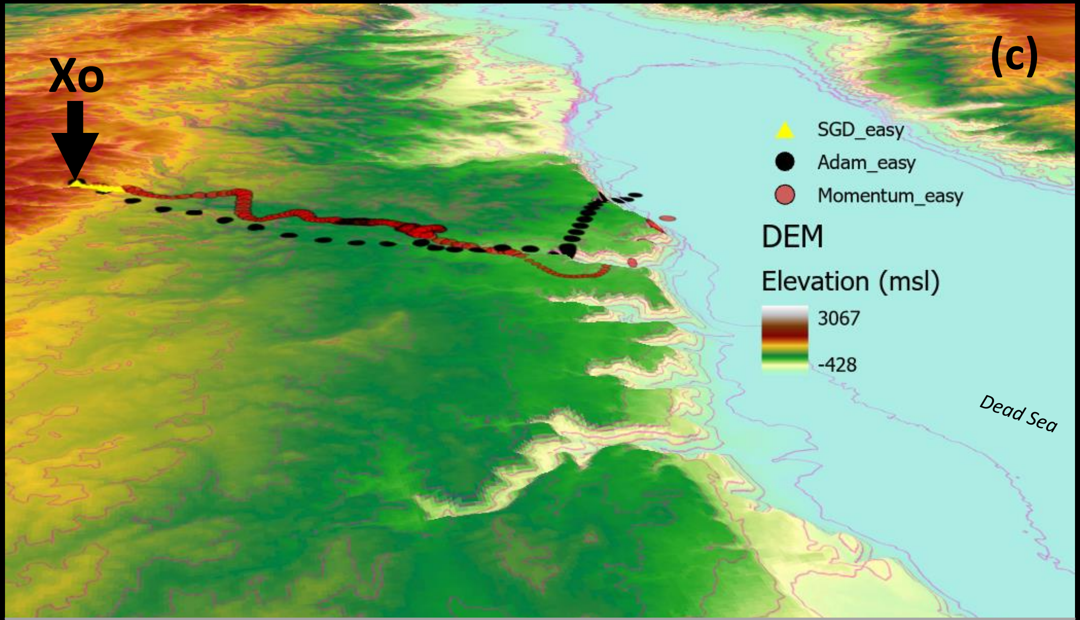
Xo = [36.2691711, -117.0563655] ; Min elevation = -86



Method	Iterations	Final Elevation (meters)
SGD	10,000+	1080
SGD with momentum	10,000+	626
SGD with Adam	26	-86

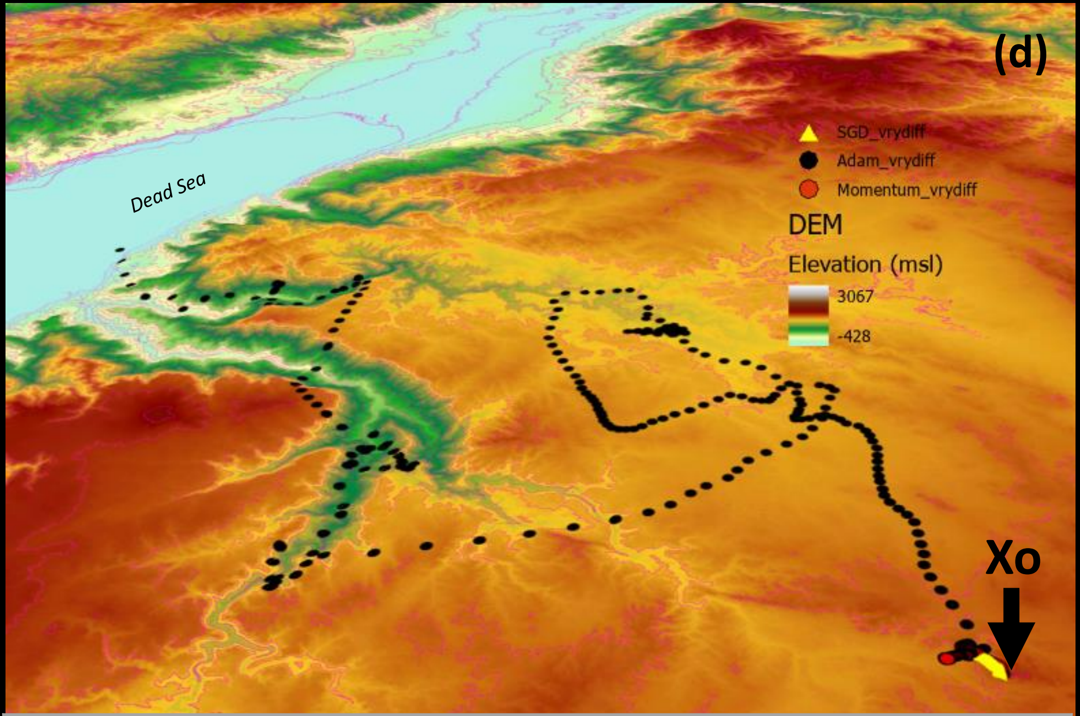
Results for site 2 – Dead Sea

Xo = [31.5575025, 35.1970443] ; Min elevation = -428



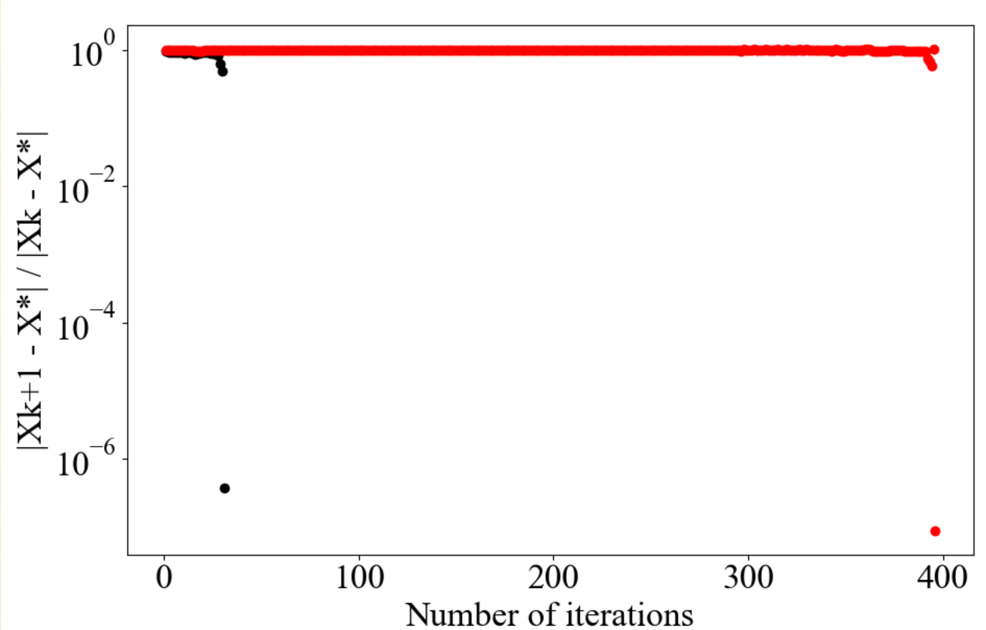
Method	Iterations	Final Elevation (meters)
SGD	10,000+	588
SGD with momentum	397	-415
SGD with Adam	32	-415

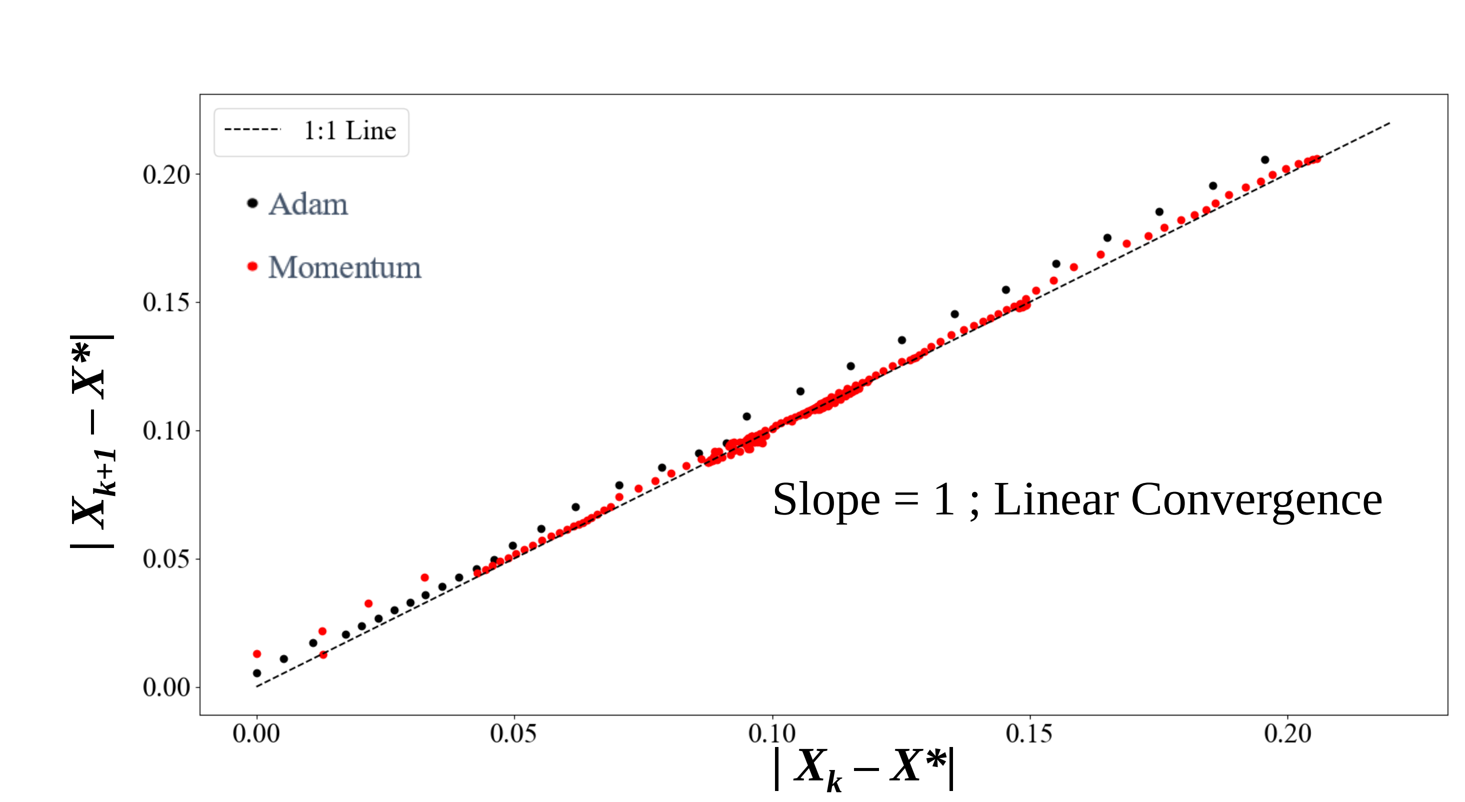
Xo = [31.3953303, 36.1588269] ; Min elevation = -428



Method	Iterations	Final Elevation (meters)
SGD	10,000+	800
SGD with momentum	10,000+	763
SGD with Adam	226	-415

Comparison of rate of convergence for (c)





Summary

- An existing code that performed gradient descent on the earth was modified to use an Adam algorithm where latitude and longitudes represented the [x,y] and elevation represented our z values. Adam can adjust the learning rate at each iteration, enabling it to slow down or speed up depending on the point in the function.
  - Note: There are so many improvements over Adam, however, I only learned this after reading more very recently (e.g. AdaGrad, AdaFom, AMSGrad, AdaMax, Nadam)
- Without calibrating, Adam outperformed gradient descent and the momentum algorithms and was consistently able to find the minimum elevation with linear rate of convergence
  - The Dead Sea true minimum could not be achieved, due to the slope of Dead Sea being too small for the algorithms to find the next point.

References

1.Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

2.Qian, N. (1999). On the momentum term in gradient descent learning algorithms. *Neural networks*, *12*(1), 145-151.

3.Shuttle Radar Topography Mission 1 Arc-Second Global (Digital Object Identifier (DOI) number: /10.5066/F7PR7TFT

4. Chen, X., Liu, S., Sun, R., & Hong, M. (2018). On the convergence of a class of adam-type algorithms for non-convex optimization. arXiv preprint arXiv:1808.02941.

chrisfosterelli/physical-gradient-descent
bwilder95/GerudoValley