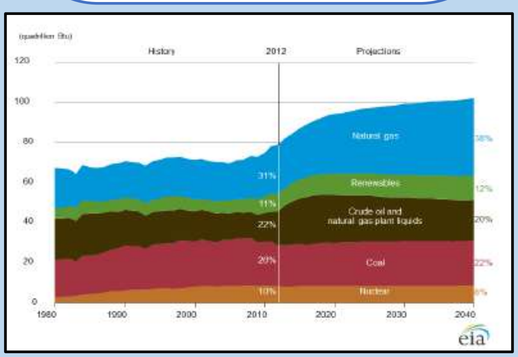
PhD RESEARCH: NUMERICAL MODELING OF COAL EXCAVATOR – ROCK FORMATION INTERACTIONS

SOMUA-GYIMAH, Godfred [PhD Research]

Abstract

- Coal contributes over 20% of the energy in the US and the rest of the world. This is not expected to change over the next 20 years.
- A recent focus of coal production research studies has been to develop new excavator bucket designs which outperform conventional buckets.
- A major bottleneck in the current bucket design approach is that every proposed design has to be physically built, tested and compared.
- A computer model, that sufficiently replicates real-life excavator-ground interactions, will lead to huge savings in time, efficiency, design & opportunity costs.
- The exact relationship between DEM parameters and actual rock properties is unknown. Hence, replicating the behavior of a **specific** setup is tough.
- Existing models adopted 'trial and error' property-matching techniques. However, these models overpredict excavation outcomes by 300% - 500%.
- This study introduces a material calibration and excavation simulation approach that reduce excavation prediction errors from 300% to 16.55%.
- The material calibration method combines DEM-based tri-axial rock testing with the XGBoost machine learning algorithm to achieve prediction accuracies of between 80.6% and 95.54%.

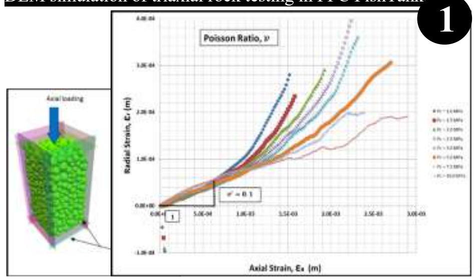


Forecast US Energy Production (EIA, 2014)

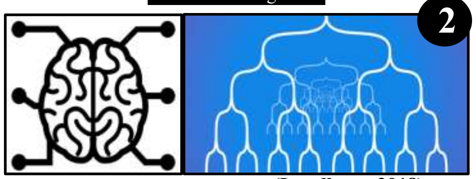
Methods

1. DEM simulations of laboratory rock test experiments are performed in PFC FishTank to generate data (1500 observations) for model training and testing.
2. A model is trained with the XGBoost ML algorithm (Chen, 2016) to predict DEM rock parameters based on corresponding laboratory rock properties.
3. The excavator bucket – rock interaction is simulated and the results are compared with field observations.

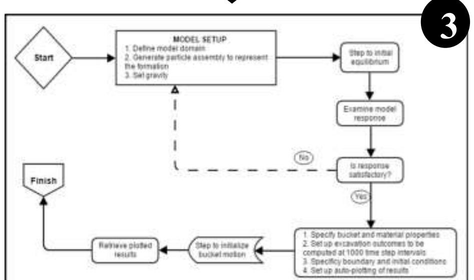
DEM simulation of triaxial rock testing in PFC FishTank



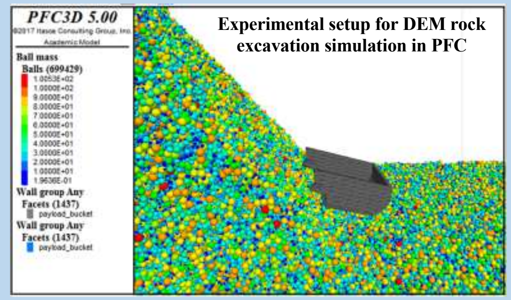
Predictive modeling of DEM properties with XGBoost ML algorithm



(Lundberg, 2018)

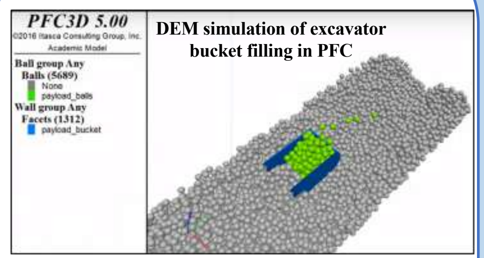


Flowchart for DEM simulation of rock excavation



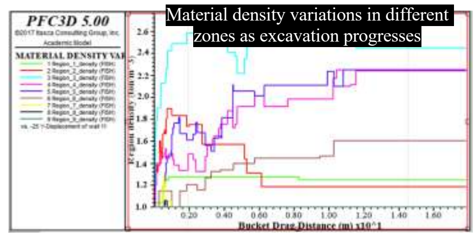
Experimental setup for DEM rock excavation simulation in PFC

Results

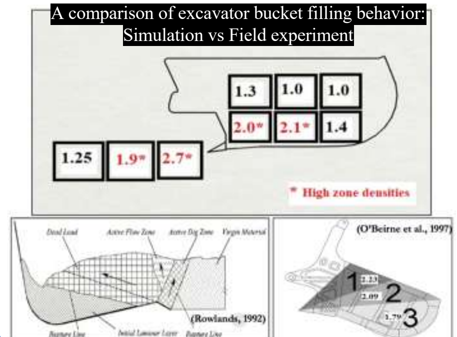


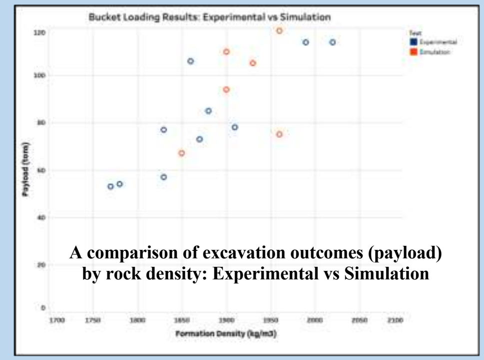
DEM simulation of excavator bucket filling in PFC

Material density variations in different zones as excavation progresses



A comparison of excavator bucket filling behavior: Simulation vs Field experiment





A comparison of excavation outcomes (payload) by rock density: Experimental vs Simulation

Conclusions

1. The relationships between DEM parameters and rock properties are more complex than can be explained by simple calibration models.
2. The proposed material calibration process is able to achieve prediction accuracies of 80.6% to 95.54%. The DEM excavation model reduces excavation prediction errors from 300% to 16.55%.
3. The optimum rock size distribution for excavation ranges from fines up to ~25% of the bucket width.
4. There is a material density distribution which develops inside and ahead of the excavator bucket during filling. This distribution decreases towards the rear of the bucket.
5. The most active material zones are typically within a distance equal to two-thirds of the bucket length.

Novelty / Impact

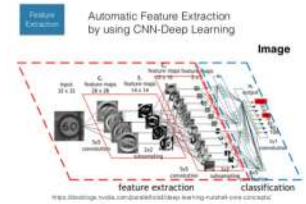
- This study introduces a new approach to DEM geomaterial calibration by combining extensive rock test simulation with the XGBoost ML algorithm.
- This DEM model complements coal studies by providing a cheap and time-efficient tool for comparing different bucket geometries for design improvements.
- This study is the first DEM attempt to investigate dragline bucket loading behavior in 3D & at full scale.

References

* All uncited images are from the PhD dissertation

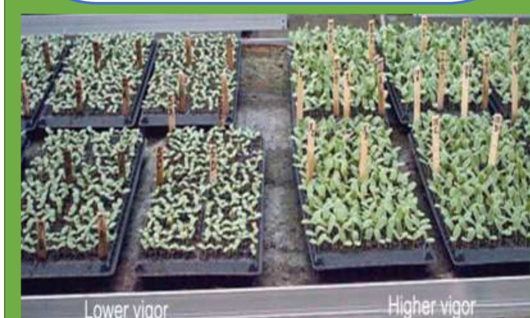
IMAGE ANALYTICS: PREDICTING SEED QUALITY FROM SEED X-RAY IMAGES

SOMUA-GYIMAH, Godfred [PoC Project]



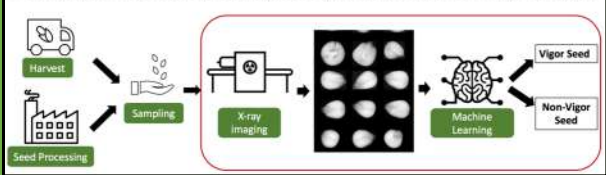
Introduction

- Seed quality is often a key consideration for seed growers, when choosing from the different seed brands available on the market.
- To the farmer, high quality seeds implies a high crop yield and revenue. For seed companies, this translates to fewer returned seeds, increased customer trust and satisfaction, customer retention, increased market share and revenue.
- Seed quality assurance and control is therefore, an important step in ensuring that the seeds which eventually reach the customer, are of acceptable quality.
- The current seed quality assessment method involve sampling some seeds (< 1%) from each batch and performing germination tests. The germination performance (in %) is then assumed to be representative of the entire seed batch.
- The combination of X-ray imaging and machine learning presents a huge opportunity for automating, scaling and speeding up the entire seed quality assessment process.
- This will provide a faster quality metric as well as the ability to screen entire seed batches. The ability to screen every single seed for quality before sending it off to the farmer is one huge opportunity that a fast, automated seed testing method may guarantee.
- This POC project seeks to determine the possibility of predicting seed germination test outcomes from seed x-ray images using machine learning.



Seed Germination Testing (Marcos-Filho, 2015)

Flowchart for predicting seed quality from X-ray Images



Methods

- Above is the proposed flowchart for seed quality predictions.
- The data set of 48000 x-ray images was randomly partitioned into training, validation and test sets using a 70:15:15 ratio respectively.
- Downsampling, upsampling and data augmentation techniques were used to mitigate the high class imbalance (92% vigor, 8% non-vigor)
- Two CNN architectures, VGG16 and ResNet50, were tried. In each case, the 1000-output fully connected layer was replaced with batch normalization, dropout and densely-connected layers with a binary output in the final layer.
- Different models were developed from the 2 architectures by varying the learning rate, L1/L2 regularization parameters, dropout rate and training time among others.
- Different models were also developed by sub-grouping the seeds according to shape and size.
- Single models and different ensemble model combinations were then compared.

Seed quality model base architectures

Layer (type)	Output Shape	Param #
vgg16 (Model)	(None, 8, 8, 512)	14714688
batch_normalization_1 (Batch	(None, 8, 8, 512)	2048
gaussian_dropout_1 (Gaussian	(None, 8, 8, 512)	0
flatten_1 (Flatten)	(None, 32768)	0
dense_1 (Dense)	(None, 8)	262152
batch_normalization_2 (Batch	(None, 8)	32
dropout_1 (Dropout)	(None, 8)	0
dense_2 (Dense)	(None, 1)	9

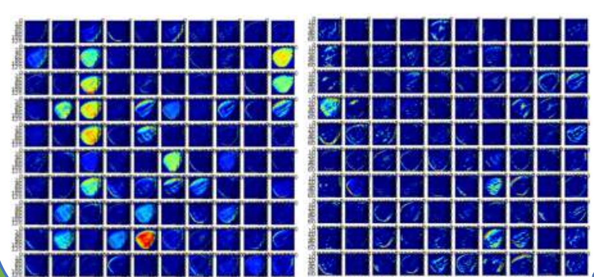
Total params: 14,978,929

VGG16

Layer (type)	Output Shape	Param #
resnet50 (Model)	(None, 1, 1, 2048)	23587712
batch_normalization_1 (Batch	(None, 1, 1, 2048)	8192
gaussian_dropout_1 (Gaussian	(None, 1, 1, 2048)	0
flatten_1 (Flatten)	(None, 2048)	0
dense_1 (Dense)	(None, 8)	16392
batch_normalization_2 (Batch	(None, 8)	32
dropout_1 (Dropout)	(None, 8)	0
dense_2 (Dense)	(None, 1)	9

Total params: 23,612,337

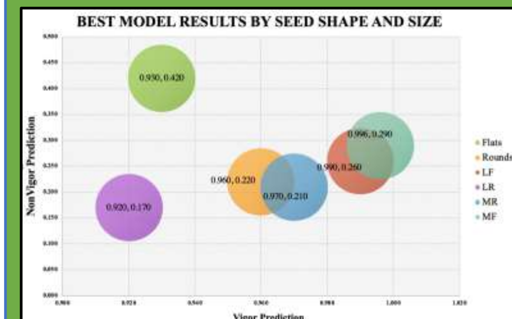
ResNet50

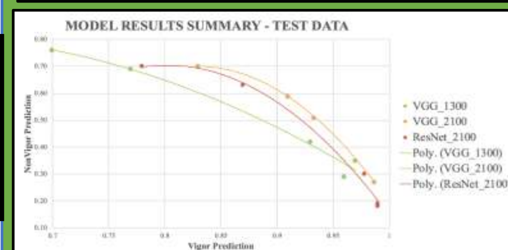


Some feature maps from the convolution filters

Results

1. The best class predictions were 99.6% and 77.0% for vigor and non-vigor classes respectively.
2. Generally, flat seed models (Flats, LF, MF) performed better than round seed models (Rounds, LR, MR) but size/shape sub-grouping did not significantly improve model performance.
3. For the same model, hyper-parameter tuning results lay on the same vigor-nonvigor inverse correlation curve.
4. Increasing from 760 to 1321 (a factor of 1.7) nonvigor training images improves non-vigor prediction by 1.3 times (0.22 to 0.28). Similarly, increasing from 760 to 2077 (a factor of 2.7) nonvigor images improves non-vigor prediction by1.8 times (0.22 to 0.39).
5. Hence, the biggest gain in model performance was achieved by increasing the number of non-vigor images, rather than through CNN architecture selection or hyper-parameter tuning.
6. The best single model achieves 99.6% & 20.0% across vigor and nonvigor classes. However, an ensemble of the 9 best models achieves 99.6% & 59.3% across vigor and nonvigor classes.





Conclusions

1. The best model (99.6% vigor | 59.3% nonvigor) is an ensemble of the top 9 single models.
2. The class imbalance affected model training, even after data augmentation.
3. Results suggest that increasing the number of nonvigor images could improve model training.

Novelty / Impact

- This study provides promising evidence that machine learning could be used for seed vigor predictions.
- There is the potential to significantly reduce the occurrence of seed returns.
- Pilot phase testing is currently underway.
- Plans are in place to obtain more non-vigor data for model improvements.

References

Marcos Filho, J. (2015). Seed vigor testing: an overview of the past, present and future perspective. *Scientia Agricola*, *72*(4), 363-374.