US corn yield in August: a scored 80% band

Desk memo. Public USDA data only, rebuilt as archived for each release morning. Every number below is drawn from files generated by the repository's `make memo`, and `cfstore trace` checks that each one appears there.

Nik Khosravipour. Code and data: https://github.com/nik110403/usda-august-corn-misses

Question

How far can USDA's January final corn yield land from its August print? We put an 80% band around the August print, built from USDA's own past August-to-January misses, and scored it on 22 seasons, 2004–2025. It covered 21 of 22; the pre-registered verdict is "too wide", with about 4 points of the excess expected from the estimator itself.

August 2025 answer

Object	Unit	USDA point	q10	q50	q90	80% band	Truth (Jan first-final)	Inside
M1	bu/ac	188.8	178.6	188.8	200.6	[178.6, 200.6]	186.5	yes
PROD	Mbu	16742	15592	16768	17893	[15592, 17893]	17021	yes

On August 12, 2025, the USDA estimated corn yield at 188.8 bu/ac, with our 80% band ranging from 178.6 to 200.6 bu/ac. The corresponding production estimate was 16,742 Mbu, with our band at 15,592 to 17,893 Mbu. By January, the final yield was 186.5 bu/ac and production was 17,021 Mbu, both within their respective bands. This single season's result is merely a check on the arithmetic, with the true test of our band's performance lying in its 22-season track record.

Stocks-to-use decision line

USDA's August balance sheet puts 2025/26 stocks-to-use at 13.3%. Our q10 production, put through that sheet with use held fixed, gives 6.1% (q50 13.4%, q90 20.5%). This is conditional on USDA's use projections and is not scored. Carry-in moves the sheet too: since July USDA revised beginning stocks by −35 Mbu, against +1,037 Mbu on production, and in 2019 the August-to-November carry-in revision was as large as the production one. Price steepens at low stocks (Wright 2011), which is why the downside tail deserves a look; this step estimates no price response.

How the band scored, 2004–2025

Band	Status	Horizon	Level	n	x	Coverage	CP90	Pre-registered verdict
Yield	exploratory	May	80%	22	20	90.9%	[74.1%, 98.4%]	not distinguishable
Yield	exploratory	Jun	80%	22	18	81.8%	[63.1%, 93.5%]	not distinguishable
Yield	exploratory	Jul	80%	22	18	81.8%	[63.1%, 93.5%]	not distinguishable
Yield	primary	Aug	80%	22	21	95.5%	[80.2%, 99.8%]	too wide
Yield	exploratory	Sep	80%	22	21	95.5%	[80.2%, 99.8%]	too wide
Yield	exploratory	Oct	80%	20	17	85.0%	[65.6%, 95.8%]	not distinguishable
Yield	exploratory	Nov	80%	22	17	77.3%	[58.0%, 90.6%]	not distinguishable
Production	exploratory	Aug	80%	22	20	90.9%	[74.1%, 98.4%]	not distinguishable

Status follows the pre-registration: only the August yield row is the primary test, the rest are exploratory. "Not distinguishable" is short for "not distinguishable from 80% at this sample size, and not shown equivalent".

In the primary pre-registered result, 21 of 22 seasons fell within the August band, a 95.5% coverage rate with a 90% Clopper-Pearson interval of 80.2%–99.8%, so the pre-registered wording is "too wide". The verdict sits on a knife-edge: one more miss would make it "not distinguishable". The estimator is conservative by construction. Under exchangeable errors, averaged over these sample sizes, the 80% band is expected to cover 84.2%, so about 4 points of the excess are expected from the estimator itself, and the interval contains 84.2%. M3 and M1 are not distinguishable at August with 22 seasons: adding crop condition (Good/Excellent share) did not improve mean pinball loss (one-sided p = 0.888), and the M3 band covered only 15 of 22. None of the 33 exploratory tests has a Holm-adjusted p-value below 0.05.

When this fails

Descriptively, in the pre-named stress years, the band breaks in drought springs and on late surprises. In 2012, a drought year, the bands from the May, June and July origins missed the final, which fell beyond both the 50% and 80% bands; the August band covered it. In 2020 the January final fell outside the 80% band from the October and November origins. The stocks-to-use line holds use fixed, so in years like 2011 and 2012 its q10 would need negative stocks and gives no number. Eleven WASDE and nine NASS releases survive only as corrected copies, with the originals not archived; two of the eleven are from 2026, outside the sample. They are flagged and used as archived, and no USDA error is claimed. The sensitivity run drops the 2 affected origins; the August band still covers 21 of 22.

Port to other origins

The same M1 calculation runs on Russian wheat production, with Russia-specific parsing and truth selection. This measures how far USDA revises its own Russia number before settling it: the settled figure is the last WASDE print before the marketing year leaves the table, in April about three years after harvest. It does not measure the harvest. The line shows the harness runs on a second series and makes no claim about effectiveness. Analyzing 18 seasons, from 2006 to 2023, the August USDA print's distance to the settled figure shows an 80% band covering 18 out of 18 seasons, with a 90% Clopper-Pearson interval of 84.7% to 100% coverage. The verdict on the band is "too wide". Averaged over the 18 seasons, the estimator alone expects 84.4% coverage, so about 4 points of the excess are expected from it. This exploratory analysis can be applied to other origins, such as EU rapeseed or Brazil safrinha, where dated vintages exist (MARS, CONAB).