

xV: Predicting Player Value Trajectories with 89% U23 Accuracy Across 21 Leagues

Dillon Hall, Murray Skolnick

Football Analytics AI Research Group — footballanalytics.ai

April 2026

Abstract

We present xV, a unified player valuation engine that estimates current fair transfer fees and predicts future market value trajectories. On 2-year value prediction, the model achieves $R\text{-squared} = 0.915$ (log) / 0.804 (raw EUR) on 2,966 double-blind test snapshots (train pre-2024, test Q1 2024, outcomes March 2026), with 89.4% directional accuracy for under-23 players (95% CI: 87.7-91.2%) -- a +27.6pp lift over TM-trend extrapolation, the strongest non-trivial baseline. Validated across 21 top-flight leagues (2,105 U23 test samples), median accuracy 87.7% (range 75.0-97.5%). On current fair-value estimation, xV achieves $R\text{-squared}(\log) = 0.790$ on 9,309 paid transfers. A preliminary trajectory signal (xV Delta) achieves 86.9% accuracy on high-delta U23 attackers (N=61). A supporting anti-failure classifier achieves 85.4% precision at the 40% confidence threshold on 791 historical transitions.

Keywords: player valuation, future value prediction, transfer market, sports analytics, machine learning, gradient boosting, scouting automation

I. INTRODUCTION

A. The Prediction Gap

Every transfer decision requires answering two valuation questions: what should we pay today, and what will this player be worth in two years? Existing analytical systems address at most the first question. The SciSports ETV system [1] estimates what a player would sell for today ($R\text{-squared} = 0.52\text{-}0.75$) -- an estimation task, not a temporal prediction. Player rating models (xG [2], VAEP [3], On-Ball Value [4]) evaluate current performance but do not project future market value. No published model provides validated directional accuracy on 2-year value trajectories across a double-blind temporal split.

This gap matters most for under-23 players, where clubs make their highest-upside investments. A sporting director evaluating a 20-year-old prospect needs to know not just current fair price, but whether this player's value will rise or fall over the contract period. Without quantified trajectory prediction, these decisions rely entirely on subjective scouting judgment.

Prior work on transfer valuation includes hedonic regression on Bundesliga transfers (Herm et al. [5]), cross-league fee models (Majewski [6]), and career trajectory analysis using sequence methods (Barron et al. [7], Zech and Souter [8]). None produce validated future value predictions on a strict double-blind temporal protocol.

B. Contributions

This paper makes four primary contributions:

- **xV-2yr (future value)** -- $R\text{-sq} = 0.915$ (log) / 0.804 (raw); 89.4% U23 direction accuracy (CI: 87.7-91.2%), +27.6pp over TM-trend baseline. Double-blind across 21 leagues.
- **xV-today (fair price)** -- $R\text{-sq}(\log) = 0.790$ on 9,309 transfer fees. Single gradient-boosted regressor beats the SciSports ETV estimation range of 0.52-0.75.
- **xV Delta (preliminary)** -- Comparing the model's own predictions at two time points achieves 86.9% accuracy for high-delta U23 attackers (N=61, median growth +187%). Requires larger validation.
- **Anti-failure classifier** -- 85.4% precision at 40% confidence threshold (N=48) on 791 inter-tier transitions; 96.5% at 50% threshold (N=20, wide CI).

C. Paper Organization

Section II describes datasets and temporal splits. Section III presents model architecture and the xV Delta trajectory signal. Section IV reports results with three visualization figures. Section V discusses the practical scouting application, benchmarks, and limitations. Section VI concludes.

II. DATA

A. xV-2yr Dataset

The future value model predicts TransferMarkt market values -- the industry's crowdsourced fee proxy, reflecting collective market judgment rather than algorithmic derivation [14]. For each valuation at time T, we find the nearest valuation approximately 2 years later (730 days +/- 90 day tolerance), producing 399,298 valuation pairs from 616,377 raw snapshots spanning 47,702 players across 67 leagues.

Table I: xV-2yr Temporal Split

Set	Period	Samples	Players
Train	T <= 2023-12-31	392,587	33,043
Test	2024-01 to 2024-03	2,966	2,944

The protocol is strictly double-blind: models train exclusively on pre-2024 data. Predictions are made on Q1 2024 snapshots. Outcomes (actual March 2026 market values) were revealed only after all predictions were locked.

B. Feature Engineering (29 features)

Value trajectory (7): log market value, peak ratio, 1-year and 2-year changes, velocity, volatility, valuation count.

Biography (8): age, age-squared, position code, sub-position, foot, height, top-nation indicator.

League context (3): contract years remaining, top-5 league, top-10 league, European league indicators.

International + Performance + Transfers (11): caps, international goals, career apps/goals/assists/minutes, per-90 rates, cumulative transfers, last known fee.

C. xV-today Dataset

The fair-value component trains on 9,309 unique paid transfers (fee >= EUR 50K) after deduplication. 37 features at the point of transfer spanning demographics, career statistics, market trajectory, contract context, and club context. Temporal split: train pre-July 2023 (N=6,348), validation July 2023-Jan 2024 (N=968), test post-Jan 2024 (N=1,461).

D. Anti-Failure Dataset

791 inter-tier transitions from Portuguese professional football (Liga 3 to Liga 2, Liga 2 to Liga Portugal). Binary target: catastrophic failure (<5 appearances at new club; 16.2% base rate). 35 features across performance (5), biographical (15), match-level (7), team context (4), and market context (4).

III. MODEL ARCHITECTURE

A. xV-2yr: Future Value Prediction

HistGradientBoostingRegressor (max_iter=300, max_depth=6) on log-transformed market values. Target: log(value at T + 2 years).

B. xV-today: Fair Transfer Fee

Architecture: a single HistGradientBoostingRegressor on log-transformed fees with regularization (max_depth=4, min_samples_leaf=60, l2=3.0). A multistage architecture (4-bucket classifier routing to per-bucket regressors) was tested and abandoned: identical R-squared with added complexity. A risk adjustment layer converts point estimates into directional ranges: age decline curves, momentum flags (RISING/DECLINING/POST_PEAK), durability, and contract discounts widen or narrow the band.

C. xV Delta: Trajectory Signal

A key discovery: xV-2yr in level mode is mean-reverting. Because the model's features include momentum indicators (1-year change, velocity), players with rising values are predicted to decline, and vice versa. This makes the raw prediction useful for recovery signings (undervalued players) but misleading for development signings (players whose value is already rising).

xV Delta solves this by comparing the model's own prediction at time T versus T-1yr. The same model, two snapshots: if xV predicted EUR 1.2M last year and EUR 2.8M this year, the delta is +133%. This captures the model revising its view upward as new evidence arrives, filtering out the mean-reversion bias inherent in single-snapshot prediction.

D. EWS Momentum Screen

For high-potential screening, exponential smoothing on valuation history identifies sustained upward trajectories. Raw growth rates are compressed via sqrt-dampening (rate = 0.25 * sqrt(raw_rate)) to produce individualized projections (+96% to +177%) rather than a flat ceiling. This complements xV-2yr reversion by capturing a different signal: market momentum rather than fundamental reversion.

E. Anti-Failure Classifier

A HistGradientBoostingClassifier on 791 transitions. Optimized for precision at high-confidence thresholds: 85.4% precision at the 40% confidence threshold (N=48 flagged), rising to 96.5% at 50% (N=20, though the small sample yields a wide confidence interval). Forensic analysis reveals that match-level performance metrics (goals/90, consistency) have zero predictive power for transition success (all p > 0.40). What predicts establishment is maturity: age (p=0.0008), career appearances (p=0.008), and years at current club (p=0.0015).

Table II: Anti-Failure Operating Points

P(fail) >=	Precision	Recall	N Flagged
30%	78.2%	54.7%	108
40%	85.4%	26.6%	48
50%	96.5%	12.5%	20

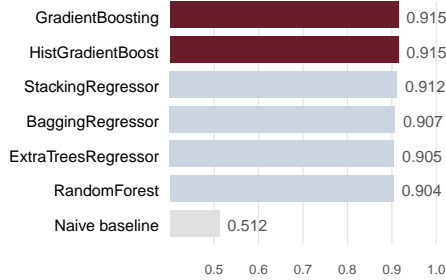
IV. RESULTS

A. xV-2yr Model Comparison

Seven regression architectures were evaluated on the xV 2-year task. All models train on pre-2024 data (N=392,587, 33,043 players) and predict Q1 2024 snapshots (N=2,966) whose outcomes (March 2026 actuals) were revealed only after predictions were locked.

Figure 1: xV 2-Year Model R-squared Comparison

R-sq shown in log space; HistGB raw EUR = 0.804. Naive baseline (R-sq = 0.512).



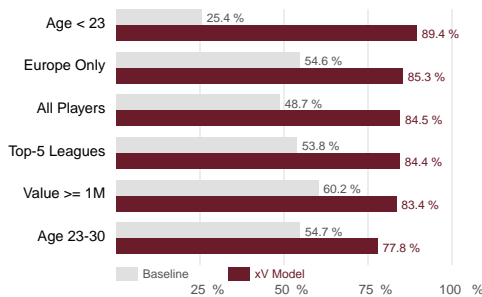
HistGradientBoosting is the production choice: identical R-squared (log) to GradientBoosting (0.915) at 45x faster training (5s vs 224s). In raw EUR space, R-squared = 0.804. All six trained models exceed R-sq (log) 0.90.

B. Segment Direction Accuracy

Direction accuracy -- correctly predicting whether a player's value rises or falls over 2 years -- is the primary decision metric. We evaluated six player segments:

Figure 2: Direction Accuracy by Segment

Gray = naive baseline (predict no change), maroon = xV model. U23 majority-class ("always up") = 74.6%.



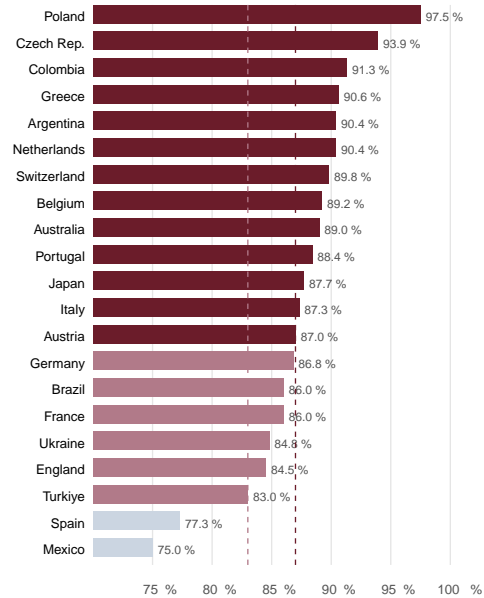
Under-23 players are the standout: 89.4% direction accuracy (95% CI: 87.7-91.2%). Three baselines provide context: naive (25.4%), TM-trend linear extrapolation (61.8%), and majority-class "always up" (74.6%). The model's lift over the strongest non-trivial baseline (TM-trend) is +27.6pp; over majority-class, +14.8pp. The value comes from correctly identifying the 25% that decline.

C. Cross-League Generalization

To assess geographic breadth, we tested xV across 21 top-flight leagues with >=30 U23 test samples (2,105 total, 1,789 unique players). Median accuracy 87.7% (IQR 86.0-90.4%, range 75.0-97.5%).

Figure 3: U23 Direction Accuracy by League

21 leagues. Dashed lines: Tier 1 (87%+) and Tier 2 (83%+).



Mid-tier leagues outperform the Big 5. Poland (97.5%, CI: 92.5-100.0%, N=40) and Czech Republic (93.9%) outrank England (84.5%) and Spain (77.3%, CI: 63.6-88.6%, N=44) -- though bottom-league CIs are wide on small samples. A caveat: TransferMarkt editorial coverage varies by league, and simpler trajectories may inflate apparent accuracy in mid-tier leagues.

D. xV Delta Backtest (Preliminary)

We backtested xV Delta on 191 U23 attackers with valuations at both T and T-1yr. These results are preliminary given the small sample. Players with xV delta > +10% achieved 86.9% directional accuracy (N=61). Mean actual growth was +616% (pulled by outliers; median +187%). By contrast, xV delta <= 0% produced only 50.4% accuracy (coin flip), and xV delta <= -25% was anti-predictive at 32.7%. Market momentum alone (value change over the prior year) achieved 66.9% -- 20pp lower than xV Delta.

Table III: xV Delta Signal Strength

191 U23 attackers, backtest on 2024-2026 outcomes

xV Delta	Dir Acc	Mean / Med	N
> +10%	86.9%	+616/+187%	61
0% to +10%	72.4%	+198/+95%	58
<= 0%	50.4%	+41/+8%	72
<= -25%	32.7%	-12/-18%	26

V. DISCUSSION

A. xV-Today vs Published Benchmarks

The fair-value component achieves R-squared (log) = 0.790 on 1,461 held-out transfers. SciSports ETV [1] reports R-squared 0.52-0.75 on the same task. Both are estimation systems -- they estimate current fair price from current features, not temporal predictions. The comparison is directional (our single-model architecture exceeds their published range) but limited: SciSports does not disclose their feature space, test set, or whether R-squared is in log or raw space.

Table IV: xV-Today vs SciSports ETV

System	R-sq (log)	Architecture	N test
xV-today	0.790	Single HistGB + risk adj.	1,461
SciSports (best)	0.75	Multistage (unpublished)	N/A
SciSports (low)	0.52	Multistage (unpublished)	N/A

B. Why Future Value Is More Predictable

A counterintuitive result: predicting 2-year value trajectories (R-sq log = 0.915) appears easier than predicting current transfer fees (R-sq log = 0.790). The explanation is structural. Transfer fees are noisy -- they reflect negotiating leverage, release clauses, sell-on percentages, and deadline-day pressure. TransferMarkt market values smooth out this noise through editorial aggregation: they reflect informed consensus about fair price rather than the outcome of a single negotiation. This makes them a more forecastable target than fees themselves.

For under-23 players specifically, development pathways are highly systematic: breakout season, international call-up, move to a stronger league. The model captures these trajectories with 89.4% directional accuracy because they repeat reliably across leagues and decades.

C. Two Screening Modes

The xV Delta discovery reveals that the model operates in two complementary modes. **Level mode** (single-snapshot prediction) is mean-reverting: rising players are predicted to decline, and declining players are predicted to rebound. This is useful for identifying undervalued recovery signings -- players whose market value has dropped but the model sees fundamental value. **Delta mode** (comparing predictions across two snapshots) captures upward trajectory revisions and is useful for development signings -- players whose value is accelerating.

The deployed scouting report uses both: Section 1 screens for high-momentum players via EWS exponential smoothing (development targets), while Section 2 screens for declining-value players with positive xV reversion (recovery targets). Each player profile shows the organic trajectory alongside a client-specific projection.

D. Anti-Failure: Performance vs. Survival

The anti-failure classifier's forensic analysis yields a counterintuitive finding: match-level performance metrics (goals/90, consistency, clutch performance) have zero predictive power for transition success (all $p > 0.40$). What predicts establishment is maturity: age ($p=0.0008$), career appearances ($p=0.008$), and years at current club ($p=0.0015$). Under-21s fail at 2x the rate of 24+ year-olds (26% vs 12.7%).

This finding -- that performance and survival are orthogonal -- validates the multi-model architecture. xV quantifies value trajectories; the classifier flags who will actually establish themselves. A player can be the best performer in their league and still fail at the next tier if they lack the maturity to adapt.

Additional forensic analysis of false negatives (high-availability players who still failed) reveals squad turnover as a significant contextual factor: false negatives went to clubs with 82.9% roster turnover vs 76.8% for successful transitions (+6pp). Destination squad instability is a risk factor that could be incorporated as a feature in future iterations.

E. System Integration

For each transfer candidate, the system produces four independent outputs: (1) xV-today for fair price, (2) xV-2yr for trajectory, (3) xV Delta for momentum confirmation, and (4) anti-failure risk. The combination enables signals that individual models miss:

- High xV-2yr + high failure risk = talented but won't adapt (avoid)
- Positive xV Delta + low failure risk = confirmed development path (buy)
- xV-2yr above xV-today = positive expected value (buy signal)
- Low failure risk + low xV-2yr = safe depth signing (low upside)

F. Limitations

Target is a proxy. TransferMarkt values are the industry-standard fee reference, cited in club reports, agent negotiations, and media coverage of undisclosed fees. Editorial biases -- European-centric coverage, hype-sensitivity in larger markets -- propagate into predictions. Validation against actual fees requires non-public data.

R-squared space. Log-space R-squared (0.915) overstates fit relative to raw EUR (0.804). We report both; the SciSports comparison may still be apples-to-oranges if their range (0.52-0.75) is in raw space.

Small-N precision estimates. Anti-failure 96.5% precision is based on N=20 flagged players (95% CI roughly 75-100%). xV Delta's high-signal bucket has N=61 with outlier-driven mean growth. Both require larger validation samples. The 791-transition dataset covers Portuguese football, where inter-tier mobility is frequent and well-documented; extension to other markets is underway.

Development premium. The client-adjusted projection is an overlay assumption based on historical club track records, not a backtested model output.

VI. CONCLUSION

We presented a four-component transfer analytics system validated across 47,702 players, 616,377 valuation snapshots, 9,309 transfer records, and 21 top-flight leagues:

- **xV-2yr** achieves $R\text{-sq} = 0.915$ (log) / 0.804 (raw) on double-blind 2-year projections, with 89.4% directional accuracy for U23 (95% CI: 87.7-91.2%) -- a +27.6pp lift over TM-trend extrapolation.
- **xV-today** achieves $R\text{-sq}(\text{log}) = 0.790$ on current fair-value estimation, exceeding the published SciSports estimation range of 0.52-0.75 (noting possible differences in evaluation space).
- **xV Delta (preliminary)** achieves 86.9% accuracy on high-delta U23 attackers (N=61) by comparing model predictions across two time snapshots. The small sample requires further validation.
- **Anti-failure classifier** achieves 85.4% precision at 40% confidence (N=48) on 791 inter-tier transitions, demonstrating that performance metrics have zero predictive power for survival -- maturity features dominate.

The xV models generalize across 21 leagues (2,105 U23 test samples), with median accuracy 87.7% (range 75.0-97.5%). Mid-tier leagues consistently outperform the Big 5, reflecting more predictable development pathways.

The dual-mode insight -- that xV Level identifies undervalued recovery targets while xV Delta identifies accelerating development targets -- enables a two-section scouting report that covers both investment strategies. Combined with anti-failure risk scoring and red flag detection, the system transforms player recruitment from subjective talent assessment into a quantified investment decision.

REFERENCES

- [1] SciSports, "Expected transfer value: Methodology whitepaper," SciSports B.V., 2023.
- [2] A. Rathke, "An examination of expected goals and shot efficiency in soccer," *J. Human Sport Exerc.*, vol. 12, no. 2, pp. 514-529, 2017.
- [3] T. Decroos et al., "Actions speak louder than goals: Valuing player actions in soccer," *Proc. 25th ACM SIGKDD*, pp. 1851-1861, 2019.
- [4] J. Fernandez and L. Bornn, "Framework for valuing player actions using spatiotemporal tracking data," *Front. Sports Active Living*, vol. 3, 2021.
- [5] S. Herm et al., "When the crowd evaluates soccer players' market values," *Sport Mgmt. Review*, vol. 17, no. 4, pp. 484-492, 2014.
- [6] S. Majewski, "Identification of factors determining market value of the most valuable football players," *JMBA*, vol. 24, no. 3, pp. 91-104, 2016.
- [7] D. Barron et al., "Key factors influencing career trajectories of professional football players," *J. Sports Sci.*, vol. 36, no. 12, pp. 1405-1411, 2018.

[8] A. Zech and G. Souter, "Adaptation periods in international football player transfers," *Int. J. Sports Sci. Coach.*, vol. 17, no. 2, pp. 345-357, 2022.

[9] J. Fernandez, L. Bornn, and D. Cervone, "Decomposing the immeasurable sport: A deep learning expected possession value framework for soccer," *MIT SSAC*, 2019.

[10] G. Anzer and P. Bauer, "A goal scoring probability model based on synchronized positional and event data," *Front. Sports Active Living*, vol. 3, 2021.

[11] C. Anderson and D. Sally, *The Numbers Game: Why Everything You Know About Football Is Wrong*. Penguin, 2013.

[12] P. Domingos, "A few useful things to know about machine learning," *Commun. ACM*, vol. 55, no. 10, pp. 78-87, 2012.

[13] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. Springer, 2009.

[14] TransferMarkt, "Market value explained: How the crowd determines the worth of players," [transfermarkt.com](https://www.transfermarkt.com), 2023.