

Climate Information and Wheat Futures: A Filtered-Information Theory of Forecast Value, Equilibrium Response, and Trading Feasibility

Arya Somu

November 15, 2025

Abstract

Climate-conditioned futures trading poses three distinct mathematical questions: whether an observation improves the conditional law of crop output, whether it changes the conditional law of a post-release futures return relative to market information, and whether any such change admits positive utility after execution costs. We formulate these questions on nested filtrations and connect a stochastic production layer to an inventory-sensitive local equilibrium and a side-specific contract payoff. Within the stated assumptions, a projection identity characterizes physical forecast value; an implicit-function result gives the local supply-price derivative; an observational-equivalence theorem proves that physical predictability alone cannot identify trading alpha; and a robust optimization theorem yields a cost- and uncertainty-dependent abstention region. Measurement-error and overlapping-horizon results further delimit inference. A point-in-time evaluation protocol specifies the empirical objects needed to test the model without using later data vintages or synthetic execution prices. The paper establishes conditional mathematical results and testable implications. It reports no completed backtest or measured profitability.

1 Introduction and mathematical question

A weather observation can alter the conditional law of wheat output without altering the conditional law of an executable futures return. The distinction is foundational. The first object is a physical forecast on a probability space, the second is an equilibrium price response to information and risk, and the third is a self-financing trading decision made at a specific bid or ask after the information becomes available. This paper formulates the three objects jointly and asks when a climate signal can improve each in succession. The resulting framework is a conditional mathematical analysis and a falsifiable empirical program. It does not report a completed backtest or infer realized profitability.

The physical motivation is substantial but does not settle the financial question. Wheat yield responds to heat and moisture in ways that vary with phenological stage, region, and adaptation (Lobell et al., 2011; Asseng et al., 2015; Tack et al., 2015). Remote sensing and reanalysis may help

reconstruct exposure, subject to sampling and revision error (Becker-Reshef et al., 2010; Hersbach et al., 2020). A futures market, however, trades anticipated scarcity, storage, financing, location and quality options, and risk-bearing capacity. The price of storage is itself endogenous (Working, 1949). Therefore the mathematically relevant signal is a revision relative to the information used by the market, and the relevant outcome is the post-release, side-specific, after-cost payoff of a specified contract.

Our deductive claims are deliberately conditional. We prove an information-value identity for crop forecasts, a nonidentification result showing that crop predictability alone cannot identify futures alpha, a local supply-price comparative static, a robust no-trade theorem, and an attenuation formula for noisy signals. We then give the observation scheme needed to evaluate their empirical premises. The detailed intermediate algebra appears in Appendix A. Each proposition identifies which quantities are observable, estimable, or assumed.

2 Probability space, filtrations, and admissible information

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Time is indexed by a calendar-aware filtration $(\mathcal{F}_t)_{t \geq 0}$. Crop year y has terminal output $Q_y \in L^2(\mathbb{P})$. A climate product S_t is an observation with an observation time t^{obs} , public release time t^{pub} , first reliable availability time t^{avail} , research decision time t^{dec} , and earliest admissible fill time t^{fill} . The basic ordering is

$$t^{\text{obs}} \leq t^{\text{pub}} \leq t^{\text{avail}} \leq t^{\text{dec}} < t^{\text{fill}}. \quad (1)$$

The strict final inequality prevents a backtest from buying at a quote that preceded the decision. An observation stamped by physical acquisition but published later is not admissible at the earlier timestamp.

Let $\mathcal{M}_t$ denote the market's public information, including prices and official reports, and $\mathcal{R}_t$ the researcher's reconstructible information. Their distinction matters because an archival file available now may not have been available at the historical t . For a fixed admissible information set $\mathcal{G}_t$, define the crop forecast $m_t = \mathbb{E}_{\mathbb{P}}[Q_y \mid \mathcal{G}_t]$. An innovation is $U_t = m_t - m_{t-}$ for nested information sets $\mathcal{G}_{t-} \subseteq \mathcal{G}_t$. Merely subtracting a climatological average from a weather variable does not make it an innovation relative to market information.

Definition 2.1 (Incremental forecast information). For a square-integrable target X , nested information sets $\mathcal{G} \subseteq \mathcal{H}$ have mean-square information gain $\mathcal{V}_X(\mathcal{H}; \mathcal{G}) = \inf_{g \in L^2(\mathcal{G})} \mathbb{E}[(X - g)^2] - \inf_{h \in L^2(\mathcal{H})} \mathbb{E}[(X - h)^2]$.

Theorem 2.2 (Projection value identity). For $X \in L^2(\mathbb{P})$ and $\mathcal{G} \subseteq \mathcal{H}$,

$$\mathcal{V}_X(\mathcal{H}; \mathcal{G}) = \mathbb{E}[(\mathbb{E}[X \mid \mathcal{H}] - \mathbb{E}[X \mid \mathcal{G}])^2] \geq 0. \quad (2)$$

Equality holds precisely when the conditional means coincide almost surely.

Proof. Conditional expectation is the orthogonal projection in L^2 . Put $m_H = \mathbb{E}[X \mid \mathcal{H}]$ and $m_G = \mathbb{E}[X \mid \mathcal{G}]$. The equality $X - m_G = (X - m_H) + (m_H - m_G)$ has orthogonal terms because $m_H - m_G$ is $\mathcal{H}$ -measurable. Squaring and taking expectations gives $\mathbb{E}[(X - m_G)^2] = \mathbb{E}[(X - m_H)^2] + \mathbb{E}[(m_H - m_G)^2]$. Each conditional expectation minimizes the respective risk. Subtract to obtain (2). $\square$

Corollary 2.3 (Innovation martingale difference). *For $m_t = \mathbb{E}[Q_y \mid \mathcal{G}_t]$ and $\mathcal{G}_{t-} \subseteq \mathcal{G}_t$, $\mathbb{E}[m_t - m_{t-} \mid \mathcal{G}_{t-}] = 0$.*

Proof. Apply the tower property: $\mathbb{E}[m_t \mid \mathcal{G}_{t-}] = \mathbb{E}[Q_y \mid \mathcal{G}_{t-}] = m_{t-}$. $\square$

This theorem concerns a physical target and a specified loss. It neither identifies a price process nor asserts that the signal was unavailable to traders. Even a large $\mathcal{V}_{Q_y}$ can correspond to zero incremental futures-return predictability.

3 The stochastic production layer

Index production regions by $r \in \{1, \dots, R\}$. Harvested area $A_{r,y}$ and yield $Y_{r,y}$ give regional production $Q_{r,y} = A_{r,y}Y_{r,y}$ and aggregate output $Q_y = \sum_r Q_{r,y}$. This is an accounting identity. Both area and yield may respond to drought or policy. For old estimates (A_0, Y_0) and updates $(\Delta A, \Delta Y)$,

$$\Delta Q_r = A_0 \Delta Y + Y_0 \Delta A + \Delta A \Delta Y. \quad (3)$$

The final interaction is exact and can be material for crop abandonment. The joint predictive law of $(A_{r,y}, Y_{r,y})_{r=1}^R$, rather than separate point estimates, determines the law of Q_y . In particular $\mathbb{E}[A_r Y_r] = \mathbb{E}[A_r] \mathbb{E}[Y_r] + \text{Cov}(A_r, Y_r)$ and

$$\text{Var}(Q_y) = \sum_r \text{Var}(Q_{r,y}) + 2 \sum_{r < s} \text{Cov}(Q_{r,y}, Q_{s,y}). \quad (4)$$

Common climate modes and correlated policy shocks make the cross-region terms relevant to tail risk.

For crop stage k , heat and moisture exposures are functionals of a weather path $W_{r,y}(u)$ and uncertain phenology $\phi_{r,y}$, rather than single seasonal averages. For temperature $T(u)$ and stage weight $g_k(u; \phi)$, define $H_k(\phi) = \int g_k(u; \phi) [T(u) - T_k^{\text{crit}}]_+ du$. A corresponding root-zone water deficit is $D_k(\phi) = \int g_k(u; \phi) [W_k^{\text{crit}} - W(u)]_+ du$. The physical aggregation and bucket-water derivation are given in Appendix A. Thresholds and stage weights require calibration; their definition does not prove crop damage.

One estimable family is a partially pooled regional response model

$$\log Y_{r,y} = a_r + g_r(y) - \beta_{H,r}^\top H_{r,y} - \beta_{D,r}^\top D_{r,y} + \gamma_r^\top X_{r,y} + u_{r,y}, \quad a_r \sim N(a_0, \tau_a^2). \quad (5)$$

Its purpose is to represent heterogeneous stages and regions, not to canonize a universal damage coefficient. Sowing dates, irrigation, soil, cultivar, harvested area, policy, and measurement error

must either enter the model, be integrated out, or appear in sensitivity bounds. The predictive distribution is a pushforward of the joint latent law:

$$\mathbb{P}(Q_y \in B \mid \mathcal{G}_t) = \int \mathbf{1}\left\{\sum_r a_r y_r \in B\right\} d\mathbb{P}_t((a_r, y_r)_{r=1}^R), \quad (6)$$

for Borel set B . Equation (6) states exactly what must be propagated to a supply forecast; it does not presume regional independence. It also clarifies why point-prediction improvements may fail to improve a tail-sensitive trading decision.

4 Market filtration, equilibrium sensitivity, and nonidentification

A crop forecast must reach price through a market-clearing and information channel. At a domestic boundary, the accounting identity $B_{y+1} = B_y + Q_y + M_y - C_y - X_y + J_y$ relates ending stocks, beginning stocks, production, imports, use, exports, and a statistical discrepancy. Introduce differentiable use $C(P, Z)$ and desired stocks $B_{y+1}(P, K)$, where Z and K are demand and inventory shifters. Define

$$G(P, Q, Z, K, B_y, M, X, J) = C(P, Z) + B_{y+1}(P, K) - B_y - Q - M + X - J. \quad (7)$$

An equilibrium is a zero of G . The model requires a coherent commodity, geographic boundary, and marketing year; aggregate world imports and exports cancel after reconciliation. A wheat futures quote is additionally affected by maturity, storage, margining, delivery options, and risk premia, so P is initially a local spot scarcity index, not automatically the exchange-traded futures price.

Proposition 4.1 (Local production-price derivative). *Suppose G is continuously differentiable near an equilibrium $(P_0, Q_0, \dots)$ and $G_P(P_0, \dots) = C_P + B_P \neq 0$. Then there is a locally unique differentiable equilibrium branch $P = p(Q, Z, K, B_y, M, X, J)$, with*

$$\frac{\partial p}{\partial Q} = \frac{1}{C_P + B_P}, \quad (C_P + B_P)dP = dB_y + dQ + dM - dX + dJ - C_Z dZ - B_K dK. \quad (8)$$

Proof. The implicit-function theorem applies because $G_P \neq 0$. Differentiate $G(p(Q, \dots), Q, \dots) = 0$ in Q : $G_P p_Q + G_Q = 0$. Since $G_Q = -1$, $p_Q = 1/G_P$. Total differentiation in all arguments and rearrangement yields the second expression. $\square$

If inventories and all other shifters are fixed, $B_P = 0$ and $dP = dQ/C_P$. With demand elasticity $\varepsilon_D = (P/C)C_P < 0$, $dP/P = (1/\varepsilon_D)(dQ/C)$. This is a local comparative static conditional on the exclusions, not a law of futures returns. Inventory release, substitution across grains, imports, currencies, and export restrictions can alter the response. The production forecast distribution therefore enters price only after a behavioral and institutional model has been specified.

The inability to infer trading alpha from crop predictability can be stated as an observational-equivalence result. Let X be a terminal contract payoff in common currency units and S a climate

signal with nonconstant $m(S) = \mathbb{E}[X | S]$. Consider the zero-risk-premium benchmark purely as a mathematical counterexample; actual futures need not satisfy it.

Theorem 4.2 (Physical forecast value does not identify return alpha). *There exist two price-information models with the same joint law of (S, X) and hence the same ability of S to predict the physical terminal payoff, but different conditional expected post-signal returns.*

Proof. In Model A, the market observes S at the entry time and sets $P_t^A = m(S)$. The terminal payoff is X , so $\mathbb{E}[X - P_t^A | S] = m(S) - m(S) = 0$. In Model B, the market does not observe S at that time and sets $P_t^B = \mathbb{E}[X]$. Then $\mathbb{E}[X - P_t^B | S] = m(S) - \mathbb{E}[X]$, which is nonzero with positive probability because $m(S)$ is nonconstant. The distribution of (S, X) is identical in the two constructions. Only the market filtration and entry price differ. $\square$

The counterexample isolates the empirical problem. A regression of yield on weather can validate a physical mechanism while being silent about the speed with which that mechanism was priced. In an actual futures model, conditional risk premia, basis, carry, and transaction costs introduce further distinctions. Under a benchmark $\pi_t = \mathbb{E}[R_{t,h} | \mathcal{M}_t]$, define incremental alpha $\alpha_t = \mathbb{E}[R_{t,h} | \mathcal{M}_t \vee \sigma(S_t)] - \pi_t$. If S_t is already $\mathcal{M}_t$ -measurable, then $\alpha_t = 0$ by equality of sigma-fields. This statement is conditional on the benchmark's information set and does not assert that all publicly released weather is instantly incorporated into all quotes.

5 From equilibrium price to an executable futures position

For a maturity T , a stylized forward carrying value under deterministic financing r , storage cost c , and convenience yield δ is $F_{t,T}^{\text{carry}} = S_t \exp(\int_t^T (r_u + c_u - \delta_u) du)$. It follows by solving $dV_u/du = (r_u + c_u - \delta_u)V_u$ for a held deliverable unit. The equality with a traded futures quote requires further assumptions on margining, stochastic rates, delivery quality and location, and no-arbitrage feasibility. The detailed conditional derivation is in Appendix A. The financial outcome must therefore be defined on the specific contract and its actual quotes, not inferred from a spot comparative static.

Write ask $A_{t,T} = M_{t,T} + s_{t,T}$ and bid $B_{t,T} = M_{t,T} - s_{t,T}$. For multiplier q and fees f , one long contract entered after the signal and liquidated at u has

$$\Pi_{t,u}^L = q(B_{u,T} - A_{t,T}) - f = q(M_{u,T} - M_{t,T} - s_{t,T} - s_{u,T}) - f. \quad (9)$$

A short enters at bid and exits at ask. A roll requires closing one maturity and opening another, and its cash P&L is the sum of the two side-specific legs. The level jump between a nearby and next contract is not itself a trading gain. Equations based on synthetic midpoint series answer a different question from attainable after-cost P&L. Market depth, impact, nonexecution, and margin funding further restrict the feasible set.

Let the conditional gross return mean be uncertain in an interval $[\underline{\mu}_t, \bar{\mu}_t]$, let $v_t > 0$ be a risk estimate, $\lambda > 0$ a risk-aversion parameter, and $c_t \geq 0$ the estimated full linear trading cost per unit

of risk-normalized exposure x . Consider the worst-case one-period objective

$$\sup_{x \in \mathbb{R}} \inf_{\mu \in [\underline{\mu}_t, \bar{\mu}_t]} \left\{ x\mu - \frac{\lambda v_t}{2} x^2 - c_t |x| \right\}. \quad (10)$$

Theorem 5.1 (Robust abstention and position size). *The unique optimizer of (10) is*

$$x_t^* = \begin{cases} (\underline{\mu}_t - c_t)/(\lambda v_t), & \underline{\mu}_t > c_t, \\ (\bar{\mu}_t + c_t)/(\lambda v_t), & \bar{\mu}_t < -c_t, \\ 0, & \text{otherwise.} \end{cases} \quad (11)$$

In particular, a positive mean point estimate does not warrant a long position when its admissible lower bound fails to exceed full costs.

Proof. For $x > 0$, the minimizing mean is $\underline{\mu}_t$, so the objective is the concave quadratic $x(\underline{\mu}_t - c_t) - \lambda v_t x^2/2$. Its positive maximizer exists exactly when $\underline{\mu}_t > c_t$. For $x < 0$, the minimizing mean is $\bar{\mu}_t$, giving $x(\bar{\mu}_t + c_t) - \lambda v_t x^2/2$ and a negative maximizer exactly when $\bar{\mu}_t < -c_t$. Otherwise both one-sided derivatives at zero point toward zero. Strict concavity on each side and continuity complete the argument. $\square$

The interval can represent parameter uncertainty, model ambiguity, or a deliberately conservative forecast bound. It is not a confidence interval unless constructed by a valid inference procedure. With current inventory x_{t-} , replace $c|x|$ by $c|x - x_{t-}|$; the subgradient gives an inaction band around existing holdings. For multiple wheat maturities and related grains, the scalar decision generalizes to a convex program with covariance matrix Σ_t , contract limits, and side-specific cost vector. The multivariate model is meaningful only when synchronized executable quotes and cross-contract margin rules are available.

6 Observation error, vintages, and identification

A historical sample is indexed by admissible decisions, not by retrospective file timestamps. For each candidate record, retain the original source payload, observation interval, public release timestamp, processing timestamp, selected contract, available quote timestamp, and rule version. An as-of join selects the latest datum satisfying (1); later revisions belong to later information sets. Reanalysis is valuable for physical reconstruction but may be revised, whereas an operational decision requires the vintage actually accessible at the decision time. Satellite compositing, cloud screening, geolocation, and crop-mask revisions can similarly create look-ahead contamination if a modern product is treated as an old real-time signal (Hersbach et al., 2020; Becker-Reshef et al., 2010).

Even when timing is correct, noisy signal measurements alter the observed financial relation. The following standard attenuation result quantifies one direction of that error; endogenous measurement error is more difficult.

Proposition 6.1 (Classical measurement-error attenuation). *Assume $R = \alpha + \beta U + \epsilon$, with $\text{Cov}(U, \epsilon) = 0$. The recorded signal is $\tilde{U} = U + \nu$, where ν has mean zero and is independent of (U, ϵ) . Then the population least-squares slope of R on $\tilde{U}$ is*

$$\tilde{\beta} = \beta \frac{\text{Var}(U)}{\text{Var}(U) + \text{Var}(\nu)}. \quad (12)$$

Proof. The population slope is $\text{Cov}(R, \tilde{U}) / \text{Var}(\tilde{U})$. Expanding the numerator and using the stated orthogonalities gives $\text{Cov}(\beta U + \epsilon, U + \nu) = \beta \text{Var}(U)$. Independence gives the denominator $\text{Var}(U) + \text{Var}(\nu)$. Division proves (12). $\square$

This result does not license correcting a null result by arbitrary rescaling. It requires a credible estimate of measurement-error variance and fails when the error is correlated with the target, the market information set, or selection into observable regions. It does, however, show why the physical signal, its historical measurement process, and its release timestamp must be modeled together.

A viable hierarchy of null hypotheses is

$$H_0^Q : \mathcal{V}_{Q_y}(\mathcal{G}_t; \mathcal{G}_{t-}) = 0, \quad (13)$$

$$H_0^R : \mathbb{E}[R_{t,h} \mid \mathcal{M}_t \vee \sigma(S_t)] = \mathbb{E}[R_{t,h} \mid \mathcal{M}_t], \quad (14)$$

$$H_0^\Pi : \sup_{x \in \mathcal{A}_t} \mathbb{E}[\Pi_{t,h}^{\text{net}}(x) \mid \mathcal{J}_t] \leq 0, \quad (15)$$

where $\mathcal{A}_t$ is a contract- and liquidity-constrained feasible set. Rejecting H_0^Q is insufficient for rejecting H_0^R , and rejecting H_0^R is insufficient for rejecting H_0^Π . The ordering is logical, not a license to test whichever hypothesis looks favorable after observing outcomes.

7 A precommitted statistical experiment

For the physical gate, compare the climate-conditioned predictive density with a timestamp-matched baseline such as a seasonal climatology, a public crop forecast, and a persistence model. Proper scores ensure that a model cannot improve merely by becoming unjustifiably sharp. For a binary drought event with true conditional probability p^* , the Brier-risk decomposition is $\mathbb{E}[(p - E)^2 \mid \mathcal{G}] = (p - p^*)^2 + p^*(1 - p^*)$; it is minimized at $p = p^*$. For a full density, expected logarithmic score differs from the true-density score by a nonnegative Kullback–Leibler divergence. Appendix A gives the algebra. Calibration, sharpness, and tail coverage must be assessed together.

For the return gate, fix a region, crop calendar, contract family, forecast lead, horizon, and signal transform before the held-out evaluation. A reduced-form specification may take the form

$$R_{t,t+h}^{\text{gross}} = \alpha + \beta U_t + \theta^\top Z_t + \varepsilon_{t,h}, \quad (16)$$

where U_t is a supply-relevant revision and Z_t consists only of predetermined public controls. Equation (16) is a statistical projection, not a structural supply-demand identity. Contract selection,

overlap, roll timing, basis, public report releases, and contemporaneous risk factors can confound β . A sign-reversed signal, past-return signal, persistence forecast, and shuffled release-time placebo probe different failure mechanisms.

Use an outer expanding or rolling origin by crop year. For each outer fold, fit exposure transformations, hyperparameters, shrinkage, thresholds, and cost models on preceding years only. Freeze one model and evaluate the next untouched year. The chronological split must be imposed at the earliest raw source and in every derived feature. Report the number of independent harvest seasons, not only the number of weekly rows. Regions and horizons chosen after looking at outcomes form one searched family; the best displayed curve is not a prespecified result.

Lemma 7.1 (Overlap-induced covariance). *If independent one-period innovations ϵ_t have common variance σ^2 and h -period returns are $R_t = \sum_{j=0}^{h-1} \epsilon_{t+j}$, then $\text{Cov}(R_t, R_{t+k}) = (h-k)\sigma^2$ for $0 \leq k < h$ and zero for $k \geq h$.*

Proof. The two sums share exactly $h - k$ identical innovations for $k < h$; all other cross-products have zero covariance by independence. For $k \geq h$ there are no shared innovations. $\square$

Actual weather and futures returns may have longer seasonal dependence, so a crop-year block bootstrap or a prespecified long-run variance estimator is preferable to independent-week standard errors. Multiple variants require familywise or false-discovery control. If each of K tests has level at most α/K , the union bound yields $\mathbb{P}(\text{one or more false rejections}) \leq \sum_{k=1}^K \alpha/K = \alpha$. This simple correction may be conservative; a maximum-statistic bootstrap can account for correlated candidate signals if its dependence structure and selection set were fixed in advance (White, 2000; Hansen, 2005).

The net-value gate must rebuild positions from side-specific bid and ask prices, apply commissions and impact assumptions, respect exchange trading hours and contract rolls, and measure realized P&L or utility after costs. A weekly price proxy can provide an exploratory directional test but cannot certify executable net performance without fill-side observations. Sensitivity to cost, latency, crop-area revision, weather-vintage choice, and different harvest regimes should be reported even when the central estimate is favorable.

8 Deductive conclusions, empirical limits, and research agenda

The mathematical results establish conditions, not a profitable strategy. Theorem 2.2 identifies the exact squared-error value of additional information for a physical target. Proposition 4.1 gives a local production-price derivative only after demand and inventory behavior are specified. Theorem 4.2 proves that a climate-to-output relation alone cannot determine return alpha. Theorem 5.1 identifies an after-cost abstention region under a chosen robust quadratic objective. Proposition 6.1 quantifies a particular classical observation-error bias. Lemma 7.1 demonstrates that repeated overlapping return windows do not constitute independent evidence.

The empirical quantities needed to cross from these conditional statements to a trading claim remain unknown here: the vintage-correct distribution of regional production revisions, the price re-

sponse conditional on market information, the relevant risk premium, historical executable quotes, and the stability of all of these across regimes. A publication-quality result would require a frozen source and release-time audit, a nested out-of-sample study, controls and multiplicity adjustment, and reproducible contract-level after-cost accounting. Until those gates pass, the paper’s contribution is a formal theory and a test design rather than measured alpha. That distinction is part of the mathematical conclusion, not a disclaimer external to it.

References

- Asseng, S., Ewert, F., Martre, P., et al. (2015). Rising temperatures reduce global wheat production. *Nature Climate Change*, 5:143–147.
- Becker-Reshef, I., Vermote, E., Lindeman, M., and Justice, C. (2010). A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. *Remote Sensing of Environment*, 114(6):1312–1323.
- CME Group (2025). Chicago SRW wheat futures contract specifications. <https://www.cmegroup.com/markets/agriculture/grains/wheat.contractSpecs.html>.
- Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business and Economic Statistics*, 23(4):365–380.
- Hersbach, H., Bell, B., Berrisford, P., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146:1999–2049.
- Lobell, D. B., Schlenker, W., and Costa-Roberts, J. (2011). Climate trends and global crop production since 1980. *Science*, 333(6042):616–620.
- Tack, J., Barkley, A., and Nalley, L. L. (2015). Effect of warming temperatures on US wheat yields. *Proceedings of the National Academy of Sciences*, 112(22):6931–6936.
- White, H. (2000). A reality check for data snooping. *Econometrica*, 68(5):1097–1126.
- Working, H. (1949). The theory of price of storage. *American Economic Review*, 39(6):1254–1262.

A Mathematical foundations and complete conditional derivations

This appendix derives the mathematical claims used in the research design. Each derivation is conditional on its stated assumptions. Weather damage functions, demand curves, and price laws are empirical objects: mathematics derives their implications but cannot establish their parameters without data. We distinguish exact identities, model-dependent results, and first-order approximations.

A.1 Notation, timing, and units

Work on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $(\mathcal{F}_t)_{t \geq 0}$. Let $\mathcal{G}_{t-}$ be information genuinely available immediately before a signal is released and $\mathcal{G}_t = \mathcal{G}_{t-} \vee \sigma(S_t)$ afterward. Availability includes publication and processing latency. Let $\mathcal{F}_t^M$ denote public market information. Area A is in hectares, yield Y in tonnes per hectare, and production $Q = AY$ in tonnes. A futures quote is currency per bushel and multiplier q is bushels per contract. Quantities must be converted to consistent units before a supply-to-price regression.

A.2 Conditional expectation as the least-squares forecast

Assume $Q \in L^2(\mathbb{P})$ and set $m = \mathbb{E}[Q \mid \mathcal{G}]$. Conditional expectation satisfies $\mathbb{E}[(Q - m)h] = 0$ for every square-integrable $\mathcal{G}$ -measurable h . For any competing $\mathcal{G}$ -measurable forecast g , expand

$$(Q - g)^2 = ((Q - m) + (m - g))^2 \quad (17)$$

$$= (Q - m)^2 + (m - g)^2 + 2(Q - m)(m - g). \quad (18)$$

Taking expectations and setting $h = m - g$ makes the cross term zero. Hence

$$\mathbb{E}[(Q - g)^2] = \mathbb{E}[(Q - m)^2] + \mathbb{E}[(m - g)^2] \geq \mathbb{E}[(Q - m)^2]. \quad (19)$$

Equality holds exactly when $g = m$ almost surely. This establishes optimality under squared-error loss, not under every economic loss. Now let $m_0 = \mathbb{E}[Q \mid \mathcal{G}_{t-}]$ and $m_1 = \mathbb{E}[Q \mid \mathcal{G}_t]$. The tower property and nested information sets give

$$\mathbb{E}[m_1 - m_0 \mid \mathcal{G}_{t-}] = \mathbb{E}[\mathbb{E}[Q \mid \mathcal{G}_t] \mid \mathcal{G}_{t-}] - m_0 = m_0 - m_0 = 0. \quad (20)$$

Thus $\Delta m = m_1 - m_0$ is an innovation relative to the old information set. Apply (19) with $\mathcal{G} = \mathcal{G}_t$ and $g = m_0$ to obtain the exact ex ante value identity

$$\mathbb{E}[(Q - m_0)^2] - \mathbb{E}[(Q - m_1)^2] = \mathbb{E}[(m_1 - m_0)^2] \geq 0. \quad (21)$$

A hot day has no innovation value if its occurrence was already known or fully forecast by the chosen prior information set.

A.3 Gaussian signal update, step by step

For an interpretable example, assume $Y \mid \mathcal{G}_{t-} \sim N(m_0, v_0)$ and $S = Y + \eta$, where $\eta \sim N(0, \tau^2)$ is conditionally independent of Y . Bayes' rule gives

$$p(y \mid s) \propto \exp \left[-\frac{(y - m_0)^2}{2v_0} - \frac{(s - y)^2}{2\tau^2} \right]. \quad (22)$$

Collecting the coefficients of y^2 and y in the exponent gives

$$-2 \log p(y \mid s) = (v_0^{-1} + \tau^{-2})y^2 - 2(m_0 v_0^{-1} + s \tau^{-2})y + \text{constant}. \quad (23)$$

Completing the square yields

$$v_1 = (v_0^{-1} + \tau^{-2})^{-1} = \frac{v_0 \tau^2}{v_0 + \tau^2}, \quad (24)$$

$$m_1 = v_1(m_0/v_0 + s/\tau^2) = m_0 + \frac{v_0}{v_0 + \tau^2}(s - m_0). \quad (25)$$

The gain $K = v_0/(v_0 + \tau^2)$ lies between zero and one. The variance reduction is $v_0 - v_1 = v_0^2/(v_0 + \tau^2)$. This equals the expected square of the revision because $\text{Var}(S - m_0) = v_0 + \tau^2$ and $\text{Var}(K(S - m_0)) = v_0^2/(v_0 + \tau^2)$. Gaussianity and independent additive noise are assumptions of this example.

A.4 Physical forecast value versus return alpha

Let $R_{t,h}$ be the gross return on an explicitly specified futures position entered after the signal. Define the public benchmark mean $\pi_t = \mathbb{E}[R_{t,h} \mid \mathcal{F}_t^M]$ and incremental alpha

$$\alpha_t = \mathbb{E}[R_{t,h} \mid \mathcal{F}_t^M \vee \sigma(S_t)] - \pi_t. \quad (26)$$

If S_t is already measurable with respect to $\mathcal{F}_t^M$, the joined information set equals $\mathcal{F}_t^M$, so $\alpha_t = 0$ by substitution. A nonzero public risk premium may still make $\pi_t \neq 0$; that is not an incremental signal edge. The projection proof for crop production therefore does not prove a futures-return edge.

A.5 Thermal-time and root-zone water exposure

For hourly temperature $T(u)$, stage weight $g_k(u; \phi) \in [0, 1]$, and heat threshold T_k^{crit} , define dose

$$H_k(\phi) = \int_{\mathcal{S}} g_k(u; \phi) [T(u) - T_k^{\text{crit}}]_+ du, \quad [x]_+ = \max(x, 0). \quad (27)$$

A partition with step Δu gives the Riemann approximation $H_k(\phi) \approx \sum_{j=1}^J g_k(u_j; \phi) [T(u_j) - T_k^{\text{crit}}]_+ \Delta u$. Replacing hourly temperature with daily maxima changes the estimand unless a within-day path is assumed. Frost dose uses $[T_k^{\text{frost}} - T(u)]_+$. Stage timing ϕ is uncertain.

Conservation of water in a root-zone bucket gives $W_{u+1}^{\text{raw}} = W_u + P_u + I_u - ET_u - R_u$, where precipitation P and irrigation I add water and evapotranspiration ET and runoff or drainage R remove it. Physical bounds give

$$W_{u+1} = \min\{W_{\max}, \max\{0, W_{u+1}^{\text{raw}}\}\}. \quad (28)$$

A deficit dose is $D_k = \int g_k(u; \phi)[W_k^{\text{crit}} - W_u]_+ du$. Unmeasured irrigation and soil depth make it a latent covariate. Let $Z = (\phi, I, W_0)$. The law of total variance gives

$$\text{Var}(Y \mid \mathcal{G}) = \mathbb{E}[\text{Var}(Y \mid \mathcal{G}, Z) \mid \mathcal{G}] + \text{Var}(\mathbb{E}[Y \mid \mathcal{G}, Z] \mid \mathcal{G}). \quad (29)$$

To derive it, write $Y - \mathbb{E}[Y \mid \mathcal{G}] = (Y - \mathbb{E}[Y \mid \mathcal{G}, Z]) + (\mathbb{E}[Y \mid \mathcal{G}, Z] - \mathbb{E}[Y \mid \mathcal{G}])$. Square, condition on $\mathcal{G}$, and note that the cross term is zero because the first parenthesis has conditional mean zero given $(\mathcal{G}, Z)$. Plugging in one stage date or irrigation estimate discards the second nonnegative term.

A.6 Yield response is an empirical hypothesis

One possible stage-specific model is

$$\log Y_{r,y} = a_r + b_r y - \sum_k \beta_{H,k} H_{r,y,k} - \sum_k \beta_{D,k} D_{r,y,k} - \sum_k \beta_{HD,k} H_{r,y,k} D_{r,y,k} + \gamma^\top X_{r,y} + u_{r,y}. \quad (30)$$

Within this model and holding other covariates fixed, the chain rule $dY = Y d \log Y$ yields

$$\frac{\partial Y}{\partial H_k} = Y \frac{\partial \log Y}{\partial H_k} = -Y(\beta_{H,k} + \beta_{HD,k} D_k). \quad (31)$$

This does not prove the coefficients are positive, stable, causal, or transferable. Adaptation, irrigation, cultivar, and measurement require empirical treatment.

A.7 Area, yield, and production uncertainty

Let $A_1 = A_0 + \Delta A$ and $Y_1 = Y_0 + \Delta Y$. Multiplying gives the exact revision

$$Q_1 - Q_0 = (A_0 + \Delta A)(Y_0 + \Delta Y) - A_0 Y_0 \quad (32)$$

$$= A_0 \Delta Y + Y_0 \Delta A + \Delta A \Delta Y. \quad (33)$$

The first two terms alone are a first-order approximation, with omitted error $\Delta A \Delta Y$. Write $A = \bar{A} + (A - \bar{A})$ and $Y = \bar{Y} + (Y - \bar{Y})$, multiply and take expectations to obtain $\mathbb{E}[AY] = \bar{A}\bar{Y} + \text{Cov}(A, Y)$. Separate area and yield forecasts cannot be multiplied without considering covariance.

Set $A = a + \epsilon_A$, $Y = y + \epsilon_Y$ with centered errors. Retaining the linear terms gives $Q - ay \approx$

$y\epsilon_A + a\epsilon_Y$. Expanding the variance yields

$$\text{Var}(Q) \approx y^2 \sigma_A^2 + a^2 \sigma_Y^2 + 2ay\sigma_{AY}. \quad (34)$$

For jointly Gaussian errors, odd centered moments vanish. The centered joint moment-generating function is $M(t, u) = \exp[(\sigma_A^2 t^2 + 2\sigma_{AY}tu + \sigma_Y^2 u^2)/2]$. Differentiating twice in each argument at zero gives $\mathbb{E}[\epsilon_A^2 \epsilon_Y^2] = \sigma_A^2 \sigma_Y^2 + 2\sigma_{AY}^2$. Since $\mathbb{E}[\epsilon_A \epsilon_Y] = \sigma_{AY}$, the variance of their product is $\sigma_A^2 \sigma_Y^2 + \sigma_{AY}^2$. Substitution into $AY = ay + y\epsilon_A + a\epsilon_Y + \epsilon_A \epsilon_Y$ gives the exact correction

$$\text{Var}(AY) = y^2 \sigma_A^2 + a^2 \sigma_Y^2 + 2ay\sigma_{AY} + \sigma_A^2 \sigma_Y^2 + \sigma_{AY}^2. \quad (35)$$

For $Q = \sum_r Q_r$, expand the centered square to get $\text{Var}(Q) = \sum_r \text{Var}(Q_r) + 2 \sum_{r < s} \text{Cov}(Q_r, Q_s)$. Spatially correlated shocks cannot be discarded because regional models were fit separately.

A.8 Physical stocks and local price response

For one geographic boundary and marketing year, conservation of wheat gives

$$B_{y+1} = B_y + Q_y + M_y - C_y - X_y + J_y, \quad (36)$$

where J explicitly records statistical discrepancy. World imports and exports should cancel after reconciliation. Let use be $C(P, Z)$ and desired ending stocks $B_{y+1}(P, K)$. Rearranging the balance yields

$$C(P, Z) + B_{y+1}(P, K) = B_y + Q + M - X + J. \quad (37)$$

Take a total differential at a reference equilibrium:

$$(C_P + B_P)dP = dB_y + dQ + dM - dX + dJ - C_Z dZ - B_K dK. \quad (38)$$

When $C_P + B_P \neq 0$, division gives the local comparative static. If stocks, trade, demand shifts, and discrepancy are fixed, $dP = dQ/C_P$. Define elasticity $\varepsilon_D = (P/C)C_P < 0$. Dividing by P yields

$$\frac{dP}{P} = \frac{1}{\varepsilon_D} \frac{dQ}{C}. \quad (39)$$

This is a restrictive partial derivative. With price-responsive storage, the denominator is $C_P + B_P$, and policy, trade, substitution, and currency can alter the observed futures response.

A.9 Derivation and limits of the carrying relation

Buy one unit at spot price S_t . Assume deterministic continuously compounded financing r_u , proportional physical carrying cost c_u , and a valued convenience flow δ_u through maturity T . Ignoring delivery options and other frictions, carrying value satisfies $dV_u/du = (r_u + c_u - \delta_u)V_u$ with $V_t = S_t$.

Separate variables:

$$\int_{S_t}^{V_T} \frac{dV}{V} = \int_t^T (r_u + c_u - \delta_u) du. \quad (40)$$

Integrating and exponentiating yields

$$V_T = S_t \exp\left(\int_t^T (r_u + c_u - \delta_u) du\right). \quad (41)$$

Cash-and-carry replication identifies this with a forward delivery price only when the assumptions and arbitrage trades hold. Equating it with a real futures quote additionally requires handling daily margin and stochastic rates. Delivery grade, location, quality options, seasonal scarcity, and constraints can break the equality. For a fixed maturity in the stylized model, $d \log F_{t,T} = d \log S_t + d[\int_t^T (r_u + c_u - \delta_u) du]$; a production shock can also change the carry component.

A.10 Bid, ask, fees, and contract rolls

Write ask $A_{t,T} = M_{t,T} + s_{t,T}$ and bid $B_{t,T} = M_{t,T} - s_{t,T}$ for midpoint M and half-spread s . A long contract bought at t and sold at u has

$$\Pi_{t,u}^L = q(B_{u,T} - A_{t,T}) - f_{t,u} \quad (42)$$

$$= q(M_{u,T} - M_{t,T} - s_{u,T} - s_{t,T}) - f_{t,u}. \quad (43)$$

The first line is executable-side accounting; the second follows by substitution. A short sells at bid and covers at ask: $\Pi_{t,u}^S = q(B_{t,T} - A_{u,T}) - f_{t,u}$. Both assume sufficient depth. Impact, partial fills, limit-order nonexecution, and margin funding need additional treatment.

Rolling a long at v from T_1 to T_2 gives

$$\Pi_{t,u}^{\text{roll}} = q(B_{v,T_1} - A_{t,T_1}) + q(B_{u,T_2} - A_{v,T_2}) - f_{t,v,u}. \quad (44)$$

There are four side-specific prices. The level difference $M_{v,T_2} - M_{v,T_1}$ is not itself roll profit. A stitched continuous series may be useful for charts yet fail to represent attainable cash P&L. Report money, bushels, contract count, and the denominator of any percentage return.

A.11 Quadratic trading objective and no-trade region

Let signed risk-normalized exposure be x , conditional gross mean μ , variance $v > 0$, risk aversion $\lambda > 0$, and full entry-exit cost $c \geq 0$. Consider $J(x) = x\mu - \lambda vx^2/2 - c|x|$. For $x > 0$, the first-order condition $\mu - \lambda vx - c = 0$ gives $x = (\mu - c)/(\lambda v)$, feasible only when $\mu > c$. For $x < 0$, $\mu - \lambda vx + c = 0$ gives $x = (\mu + c)/(\lambda v)$, feasible only when $\mu < -c$. At zero, the subgradient of $c|x|$ is $c[-1, 1]$, so zero is optimal for $|\mu| \leq c$. Concavity makes the conditions sufficient:

$$x^* = \frac{\text{sgn}(\mu)(|\mu| - c)_+}{\lambda v}. \quad (45)$$

For an existing holding x_- and cost $c|x - x_-|$, the subgradient at x_- gives the inaction band $|\mu - \lambda v x_-| \leq c$. Estimation uncertainty can justify a wider abstention policy, but its width requires calibration.

A.12 Proper scores derived from conditional risk

For a binary event E with true probability $p^* = \mathbb{P}(E = 1 \mid \mathcal{G})$, the expected Brier score of report p is

$$\mathbb{E}[(p - E)^2 \mid \mathcal{G}] = p^*(p - 1)^2 + (1 - p^*)p^2 \quad (46)$$

$$= (p - p^*)^2 + p^*(1 - p^*). \quad (47)$$

The only term depending on p is minimized at $p = p^*$. For a discrete outcome with true probabilities p_i^* and reported $p_i > 0$, expected log loss is

$$-\sum_i p_i^* \log p_i = -\sum_i p_i^* \log p_i^* + \sum_i p_i^* \log(p_i^*/p_i) \quad (48)$$

$$= H(p^*) + D_{\text{KL}}(p^* \| p) \geq H(p^*). \quad (49)$$

To see the inequality, apply $\log z \leq z - 1$ to $z = p_i/p_i^*$, multiply by p_i^* , and sum; $\sum_i p_i = \sum_i p_i^* = 1$. The density version replaces sums with integrals where defined. Better crop scores do not prove trading profit.

A.13 Overlapping returns and multiple testing

Let one-week innovations ϵ_t be independent with variance σ^2 , and define an h -week return $R_t = \sum_{j=0}^{h-1} \epsilon_{t+j}$. For lag $0 \leq k < h$, R_t and R_{t+k} share exactly $h - k$ innovations, so

$$\text{Cov}(R_t, R_{t+k}) = (h - k)\sigma^2. \quad (50)$$

The covariance is zero for $k \geq h$. Thus independent-week standard errors understate uncertainty even in this idealized case. Resample whole crop years or use a prespecified dependence-robust long-run variance, and report the number of independent harvest seasons.

For K tested specifications, each with marginal false-rejection probability at most α/K , the union bound gives

$$\mathbb{P}(\text{at least one false rejection}) \leq \sum_{k=1}^K \mathbb{P}(\text{false rejection of } k) \leq \alpha. \quad (51)$$

Bonferroni needs no independence but can be conservative. A maximum-statistic bootstrap may exploit dependence if resampling respects releases and crop years. All regions, leads, signals, rolls, and costs tried during research belong to the examined family.

A.14 Boundary of the derivation

Equations (19), (21), (33), (36), (43), and (44) are mathematical or accounting identities under their definitions. Equations (30), (39), (41), and (45) depend on the stated statistical, equilibrium, financing, or utility assumptions. None implies observed yield predictability, return predictability, or net profitability. Those claims require point-in-time data, a frozen test, and executable observations.

B Reproducibility and release-time audit

The unit of analysis is a forecast issuance, not a row in a revised historical data table. Each observation must carry a source identifier, geographic footprint, physical interval, issue time, receipt time, quality status, and revision identifier. Store the bytes or immutable hash of every file used to create a signal. At decision time t , the as-of join is

$$X_t(v, r) = X_{j^*}(v, r), \quad j^* = \arg \max_{j: t_j^{\text{avail}} \leq t} t_j^{\text{avail}}. \quad (52)$$

If no admissible j exists, X_t is missing. Forward-filling is allowed only for a variable with a prespecified maximum age, and age is itself a feature. Any imputed value must be reproducible from the training fold and inputs available at t . Timestamp tolerance is zero for scheduled report releases and predeclared for feeds whose delivery metadata are coarse. The audit records every exception.

A candidate event passes the release-time gate only if the climate observation is published, received, transformed, and frozen before the first price quote used for execution. A backtest that uses a close price from the same day as an after-close satellite update fails this gate. For monthly USDA data, use the exact report publication instant; for a multi-day remote-sensing composite, do not use the last image date as a publication proxy. For reanalysis, retain the actual preliminary or forecast vintage, never the later final series, in a confirmatory trading test.

Table 1: Required evidence ledger for every tested signal. Empty fields block an executable claim.

Evidence item	Required stored fields
Climate input	Source, geography, observation interval, issue and receipt times, revision, quality flag
Public baseline	Forecast or report vintage and timestamp, curve and inventory snapshot
Signal	Code hash, model version, training cutoff, feature vector, predicted distribution
Trade	Contract, maturity, bid and ask, size, order time, fill rule, spread and fees
Outcome	Exit quote, roll cash flow, realized cost, benchmark return, event and season ID

C Forecast and trading validation protocol

The outer test is organized by complete growing season. On each outer fold, train on earlier seasons; tune on a strictly earlier inner sequence; issue predictions without revising past signals; then evaluate on the held-out season. Return labels overlapping a training or validation boundary are purged, and an embargo extends beyond the maximum holding horizon. A final holdout remains untouched until all feature definitions, cost assumptions, region weights, thresholds, and hypothesis families are frozen. Spatial generalization is evaluated separately with leave-one-region-out folds because regional crop data can share common weather and policy shocks.

Let $L_{k,t}$ be the proper loss of model k at forecast issuance t and $L_{0,t}$ that of its public-information benchmark. Define $d_{k,t} = L_{0,t} - L_{k,t}$. The primary forecast statistic is $\bar{d}_k$, aggregated first within a harvest season and then across seasons so dense daily issuances do not dominate. Bootstrap resampling occurs at the season or weather-event cluster level and is repeated over the full registered candidate family. Report the family-wise adjusted p -value and the unadjusted estimate with its interval. If the number of independent seasons is too small for a credible interval, the result remains descriptive.

For trading, report net wealth recursively with actual contract counts and cash flows. With n_t contracts and price in dollars per bushel, the gross variation for a price move is $5,000n_t(P_{t+1} - P_t)$ dollars; an exchange tick of 0.0025 dollars per bushel is 12.50 per full-size contract under the current Chicago SRW specification (CME Group, 2025). A roll consists of closing one maturity and opening another, each at its executable side of the market. It is not modeled as an arbitrary subtraction of the futures-price difference. Include commissions, exchange fees, spread, impact, and financing or cash yield according to the stated return convention.

A defensible cost sensitivity reports observed quotes where available and multiplies the estimated impact component by at least 0.5, 1, and 2. Stress missing depth and delayed fills separately. The selected strategy is rejected if its upper credible or confidence bound under the conservative cost case does not clear zero net excess return, or if missing quotes prevent reconstruction. These thresholds must be registered before looking at the final holdout.

D Model complexity and interpretability

The primary yield model is the partially pooled stage-exposure model in (5); the primary price model is (16). Secondary candidates may include penalized regression, a modest tree ensemble, and a sequence model. Their tuning budgets and search spaces are fixed in advance. A large Cartesian hyperparameter grid is not evidence of rigor when the number of independent harvest seasons is small. Model complexity is accepted only if it improves the public baseline on outer folds, maintains forecast calibration, and does not concentrate apparent value in a single event or region.

The final report must show the full candidate count, all failed variants, training windows, predeclared selection rule, seed or deterministic settings, and data-license restrictions. Feature

attributions are descriptive and do not substitute for the release-time and placebo tests. The success claim is bounded to the observed instruments, vintages, costs, and regimes; extrapolation to other wheat classes or future climate regimes requires a new test.