Construction Rules for the Morningstar Consolidated Market Price Methodology for Venture Capital-Backed Companies

Morningstar Inc.
15 April 2026

Contents

Introduction

These Construction Rules for the Morningstar Consolidated Market Price Methodology for Venture Capital-Backed Companies has been developed to amalgamate pricing data from two distinct data vendors into a singular, integrated price to achieve a market value. This price aims to reflect the prevailing market sentiment regarding the valuation of the targeted companies.

In this document, the term "model" refers to the systematic process and algorithms employed to calculate the consolidated market price, including the Kalman filter. The Kalman filter is a well-defined and widely accepted methodology in the financial domain, aimed at refining price estimates by reducing noise and integrating multiple data sources.

The primary purpose of this document is to delineate the methodologies employed, the data integration process, and the overall functionality of the Morningstar Consolidated Market Price Methodology for Venture Capital-Backed Companies.

Background

The model addresses a significant challenge in private markets: the lack of publicly accessible, daily pricing data for companies not listed on public exchanges. These private entities often engage in groundbreaking work which may present attractive investment opportunities. However, the absence of frequent market valuations makes it difficult for investors to track or invest in these companies effectively.

Our model seeks to establish a consensus price that reflects the value of a company on any given day. This is achieved by leveraging data from two distinct sources, providing a broader perspective of the company's valuation despite the inherent limitations of private market data.

The reliability, robustness and consistency of data inputs are crucial. Each data provider offers unique insights, which can introduce idiosyncratic noise into the analysis due to the subjective methodologies used to estimate company valuations. To mitigate these variations, we employ advanced noise-reduction techniques, specifically the Kalman filter.

Data Sources and Relevance

Zanbato

Zanbato is one of our key data vendors, providing crucial inputs for our Morningstar Consolidated Market Price Methodology through their ZX Index Value and ZXData Robustness Score. The ZX Index Value is an algorithmically derived metric that reflects secondary prices for private companies, combining executed transaction prices and current bid-ask pricing. This approach offers a timely representation of market conditions, considering factors like executed prices, inside market prices, and transaction sizes.

In addition to the ZX Index Value, Zanbato's ZXData Robustness Score assesses the confidence levels in these valuations. This score, ranging from 1 to 10, evaluates metrics such as the number and age of trades, transaction volumes, bid-ask spreads, and order book symmetry.

Caplight

Caplight, a platform aggregating secondary market transactional data from various institutional investors, is a key data vendor for our Morningstar Consolidated Market Price Methodology. With an average transaction size of $2.5 million, Caplight collects and sanitizes data, including closed secondary transactions, bid and ask indications of interest, and other trading contexts. This extensive network ensures a robust and comprehensive dataset, crucial for market valuation.

Caplight's methodology involves real-time data collection and verification processes. They utilize various data types, such as venture capital funding rounds, secondary market closed trades, secondary market bid and ask orders, price events, and public comparable. Each data point undergoes rigorous filtering and validation to ensure high quality. For instance, closed transaction prices are adjusted for SPV fees. Additionally, Caplight uses smoothing techniques like thin plate regression splines to de-noise the data and maintain realistic price trends.

Rationale for Combining Prices

Integrating data from both Zanbato and Caplight provides a comprehensive and nuanced view of private company valuations. Zanbato contributes high-value, institutional trade data through their ZX Index Value and ZXData Robustness Score, reflecting significant investor confidence and robust transaction histories. This data is instrumental in capturing high-end market sentiment and price stability. On the other hand, Caplight offers a broader spectrum of market activity, including real-time bid and ask orders, secondary market closed transactions, and public comparable. Caplight's data adds granularity and immediacy, capturing daily market fluctuations and a wider range of investor behavior.

Assumptions and Prerequisites of the Kalman Filter

- Data Quality*:* It is assumed that the data from both sources is reasonably reliable and free from gross errors. Any significant anomalies should be identified and corrected before feeding into the model.
- Stationarity*:* The model assumes that the price series follows a random walk process, meaning that the changes in the price series are stochastic and do not follow a deterministic trend.

- Rolling Window and Span Parameters: Based on empirical evidence, we have decided to use a 6-month rolling window and a 5-day span for the dynamic adjustment of the Kalman filter. This combination has been shown to produce the most consistent and robust results in our price calculations.
- Data Frequency: The price data from both sources is in daily frequency, as this is a prerequisite for the methodology to work effectively.

Consolidated Market Price Methodology

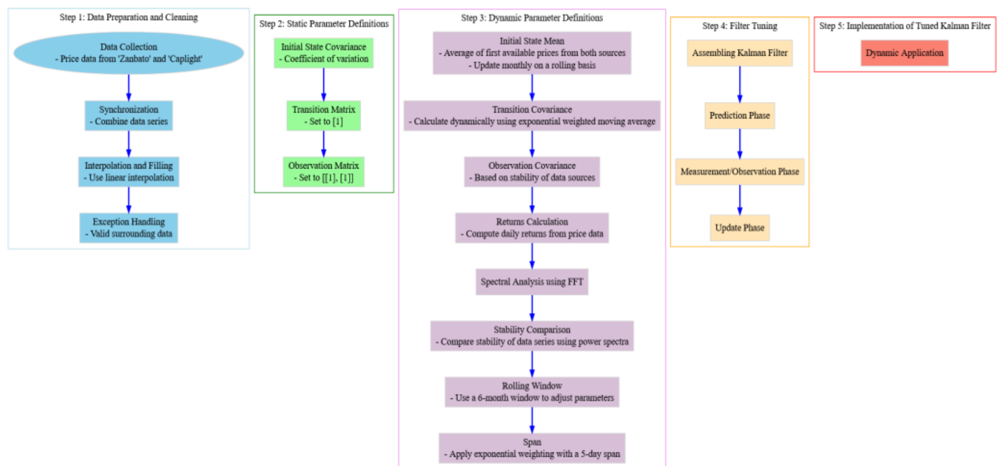
Key parameters of Kalman Filter

This section will look at 6 key parameters in the Kalman filter essential to understanding our Morningstar Consolidated Market Price Methodology.

1. *Initial State Mean:* The initial state mean is the starting point for estimating the true value of the variable being analyzed. This value is derived from the initial data points and sets the baseline for subsequent estimations. It is crucial because the accuracy of the Kalman filter's predictions relies heavily on a good initial estimate. A poor initial state mean can lead to significant deviations in the early stages of the filtering process.
2. *Initial State Covariance:* The initial state covariance represents the degree of certainty around the initial state mean. It quantifies the initial uncertainty in the system. A higher covariance indicates greater uncertainty about the initial state, while a lower covariance suggests more confidence. This parameter helps determine how much weight should be given to the initial state mean versus the new measurements.
3. *Transition Matrices:* Transition matrices describe how the state of the system evolves from one-time step to the next. These matrices encode the dynamics of the system, specifying how the current state influences the future state. In the case of a random walk model, a value of [1] implies that the next state is directly derived from the current state with some stochastic noise. This reflects the assumption that changes in the system are random but follow a predictable pattern.
4. *Observation Matrices:* Observation matrices define the relationship between the observed measurements and the true state of the system. These matrices determine how much influence each data source has during the update step. By assigning weights to different observations, the filter can balance the contributions of various data sources. Equal weighting ([1, 1]) means that both sources are considered equally reliable, which is crucial for integrating multiple data streams.
5. *Transition Covariance:* Transition covariance measures the expected variability or randomness in the system's progression over time. This parameter captures the process noise, which represents the uncertainty inherent in the system's evolution. A higher transition covariance indicates that the system is more unpredictable, while a lower value suggests a more stable progression. This helps in adjusting the filter's sensitivity to changes in the state.
6. *Observation Covariance:* Observation covariance represents the expected noise in the data measurements. It quantifies the measurement noise, with smaller values indicating higher trust in the data. This parameter is critical because it determines how much confidence the filter should place in the observations versus the predicted state. Lower observation covariance means the data is more reliable.

Model Implementation Steps

Exhibit 1: Model Implementation Steps



Source: Morningstar.

1. Data Preparation and Cleaning
    a. Data Collection: Price data is collected from two distinct sources, 'Zanbato' and 'Caplight'.
    b. Synchronization: The data series from both sources are combined into a single data frame to synchronize their indices.
    c. Interpolation and Filling: To ensure continuous data series, we use linear interpolation to handle missing values. Any remaining NaNs are filled with corresponding values from other series or with the mean of non-NaN values. This data imputation technique helps maintain data continuity, preventing any interruptions that could hinder analysis.
    d. Exception Handling: We ensure interpolation is applied only when non-NaN values exist both before and after a gap. This approach prevents any form of extrapolation, such as forward filling, thereby avoiding the introduction of data that lacks empirical support.

2. Defining Static Parameters
    a. Transition Matrix: Set to [1], assuming a random walk model where the future state is a stochastic step from the current state.
    b. Observation Matrix: Set to [[1], [1]], giving equal weight to both data sources during the update step.
3. Defining Dynamic Parameters
    a. Initial State Mean: The initial state mean is determined by averaging the first available prices from both data sources. This provides a baseline estimate to start the filter. As more data becomes available over time, the initial estimate is continuously updated

based on the latest month's data, calculated on a rolling basis. This approach ensures that the model's estimates adapt to reflect the changing dynamics of the market prices. This dynamically updated mean serves as the foundational starting point for the Kalman filter, allowing it to begin with the most current reflection of the market conditions.

b. Initial State Covariance: Calculated as the coefficient of variation of the price differences between the two sources over the last 10 days from available data at each time step. This quantifies the initial uncertainty in our estimates. [1]
c. Transition Covariance: Dynamically calculated using an exponential weighted moving average of the daily price changes. Reflects the expected variability in the system's progression over time.
d. Observation Covariance: Determined based on the stability of each data source, with lower values indicating higher trust in the respective source. Steps to set observation covariance include:
    i. Returns Calculation
        1. Computes the daily returns from the price data.
        2. Calculates the percentage change between consecutive data points and removes any NaN values resulting from this calculation.
    ii. Spectral Analysis
        1. Performs spectral analysis on a given data series to calculate its power spectrum.
        2. Uses Fast Fourier Transform (FFT) to convert the time-series data into the frequency domain.
    iii. Stability Comparison
        1. Compares the stability of the two data series based on their power spectra.
        2. Integrates the power spectrum over a range of frequencies to quantify stability (stability is determined by summing the power spectrum values within a specified frequency range. The series with the lower sum is considered more stable).
    iv. Observation Covariance Calculation
        1. Calculates the observation covariance based on the stability comparison.
        2. Assigns lower covariance to the more stable data source, indicating higher trust.
        3. If both sources are equally stable, equal covariance values are assigned.
e. Rolling Window: A rolling window of 6 months is used to dynamically adjust parameters based on recent data.
f. Span: An exponential weighting scheme with a span of 5 days is applied to manage the decay of weights, emphasizing recent data more heavily.

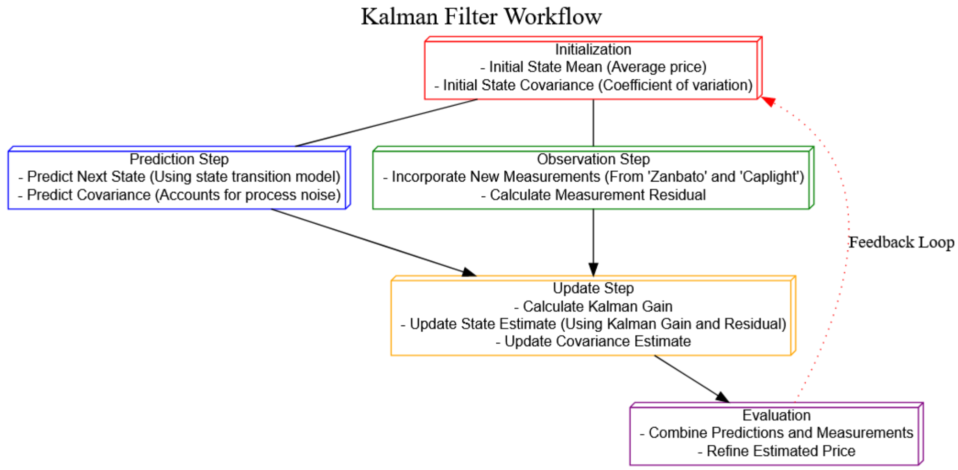
4. Filter Tuning
    a. Assembling the Kalman Filter: The Kalman Filter is assembled using the defined parameters, incorporating user inputs for the rolling window (length of historical data) and span (for exponential weighting). The filter is dynamically configured at each time step with the rolling window approach, ensuring parameters reflect recent data trends.
    b. Prediction and Update Phases
        i. Prediction Phase: The filter forecasts the next state using the state transition model.
        ii. Measurement/Observation Phase: The filter compares new measurements from 'Zanbato' and 'Caplight' against the predicted state.
        iii. Update Phase: Involves calculating the Kalman Gain, which balances the uncertainties in the predictions and the new data. The gain is then used to adjust the state estimate, updating both the state and covariance estimates to refine the overall prediction.
5. Implementation of the Tuned Kalman Filter
    a. Dynamic Application: The filter is applied dynamically using the rolling window and span parameters, starting from the fifth data entry to ensure enough data points for filtering. This iterative process involves continuously refining the price estimate by integrating both historical data and new observations.[1]

The Theory behind Kalman Filter

The Kalman filter is a recursive algorithm designed to estimate the state of a dynamic system from a series of incomplete and noisy measurements. It is widely used in various fields, including signal processing, econometrics, and robotics, due to its robustness in handling uncertain data. In our Morningstar Consolidated Market Price Methodology, we utilize the Kalman filter to integrate price data from two sources ('Zanbato' and 'Caplight') and generate a single, noise-reduced consensus price.

[1] For a sample plot of consolidated price, refer to Appendix A.

Exhibit 2: Kalman Filter Process Flow



Source: Morningstar.

Components of the Kalman Filter

State Vector (x):

Represents our ongoing estimate of market value, or "true price," of a privately held company. This is the price we believe the company would fetch in an open market based on the data we receive. Because private companies don't have publicly available market prices, this estimate is continuously refined as new data comes in, allowing us to closely approximate the company's actual value.

Prediction Step:

- Forecasts the next state based on the current state estimate.
- The predicted state (prior estimate) is calculated using the state transition model. Assuming a random walk model, the prediction is a simple carry-over of the current state, adjusted by the transition covariance.

$$Predicted\ State = x_{t|t-1}$$

The predicted covariance accounts for the previous uncertainty, the system's progression, and the process noise.

$$Predicted\ Covariance = P_{t|t-1} = F\ \times P_{t-1} \times F^T + Q$$

Where F is the state transition matrix, FT is its transpose, and Q is the transition covariance, reflecting the expected variability in the system's progression over time.

Measurement (Observation) Step:

- Incorporates new measurements from the data sources ('Zanbato' and 'Caplight').
- The actual measurement is compared against the predicted state.

$$Measurement\ Residual = z_t - H \times x_{t|t-1}$$

Where zt is the actual measurement at time t, and H is the observation matrix.

Update Step:

- Adjusts the predicted state based on the measurement residual and the Kalman Gain.
- The Kalman Gain (K) determines how much weight to give to the predicted state versus the observed state.

$$Kt = P_{t|t-1} \times H^T \times \left(H \times P_{t|t-1} \times H^T + R\right)^{-1}$$

Where R is the observation covariance.

The updated state estimate is calculated by adjusting the predicted state using the Kalman Gain and the measurement residual.

$$Updated\ State = x_{t|t} = x_{t|t-1} + K_t \times Measurement\ Residual$$

The updated covariance estimates update the uncertainty associated with the state estimate.

$$Updated\ Covariance = P_{t|t} = (I - K_t \times H) \times P_{t|t-1}$$

Here:

- I is the identity matrix, ensuring that the dimensionality and structure of the matrices are preserved during the matrix operations.
- Kt is the Kalman Gain, which adjusts the weight given to the prediction versus the observation.
- H is the observation matrix, which maps the true state space into the observed space.
- Pt∣t−1 is the predicted covariance matrix before incorporating the new observation.

Explanation and Rationale

Kalman Gain:

- The Kalman Gain is a critical factor that balances the uncertainty in the predictions (from the model) and the uncertainty in the observations (from new data).
- A higher Kalman Gain indicates that more weight is given to the new measurements, which happens when the measurements are considered reliable (low observation covariance R).

Dynamic Updating:

- As each new set of measurements is processed, the state estimates and covariances are updated, making the model adaptive to changes observed in the input data.

- This continuous refinement ensures that the filter remains robust, even in the face of noisy or incomplete data.

Balancing Act:

- The Kalman Filter effectively balances historical data (predictions based on the model) and new information (measurements), minimizing the overall estimation error.
- This balance is achieved through the dynamic calculation of the Kalman Gain, which adjusts the weight given to the model's predictions and new measurements based on their respective uncertainties.

In our Morningstar Consolidated Market Price Methodology, the Kalman filter plays a crucial role in integrating and smoothing price data from multiple sources.

Challenges and Limitations of the Morningstar Consolidated Market Price Methodology:

- Data Quality: Assumes data accuracy from both sources, with significant anomalies corrected before use.
- Stationarity: Assumes price series follow a random walk without deterministic trends.

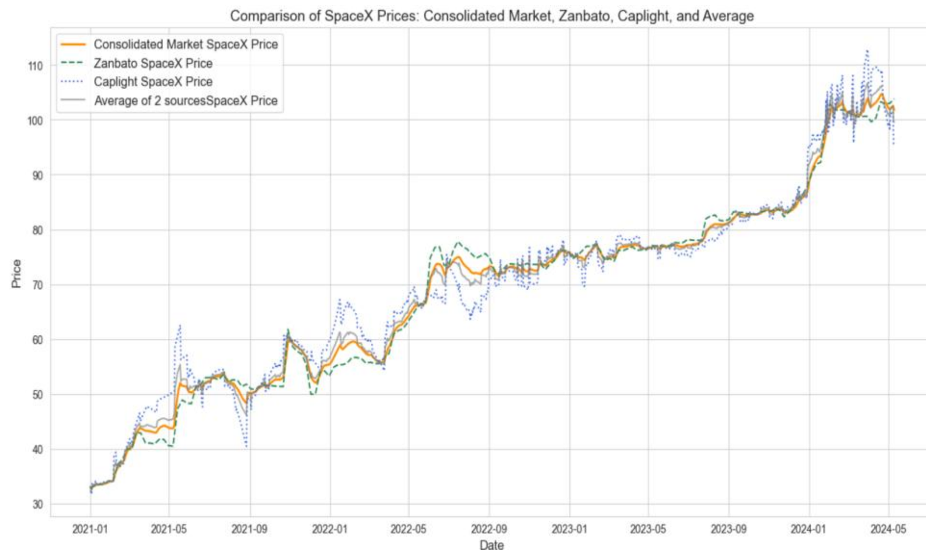
Rolling Window and Span Parameters: Utilizes a six-month rolling window and a five-day span for the Kalman filter, based on empirical evidence.

- Data Frequency: Requires daily frequency for effective methodology.
- Market Discrepancies: Factors like issuer restrictions, trading inconsistencies, and bid-ask spreads may cause differences between prices and actual market prices.
- Differences From General Partners: Prices may differ from those reported by general partners with access to nonpublic information.
- Structure-Agnostic Data: Caplight and Zanbato pricing may not match the preferred stock prices from recent financing rounds and do not account for liquidation preferences.
- Regulatory Evolution: Changes in regulations could impact private secondary marketplaces like Zanbato.

Excludes Company-Sponsored Rounds: The methodology does not include company-sponsored financing rounds.

Appendix

Appendix A: Sample Plot of Consolidated Market Price



Appendix B: Working Example

To illustrate how our Consolidated Pricing Model operates, we'll walk through an example using hypothetical data from our two sources, Vendor_1 and Vendor_2. This example introduces variability in the observed prices to demonstrate how the Kalman Filter refines the final consolidated price by favoring the more stable data source.

Example Data with Noise:

Date	Vendor_1 Price	Vendor_2 Price	Variability (Vendor_1)	Variability (Vendor_2)
Day 1	$100	$102	Low	High
Day 2	$101	$104	Low	High
Day 3	$102	$106	Low	High
Day 4	$103	$103	Low	High
Day 5	$104	$107	Low	High
Day 6	$105	$110	Low	High

Step 1: Initial State Estimate

We begin by calculating an initial estimate of the price on Day 1, using the prices from Vendor_1 and Vendor_2:

Initial Estimate (Day 1): $(100 + 102)/2 = 101$

Step 2: Predicting the Next State

For Day 2, the Kalman Filter predicts the price based on the previous day's estimate. The prediction also considers the observed variability in the prices from both sources.
Predicted Price for Day 2: $101 (carried forward from Day 1)

Step 3: Observation/Measurement Update

On Day 2, the filter compares the predicted price ($101) with the actual prices from Vendor_1 ($101) and Vendor_2 ($104). The Kalman Filter evaluates the variability in the data from both sources:
Vendor_1: Consistent, low variability.
Vendor_2: Higher variability, indicating more fluctuations.
Given the higher variability in Vendor_2 prices, the Kalman Filter places more weight on the more stable Vendor_1 price:
Updated Price Estimate (Day 2): $101.5 (leaning towards Vendor_1)

Step 4: State Update

The filter updates the price estimate for Day 2 to $101.5, which will now serve as the starting point for Day 3.

Step 5: Repeating the Process

This process is repeated for each subsequent day. The Kalman Filter continuously adjusts the weights based on the observed variability in the prices:

Day 3:
Predicted Price: $101.5
Observed Prices: Vendor_1 $102, Vendor_2 $106
Variability: Vendor_2 remains highly variable, so Vendor_1 price is favored.
Updated Price Estimate: $103 (heavier weighting towards Vendor_1)

Day 4:
Predicted Price: $103
Observed Prices: Vendor_1 $103, Vendor_2 $103
Variability: Vendor_2 is stable today.
Updated Price Estimate: $103 (even weighting)

Day 5:
Predicted Price: $103
Observed Prices: Vendor_1 $104, Vendor_2 $107
Variability: Vendor_2 variability increases again, so Vendor_1 is favored.
Updated Price Estimate: $104.25

Day 6:
Predicted Price: $104.25

Observed Prices: Vendor_1 $105, Vendor_2 $110
Variability: High variability in Vendor_2, consistent Vendor_1.
Updated Price Estimate: $105.5 (weighted towards Vendor_1)

Final Consolidated Prices:

Date	Vendor_1 Price	Vendor_2 Price	Consolidated Price
Day 1	$100	$102	$101
Day 2	$101	$104	$101.50
Day 3	$102	$106	$103
Day 4	$103	$103	$103
Day 5	$104	$107	$104.25
Day 6	$105	$110	$105.50

About Morningstar Indexes

Morningstar Indexes was built to keep up with the evolving needs of investors—and to be a leading-edge advocate for them. Our rich heritage as a transparent, investor-focused leader in data and research uniquely equips us to support individuals, institutions, wealth managers and advisors in navigating investment opportunities across major asset classes, styles and strategies. From traditional benchmarks and unique IP-driven indexes, to index design, calculation and distribution services, our solutions span an investment landscape as diverse as investors themselves.

Please visit indexes.morningstar.com for more information.

Contact:
indexes@morningstar.com

MORNINGSTAR Indexes

22 West Washington Street
Chicago, IL 60602 USA