Reported Production - GFO

Forecasts

Models Compared

- Moving Average
- ARIMA
- Neural Network (LSTM)

- The estimates are done for the Year: **2023 – 24**
  - From Aug 1, 2023 until July 1, 2024.
  - Actuals are compared to predicted.

- Note that other years for ARIMA and NN are also possible, but then we are just losing observations.
  - For example, if we do it for 2022 (a year before), then a further 12 observations will be lost leaving the model with just 145 observations to work with.
  - Econometric and Machine Learning models lose power with lower observations.

Results

	Barley	Corn	Mixed O/B	Oats	Soybean	Wheat
Actual (Aug 23' – July 24')	**50,907**	**7,840,968**	**4,545**	**54,588**	**3,697,987**	**3,070,699**
Moving Avg (MA) (5-yr)	51,244	7,375,225	5,888	51,197	3,685,863	2,168,876
ARIMA	47,432	7,856,412	3,576	63,091	3,937,989	2,404,240
Neural Network (NN – LSTM)	48,324	6,720,761	6,365	36,899	3,324,519	2,251,709

Difference between actual and predicted

	Barley	Corn	Mixed O/B	Oats	Soybean	Wheat
Actual (Aug 23' – July 24')	-	-	-	-	-	-
Moving Avg (5-yr)	**-337**	465,743	-1,343	**3,391**	**12,124**	901,823
ARIMA	3,475	**-15,444**	**969**	-8,503	-240,002	**666,459**
Neural Network	2,583	1,120,207	-1,820	17,689	373,468	818,990

- A negative sign indicates that the model over estimated and a positive sign indicates underestimation.

- The magnitude in either direction also dictates model performance.

The Green color highlights, which model performed best for that specific commodity. Summarized in the next slide

Summary

The table below shows the best model for each crop, based on the previous slide:

	Barley	Corn	Mixed O/B	Oats	Soybean	Wheat
Best Model	MA	ARIMA	ARIMA	MA	MA	ARIMA
Second-best Model	NN - LSTM	MA	MA	ARIMA	ARIMA	NN-LSTM

- ARIMA is the best model for Corn and Wheat.

- MA is working especially poorly for Wheat.
    - Yet is better for Barley, Oats and Mixed O/B.
    - Why is ARIMA working worse for Barley and Oats (I would argue, just lack of observations).

- While MA predicted Soybeans better too – I fear it is discounting trend movement, I think ARIMA's prediction for soybean (see previous slide) is "better" even though it overestimates by 240k MT.
    - I envision that going forward (especially for 24) – ARIMA will predict Soybean better too (but I could be wrong).

Assumptions and understanding results

Why are the models performing the way they are:

- A "Moving Average" (MA) ignores inter-period variation entirely.
    - However, is suitable for data where there is little variance.
    - Also, might be better for its simplicity.

- ARIMA and NN both incorporate past values, overall trend and variance in the data.

- A data with "high variation" will always be trickier to predict. Case in point: wheat.

- For example: see table below – Coefficient of Variation (CV) is a measure which shows the data's dispersion. A higher percentage means more of the data is dispersed farther away, relative to the mean. This is proof to show why for eg: MA performs poorly for wheat and better for soybeans. Wheat production is a lot more erratic as can be seen.

	Barley	Corn	Mixed O/B	Oats	Soybean	Wheat
CV	20.61%	11.08%	35.59%	30.11%	9.26%	22.39%