

METHODOLOGY DOCUMENTATION

Carbon Credit Price Estimation Methodology

Machine-learning based price estimation for projects and individual credit transactions

DATE

September 2026

TYPE

SQL + Python

Overview

AlliedOffsets estimates a fair market price for every carbon credit project, and separately for individual credit transactions, using a machine-learning model trained on real broker-reported and marketplace pricing data. Rather than relying on a single average, the model blends two independent estimation techniques and then applies a structured set of registry-, sector- and compliance-specific business rules on top, to capture known pricing dynamics that a purely statistical model would otherwise miss — such as CORSIA eligibility, registry-specific pricing patterns, and known individual outlier projects. This produces the estimated price shown against every project and the estimated price attached to every individual credit, which together feed AlliedOffsets' price explorer, portfolio valuations, and index products.

Data Inputs

Field / Parameter	Description	Source / Notes
Broker-reported prices	Bid, ask and actual transaction prices submitted by carbon credit brokers, collected on a recurring schedule.	Broker price submissions
Marketplace sample prices	Live listed prices scraped from an online carbon credit marketplace.	Emsurge marketplace
Project reference data	Registry, country, sector, sub-sector, methodology, project status and estimated annual emission reduction for every project.	AlliedOffsets project database (project_all)
Credit-level reference data	Vintage year, issuance and retirement history, transaction volume, first/last issuance year, and retirement-related metrics for every credit.	AlliedOffsets credit database (credit_all)
Compliance eligibility flags	CORSIA phase, Article 6 authorization, Core Carbon Principles (CCP) approval, and jurisdiction-specific compliance scheme eligibility for each project.	AlliedOffsets compliance tagging (see CORSIA Phase & Article 6 Authorization Tagging Methodology)
Sustainable Development Goal (SDG) tags	Which of the 17 UN SDGs a project claims to support.	AlliedOffsets project database
T20 index membership list	A maintained list of the most liquid, actively-traded projects.	AlliedOffsets T20 liquidity index
Woodland-specific price estimates	Output of a separate, purpose-built price model used to override prices for Woodland-registry projects.	Woodland price estimates companion model

Field / Parameter	Description	Source / Notes
Carbon-removal manual price overrides	Analyst-curated override prices for select carbon-removal projects.	Maintained override spreadsheet
Manual price override list	A maintained list of specific projects whose price is set directly by an analyst rather than estimated by the model.	Maintained internal reference file

Logical Flow

1

Assemble a training dataset from recent real transactions

Combine broker-reported and marketplace prices from the trailing six months; for any project not represented in that recent window, fall back to its single most recent price from the prior six months. Where a marketplace price and a broker price cover the same project and vintage in the recent window, the marketplace price is dropped in favor of the broker price.

2

Normalize bid, ask and actual prices onto a common basis

Since brokers report several different transaction types, calculate each broker's typical premium between actual sale prices and their bid/ask quotes, using only brokers that report all three types, and apply that premium to convert every price observation onto a consistent offer-price basis before it is used to train the model.

3

Train two independent price-estimation models

Fit one gradient-boosting model and one neural network model on the same set of project and credit characteristics — registry, country, sector, methodology, compliance flags, SDG tags, volume, vintage, issuance history and more — each predicting a price from those characteristics. Average the two models' predictions together as a blended baseline estimate.

4

Route each project or credit to the best-suited estimation method

Rather than using the same blended estimate everywhere, apply a defined decision path based on sector, registry and compliance status to decide whether to use the gradient-boosting estimate alone, the neural network estimate alone, or the blended average — since certain project types are known to be better represented by one technique than the other.

5

Apply registry- and scheme-specific business rule overrides

Layer on a series of known, fixed adjustments: force specific registries to a fixed observed market price where one exists; nudge Biochar-sector prices toward the latest Biochar market index; lift CORSIA-eligible credits that would otherwise be priced unrealistically low; exclude one registry's non-Biochar credits from pricing entirely, since suitable input data doesn't exist for them; split pricing up or down for projects issuing both emission-reduction and emission-removal credit types; and apply specific overrides requested for a handful of individual projects.

6

Substitute a specialist sub-model's output for one registry

For projects belonging to a specific forestry registry, replace the general model's price with the output of a separate, purpose-built price model built specifically for that registry.

7

Correct for vintage-price inversions

Since later vintages of an otherwise comparable project should not typically be priced lower than earlier vintages, group projects into comparable clusters and, wherever a later vintage's price is lower than an earlier one's, adjust it upward by a small fixed step, cascading the correction forward through any subsequent vintages that would otherwise remain inverted.

8

Apply a retirement volume and timing discount

For credits that have already been retired or cancelled, rather than newly issued and unretired, adjust the estimated price up or down depending on how the retired volume compares to the typical volume for similar credits, capped at a fixed maximum adjustment in either direction.

9

Apply manual overrides for specific external pricing sources

Fully replace the model's estimate with an analyst-curated manual price wherever one has been set for a project — used for select carbon-removal projects with dedicated market pricing — and similarly override for a maintained list of specific projects tracked in a manual reference file.

10

Guard against zero or negative prices

If any final estimated price ends up at or below a negligible threshold, replace it with the smallest positive price present elsewhere in that run's results, rather than publishing an unrealistic near-zero or negative price.

11

Publish the results

Write the resulting price per project to the project-level price field, write the average resulting price per credit (accounting for any duplicate estimates for the same credit) to the credit-level price field, and append a running weekly snapshot of prices per project and vintage to the shared pricing table that powers the price explorer, portfolio valuations and index products.

Key Calculations

Blended baseline estimate

The simple average of the gradient-boosting model's predicted price and the neural network model's predicted price for the same project or credit.

Broker offer-conversion premium

For each broker that reports bid, ask and actual transaction prices, a weighted blend of that broker's average and median premium between actual and quoted prices — weighted 75% toward the average and 25% toward the median — applied to convert bid/ask quotes onto an equivalent offer-price basis.

Vintage-inversion adjustment

For each cluster of comparable projects (grouped by shared registry, country, sector, sub-sector, offset type, methodology, verification status, compliance flags and vintage), where a later vintage's typical price is lower than an earlier vintage's, the later vintage is pushed up to 1% above the earlier vintage's adjusted price. This correction cascades forward through any subsequent vintages that would otherwise remain inverted.

Retirement volume and timing discount

A logarithmic adjustment based on the ratio between a credit's actual retired volume and the typical volume for comparable credits, capped at plus or minus 7.5% of the credit's base estimated price.

CORSIA low-price floor

For CORSIA-eligible credits from vintage year 2021 onward, if the estimated price would fall below \$10, a fixed \$5 uplift is added.

Biochar index alignment

For Biochar-sector credits, the neural-network price estimate is moved 20% of the way toward a defined band around the most recent Biochar market index price, keeping the model's estimate broadly anchored to observed Biochar market pricing without fully overriding it.

Assumptions

- The model is retrained on a rolling six-month window of real transaction data, with a secondary six-to-twelve-month lookback used only to fill in a single most-recent price for projects with no very recent activity; a project with no transaction history at all in the past year relies entirely on the model's general characteristics-based prediction rather than its own price history.
- Marketplace sample prices are treated as lower priority than broker-reported prices: where both exist for the same project and vintage within the recent window, the broker price is used and the marketplace price is dropped.
- One broker's price data is excluded from training unless that broker is a project's only available price source, reflecting an analyst judgment that this broker's pricing is less reliable as a general input.
- A small number of registries are assigned fixed, hardcoded prices (for example, \$30 for one registry and roughly \$30.43 for another) rather than a model-derived estimate, reflecting that these registries' credits are understood to trade in a narrow, well-known price band.
- One registry (Acorn) is always priced using the neural-network estimate rather than the blended or gradient-boosting estimate, with additional fixed downward adjustments applied above certain price and vintage-age thresholds.
- Credits from one registry (Puro.earth) are excluded from pricing entirely outside the Biochar sector, because the available training data for that registry only meaningfully covers Biochar-type credits.
- A small number of individual projects have their pricing method fixed to a specific model output at explicit client request, overriding the otherwise sector-driven routing logic.
- For projects issuing both emission-reduction and emission-removal credit types from the same underlying activity, the reduction-type credits are priced 20% below, and the removal-type credits 20% above, the neural-network estimate — reflecting a market premium for removal credits over avoidance/reduction credits from the same project.
- A defined list of especially liquid, frequently-traded projects is always priced using the neural-network estimate rather than the routed default, on the basis that these projects have enough real transaction history for that estimate to be trusted directly.

Limitations

- The model does not incorporate a single very recent transaction in isolation; a project's price only updates meaningfully once enough recent transaction data accumulates in the rolling window, so a sudden one-off deal can take time to be reflected.
- A known issue exists in the fallback logic for one forestry registry (Woodland): where that registry's specialist companion model doesn't produce an estimate for a specific project, the intended fixed default price is not currently being applied due to a coding error, meaning those specific projects silently keep whatever the general model produced instead of the intended fallback value. This should be corrected before being relied upon for that registry's pricing.

- The specialist companion model used for Woodland-registry projects is itself trained in a way that uses each project's known price as an input to help estimate that same project's price, so it can only really be validated for projects that already have observed pricing; its vintage-year and per-method adjustment coefficients are not derived from any logic visible in this pipeline and should be periodically reconfirmed with whoever maintains that model.
- The routing logic that decides which model (or blend) to use for a given project is a fixed decision tree based on sector, registry and compliance status; it is not re-evaluated statistically and may not adapt automatically if a sector's underlying price dynamics shift.
- The vintage-inversion correction assumes that, within a cluster of otherwise-comparable projects, later vintages should never be cheaper than earlier ones; in cases where a genuine market reason exists for a later vintage to be cheaper, this correction will overwrite that true observed pattern.
- Manual overrides — for specific carbon-removal projects and a maintained reference list — take absolute precedence over the model's estimate; the accuracy of those specific projects' prices depends entirely on how current the manually maintained figures are, not on the underlying model.
- Compliance-eligibility inputs (CORSIA phase, Article 6 authorization, CCP approval) reflect the separate compliance tagging methodology and its own known limitations; any gap in that tagging (for example, one major registry's Article 6 authorization flags never being set to true) will propagate into this pricing model's feature set.

Output

Two persisted price fields and one shared time-series table. A project-level field carries that project's current estimated market price. A credit-level field carries the estimated price for each individual credit, averaged across any duplicate estimates for the same credit. A running table of estimated prices by project, vintage year and transaction volume is appended to on a recurring basis, forming the shared time-series feed that powers the price explorer dashboard, portfolio valuation, and downstream index products.