When Linear Regression Isn't Enough: Building Optimization Models for Change-Point Detection and Trend Analysis

Abstract

Organizations routinely collect time-series data to monitor systems and support operational decision making. Whether analyzing financial markets, equipment degradation, or other evolving processes, an important challenge is identifying when underlying trends change and how to analyze those. This teaching case introduces piecewise-linear fitting as a framework for addressing this problem, guiding students from linear regression to dynamic programming and mixed-integer optimization while comparing alternative methodologies based on computational efficiency, scalability, and usefulness.

1 Introduction

Many scientific and operational problems require understanding how an observed system evolves over time. Figure 1 presents a representative time series from such a setting. At first glance, the overall trend appears to be increasing, suggesting that a single linear regression model may adequately describe the data. However, a closer inspection reveals several distinct phases, each exhibiting a different linear trend.

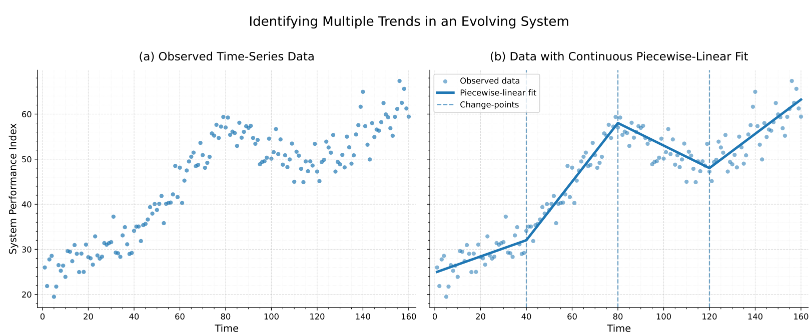


Figure 1: Illustrative time series exhibiting four continuous piecewise-linear segments. The observations follow distinct local trends separated by change-points, motivating the need for methods that jointly identify the breakpoint locations and estimate the corresponding linear segments.

A common first approach is to visually identify the locations where the trend appears to change (i.e., breakpoints), partition the data accordingly, and fit a separate linear regression model to each segment. While this may produce a reasonable visual approximation, it raises several important questions. Where should the breakpoints be placed? Are these breakpoints truly optimal, or are they simply a consequence of subjective judgment? More importantly, if different analysts choose different breakpoints, which solution should be trusted?

These questions arise in many practical settings including healthcare (Tobiasz et al. [2023]), power systems (Buason et al. [2025], Ahmadi et al. [2013]), manufacturing (Gao et al. [2018]), and public policy (Wagner et al. [2002], Dehning et al. [2020]). Analysts seek to identify changes in financial markets, engineers monitor the degradation of industrial equipment, power utility companies study the evolution of power outages during extreme weather, and healthcare researchers investigate changes in patient outcomes. In each of these applications, the objective is not merely to fit the data, but to identify meaningful changes in the underlying trend in a principled and reproducible manner.

Piecewise-linear (PWL) fitting provides a natural framework for addressing this problem. Rather than fitting a single regression line to the entire dataset, the observations are partitioned into contiguous segments, with a separate linear function fitted to each segment. The locations at which the trend changes are known as change-points, and identifying these change-points becomes a central modeling challenge. In the optimization literature, the locations separating adjacent linear segments are commonly referred to as breakpoints, whereas the statistical literature typically uses the term change-points. Since both terms refer to the same concept in the context of this case, they are used interchangeably throughout.

When the breakpoint locations are known in advance, the problem decomposes into a collection of independent regression models and is computationally straightforward. In practice, however, the breakpoint locations are rarely known beforehand. Instead, they must be determined simultaneously with the parameters of the fitted line segments. The number of possible segmentations grows rapidly with the size of the dataset, transforming what initially appears to be a regression problem into a challenging combinatorial optimization problem.

This case develops solution approaches through a progressive sequence of models and algorithms. Students first examine classical linear regression, before extending it to piecewise-linear fitting with fixed breakpoints. Unknown breakpoints naturally motivate dynamic programming for discontinuous piecewise-linear fitting. The case then shows how requiring adjacent line segments to join continuously fundamentally changes the mathematical structure of the problem, motivating mixed-integer optimization formulations. Students compare several mathematically equivalent formulations and observe how formulation design influences computational performance, scalability, and the ability to obtain globally optimal solutions.

Examples from financial time-series analysis and hurricane-induced power outages are used to illustrate the broad applicability of the framework. Rather than emphasizing a particular application, the case focuses on the development of quantitative modeling, algorithmic thinking, and optimization-based decision making for analyzing changing trends in ordered data.

2 Methodology

This case develops piecewise-linear (PWL) fitting through a sequence of increasingly sophisticated optimization models. Beginning with ordinary linear regression, students examine how different loss functions, breakpoint assumptions, and continuity requirements change the mathematical and computational structure of the problem. The sequence progresses from convex optimization and dynamic programming to mixed-integer optimization, emphasizing how modeling choices determine the appropriate solution approach.

Linear fitting. Considering ordered observations $\{(x_t, y_t)\}_{t=1}^{T}$, fitting a single linear function by minimizing squared error leads to the unconstrained convex quadratic problem

$$\min_{m,c} \sum_{t=1}^{T} (y_t - mx_t - c)^2.$$

If absolute error is used instead, the problem becomes

$$\min_{m,c} \sum_{t=1}^{T} |y_t - mx_t - c|,$$

which can be reformulated as a linear program using auxiliary variables. These two models allow students to compare quadratic and linear optimization formulations for the aforementioned regression objectives.

Piecewise-linear fitting with fixed breakpoints. When the breakpoint locations are known and continuity is not required, each segment can be fitted independently using ordinary least squares or least absolute deviations. The resulting problem therefore decomposes into separate regression models. If continuity is imposed, neighboring segments become linked. Nevertheless, the model can be reparameterized using hinge functions and solved as a convex optimization problem.

Piecewise-linear fitting with unknown breakpoints. When breakpoint locations are unknown, the segmentation itself becomes a decision. Without continuity, the fitting error is additive across segments, allowing the problem to be solved using dynamic programming, following the classical approach of Bai and Perron [1998, 2003]. For example, if $C(s, t)$ denotes the optimal fitting cost for observations $s, \dots, t$, then a dynamic programming recursion can be written as

$$D(k, t) = \min_{s} \{D(k-1, s-1) + C(s, t)\},$$

where $D(k, t)$ is the minimum cost of fitting the first t observations with k segments.

When continuity is required, adjacent segments are coupled through their unknown intersection points. The segment costs are no longer independent, and the additive structure required by dynamic programming is lost. This leads to an NP-Hard problem, and motivates a mixed-integer optimization formulation.

2.1 Mixed-Integer Formulations for Continuous PWL Fitting

The continuous PWL fitting problem can be expressed using several mathematically equivalent formulations. Although these formulations describe the same underlying fitting problem, they differ in their variables, constraint structure, relaxation strength, and computational performance.

Basic formulation The Basic formulation, proposed by Goldberg et al. [2021], models the segmentation and regression problem simultaneously using mixed-integer optimization.

The primary combinatorial decision is determining which observations belong to each linear segment. For every segment $j \in \{1, \ldots, K\}$ and observation $t \in \{1, \ldots, T\}$, a binary variable $\delta_{j,t} \in \{0, 1\}$ is introduced, where

$$\delta_{j,t} = \begin{cases} 1, & \text{if observation } t \text{ is assigned to segment } j, \\ 0, & \text{otherwise.} \end{cases}$$

The formulation also determines the slope and intercept, $m_j,\ c_j$ respectively, of every segment, together with breakpoint variables r_j that specify the locations where one linear segment ends and the next begins. Finally, continuous variables $\hat{y}_t$ represent the fitted value associated with each observation. Each observation must belong to exactly one segment,

$$\sum_{j=1}^{K} \delta_{j,t} = 1, \qquad t = 1, \ldots, T,$$

and big-M constraints ensure that the fitted value equals the corresponding segment equation whenever an observation is assigned to that segment. Explicit breakpoint variables enforce contiguous segment assignments.

Continuity between adjacent segments is imposed through

$$m_j r_j + c_j = m_{j+1} r_j + c_{j+1}, \qquad j = 1, \ldots, K-1,$$

which requires the two neighboring line segments to intersect at each breakpoint. Since both the slopes and breakpoint locations are decision variables, these equations are bilinear, yielding a mixed-integer nonlinear optimization problem.

Alternative formulation The Alternative formulation of Rebennack and Krasko [2020] uses the same decision variables for segment-assignments, slopes, intercepts, and fitted values as the Basic formulation. However, it removes the explicit breakpoint variables from the model. Instead, the formulation introduces binary variables

$$\gamma_j \in \{0, 1\},$$

to indicate whether the slope decreases or increases between consecutive segments, together with auxiliary activation variables

$$\delta^+_{j,t}, \qquad \delta^-_{j,t},$$

which determine which continuity constraints become active. Suppose the transition from segment j to segment $j+1$ occurs between observations x_t and x_{t+1}. The breakpoint is the intersection of the two fitted lines,

$$r = \frac{c_{j+1} - c_j}{m_j - m_{j+1}}.$$

Requiring this intersection to lie between the two observations $[x_t, x_{t+1}]$ yields linear inequalities once the sign of $m_j - m_{j+1}$ is known. For example, when $m_j \geq m_{j+1}$, continuity is enforced through

$$c_{j+1} - c_j \geq x_t(m_j - m_{j+1}) - M(1 - \delta^+_{j,t}),$$
$$c_{j+1} - c_j \leq x_{t+1}(m_j - m_{j+1}) + M(1 - \delta^+_{j,t}),$$

while a second pair of inequalities is used when the slope increases. The formulation also replaces explicit breakpoint variables with recursive assignment constraints such as

$$\delta_{j+1,t+1} \leq \delta_{j,t} + \delta_{j+1,t},$$

which guarantee that each segment occupies one contiguous block of observations. Consequently, continuity is modeled using only linear mixed-integer constraints.

Extended formulation The Extended formulation proposed by Narula et al. [2026] retains the regression variables

$$m_j,\ c_j,\ \hat{y}_t,$$

but replaces the segment-assignment part of the model. Instead of assigning observations directly to segments, it introduces binary variables

$$X_{j,t} \in \{0,1\}, \qquad j = 1,\ldots,K-1,$$

where

$$X_{j,t} = 1$$

indicates that observation t belongs to one of the first j segments, while

$$X_{j,t} = 0$$

indicates that it belongs to segment $j+1$ or any later segment.

The original assignment variables are recovered through

$$\delta_{j,t} = \begin{cases} X_{1,t}, & j = 1, \\ X_{j,t} - X_{j-1,t}, & j = 2,\ldots,K-1, \\ 1 - X_{K-1,t}, & j = K. \end{cases}$$

To ensure valid segmentations, the variables satisfy monotonicity constraints over both time and segment indices,

$$X_{j,t} \geq X_{j,t+1}, \qquad X_{j+1,t} \geq X_{j,t},$$

which guarantee that every segment forms one contiguous interval and that segments appear in increasing order. The Extended formulation can replace the assignment block of either the Basic or Alternative formulation, producing the Extended Basic and Extended Alternative models. Although mathematically equivalent to the original formulations, the stronger assignment structure yields a tighter linear programming relaxation and is shown to substantially improve computational performance.

These three formulations allow students to study how the same optimization problem can be represented in different ways. The Basic formulation is geometrically intuitive but contains bilinear constraints. The Alternative formulation linearizes continuity, while the Extended formulation strengthens the segment-assignment structure. Their comparison demonstrates that mathematically equivalent formulations can have substantially different theoretical properties and computational performance.

Overall, the case presents linear regression, convex optimization, dynamic programming, mixed-integer nonlinear optimization, linearization, and formulation strengthening as a connected progression. Students gain exposure not only to formulating different versions of piecewise-linear fitting, but also to understanding how optimization formulations evolve as additional modeling requirements are introduced, and how careful formulation design can improve computational performance.

3 Experiments

The examples covered in this section illustrate how optimization-based change-point detection can identify common structural changes across multiple correlated time series, producing an interpretable representation of the underlying process evolution. All datasets used throughout this case study are publicly available and obtained from open-source sources. To facilitate reproducibility and classroom use, the compiled datasets used in the experiments are available in the following GitHub repository:

`https://github.com/ApoorvaGOR/change-point-detection-case-data`

3.1 Example: Power Outage Trend Analysis During Extreme Weather

Power outages caused by hurricanes evolve over time through several distinct phases, including outage escalation as the storm makes landfall, periods of sustained disruption, and gradual restoration as repair efforts progress. Neighboring counties are often affected by the same storm system and therefore exhibit similar trend changes, although the magnitude of the outages may differ substantially. Consequently, it is natural to model these outage trajectories jointly rather than fitting each county independently. In this example, we study Hurricane Nicholas (2021) using two publicly available datasets. Hurricane trajectory and wind-field information are obtained from NOAA's HURDAT2 archive, while county-level power outage counts are obtained from the U.S. Department of Energy's EAGLE-I platform. Each county is represented by a time series of the number of customers without power at 15-minute intervals. The objective is to capture the outage trend across the severely impacted counties, and identify a set of shared change-points to fit a continuous piecewise-linear function to each outage trajectory. Formally, if $\{(x_t, y_t^d)\}_{t=1}^T$ for $d = 1, \ldots, D$ denotes the outage time series for D counties, we seek a collection of continuous piecewise-linear functions that share the same breakpoint locations while allowing each county to have its own slopes and intercepts. This corresponds directly to a multi-dimensional continuous PWL fitting problem, and is solved using the discussed MIP formulations.

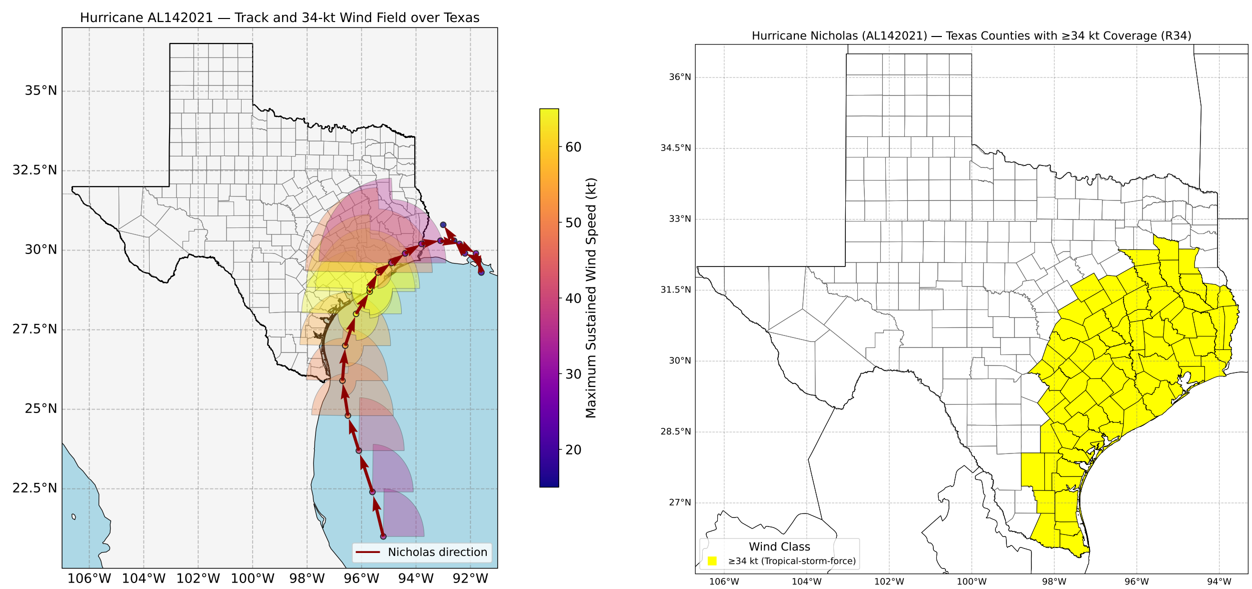


Figure 2: Hurricane Nicholas hazard context. **(a)** Six-hourly storm-center locations over Texas obtained from the HURDAT2 database. The dashed curve traces the hurricane trajectory, while the storm-center markers and their associated 34-kt wind fields are colored according to the maximum sustained wind speed (knots). Arrows indicate the direction of storm movement. **(b)** Texas counties intersecting the hurricane's 34-kt (tropical-storm-force) wind field during the event. The highlighted counties represent regions exposed to sustained winds of at least 34 knots and provide the geographic context for the subsequent power outage analysis.

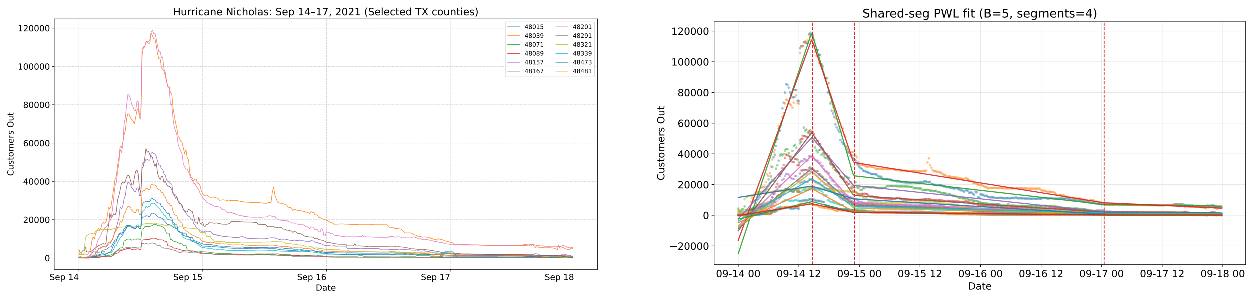


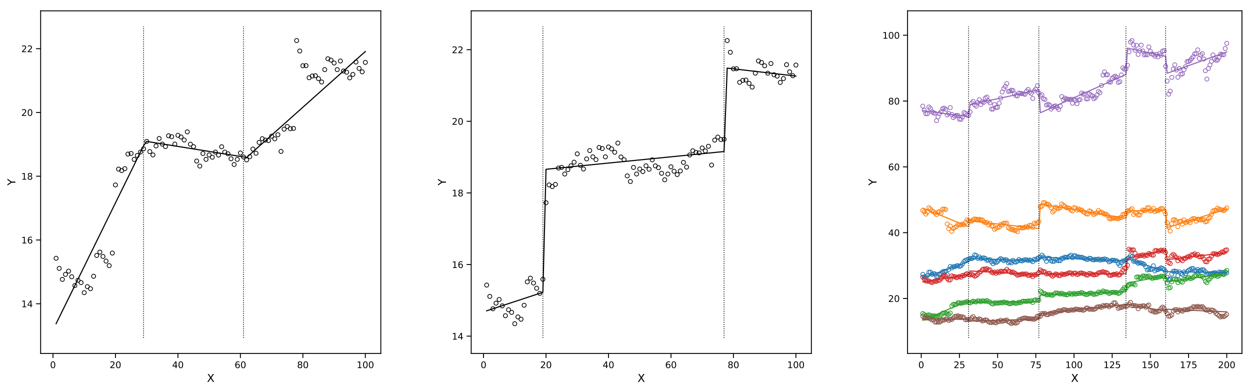
Figure 3: Multi-dimensional continuous piecewise-linear fitting of county-level power outages. **(a)** Observed outage trajectories for the 12 most severely impacted Texas counties (based on total number of customer hours lost due to power outage) obtained from the EAGLE-I platform at 15-minute resolution. Each curve represents the number of customers without power throughout the duration of Hurricane Nicholas. **(b)** Continuous multi-dimensional piecewise-linear approximation with shared breakpoints across all county time series. While each county has its own piecewise-linear trend, the breakpoints are constrained to occur at common time instants, producing an interpretable representation of the shared outage escalation and recovery phases during the hurricane.

Figure 2 illustrates the hurricane trajectory and corresponding wind footprint over Texas. Figure 3 compares the observed outage trajectories with the fitted multi-dimensional continuous PWL approximation obtained using the mixed-integer programming formulation.

3.2 Example: Detecting Common Trend Changes in Financial Time Series

Financial markets often consist of many correlated assets that respond simultaneously to major economic events, company announcements, or broader changes in market sentiment. Although each stock follows its own price trajectory, groups of assets frequently undergo changes in trend at similar points in time. Identifying these common structural changes can help analysts summarize market behavior, detect regime shifts, and study periods of synchronized movement across multiple securities. In this example, we consider the daily closing prices of six large technology companies, namely Apple Inc. (AAPL), Microsoft Corporation (MSFT), Amazon.com Inc. (AMZN), Alphabet Inc. (GOOGL), Meta Platforms, Inc. (META), and Tesla, Inc. (TSLA) starting from January 1, 2015. The objective is not to predict future prices, but rather to obtain an interpretable piecewise-linear approximation of the observed price trajectories and identify major changes in their underlying trends.

Figure 4 illustrates several variants of the PWL fitting problem. Figures 4a and 4b both fit the same Amazon (AMZN) closing-price series, differing only in whether continuity is enforced between adjacent linear segments. While the continuous formulation requires neighboring segments to meet smoothly at each breakpoint, the discontinuous formulation permits abrupt changes in the fitted values. Despite producing noticeably different fitted functions, both models provide visually reasonable approximations of the underlying data. This highlights that seemingly minor modeling assumptions, such as enforcing continuity between adjacent segments, can lead to fundamentally different solutions. It also serves to highlight that, in general, an optimal continuous piecewise-linear fit cannot be recovered or even closely approximated by simply adjusting or smoothing a discontinuous fit. Finally, Figure 4c extends the framework to a multi-dimensional setting by jointly fitting all six stock price series using shared breakpoint locations. Although each company has its own piecewise-linear trend, the breakpoints are constrained to occur at common time instants, allowing the model to identify broader market-wide structural changes while preserving company-specific behavior.



(a) $T = 100$, $K = 3$ **(b)** $T = 100$, $K = 3$ **(c)** $T = 200$, $K = 5$, $D = 6$

Figure 4: Illustration of piecewise-linear fits obtained through solving the MIP formulations while minimizing the ℓ_1 error norm (Total Absolute Error): (a) univariate fitting with continuity based on Amazon (AMZN) closing prices, (b) univariate fitting without continuity using the same data as (a), and (c) multi-dimensional fitting with shared change-points across 6 time series.

4 Conclusion

This case demonstrates how a seemingly modest extension of linear regression can lead to fundamentally different mathematical and computational problems. When the breakpoint locations are fixed, piecewise-linear fitting reduces to a collection of standard regression problems. When the breakpoints are unknown, however, the model must determine both the segmentation of the data and the parameters of the fitted linear pieces. In the discontinuous setting, the additive structure of the segment costs permits an efficient dynamic-programming approach. Requiring continuity couples neighboring segments, removes this separability, and motivates the use of mixed-integer optimization. The progression from the Basic to the Alternative and Extended formulations further illustrates that constructing a correct optimization model is only part of the modeling process. Mathematically equivalent formulations can differ substantially in their constraint structure, relaxation strength, and computational performance. The financial and hurricane-outage examples demonstrate that the same methodological framework can be applied across different domains. More broadly, the case highlights the close relationship between statistical modeling, algorithm design, and mathematical optimization. The choice among linear regression, dynamic programming, nonlinear estimation, and mixed-integer programming should depend on the structural assumptions of the problem, the required level of interpretability, the available computational resources, and the value of a global-optimality guarantee. The central lesson is that when linear regression is no longer sufficient, careful exploitation of problem structure and thoughtful formulation design are essential for building models that are both meaningful and computationally effective.

References

Hamed Ahmadi, José R Martí, and Ali Moshref. Piecewise linear approximation of generators cost functions using max-affine functions. In *2013 IEEE Power & Energy Society General Meeting*, pages 1–5. IEEE, 2013.

Jushan Bai and Pierre Perron. Estimating and testing linear models with multiple structural changes. *Econometrica*, pages 47–78, 1998.

Jushan Bai and Pierre Perron. Computation and analysis of multiple structural change models. *Journal of applied econometrics*, 18(1):1–22, 2003.

Paprapee Buason, Sidhant Misra, and Daniel K Molzahn. Sample-based piecewise linear power flow approximations using second-order sensitivities. *arXiv preprint arXiv:2501.13825*, 2025.

Jonas Dehning, Johannes Zierenberg, F Paul Spitzner, Michael Wibral, Joao Pinheiro Neto, Michael Wilczek, and Viola Priesemann. Inferring change points in the spread of covid-19 reveals the effectiveness of interventions. *Science*, 369(6500):eabb9789, 2020.

Xiaoyong Gao, Zhenhui Feng, Yuhong Wang, Xiaolin Huang, Dexian Huang, Tao Chen, and Xue Lian. Piecewise linear approximation based milp method for pvc plant planning optimization. *Industrial & Engineering Chemistry Research*, 57(4):1233–1244, 2018.

Noam Goldberg, Steffen Rebennack, Youngdae Kim, Vitaliy Krasko, and Sven Leyffer. Minlp formulations for continuous piecewise linear function fitting. *Computational Optimization and Applications*, 79(1):223–233, 2021.

Apoorva Narula, Santanu S Dey, and Yao Xie. Mixed-integer programming for change-point detection. *arXiv preprint arXiv:2602.11947*, 2026.

Steffen Rebennack and Vitaliy Krasko. Piecewise linear function fitting via mixed-integer linear programming. *INFORMS Journal on Computing*, 32(2):507–530, 2020.

Joanna Tobiasz, Najla Al-Harbi, Sara Bin Judia, Salma Majid Wakil, Joanna Polanska, and Ghazi Alsbeih. Multivariate piecewise linear regression model to predict radiosensitivity using the association with the genome-wide copy number variation. *Frontiers in Oncology*, 13:1154222, 2023.

Anita K Wagner, Stephen B Soumerai, Fang Zhang, and Dennis Ross-Degnan. Segmented regression analysis of interrupted time series studies in medication use research. *Journal of clinical pharmacy and therapeutics*, 27(4):299–309, 2002.