INFERRING DAILY GAS CONSUMPTION FROM MULTIPLE NONUNIFORMLY SAMPLED BILLING CYCLES WITH HIERARCHICAL CONSTRAINTS

by

Colin O. Quinn, B.S., M.Sc.

A Dissertation submitted to the Faculty of the Graduate School,

Marquette University,

in Partial Fulfillment of the Requirements for

the Degree of Doctor of Philosophy

Milwaukee, Wisconsin

May 2025

ABSTRACT

INFERRING DAILY GAS CONSUMPTION FROM MULTIPLE NONUNIFORMLY SAMPLED BILLING CYCLES WITH HIERARCHICAL CONSTRAINTS

Colin O. Quinn, B.S., M.Sc.

Marquette University, 2025

Local natural gas demand forecasting is essential for local distribution companies (LDCs) to optimize resources, maintain efficiency, and meet regulatory requirements. However, inconsistencies arising from nonuniform billing cycles, hierarchical data structures, and aggregated forecasts often degrade accuracy. This dissertation introduces three methodological advancements to address these issues: the Multi-Source Iterative Shifting Disaggregation (ISD) algorithm, Single-Dimension Hierarchical Reconciliation, and Cross-Temporal Hierarchical Forecast Reconciliation.

The ISD algorithm disaggregates multiple, overlapping, nonuniformly sampled time series into a coherent high-frequency signal. Unlike existing methods, ISD iteratively refines estimates through constrained load shifting, improving forecast accuracy. Applied to billing cycle data, ISD reduces the Weighted Mean Absolute Percentage Error (WMAPE) by 1.4–4.3%, while residential consumption disaggregation sees a 4.6–10.4% improvement.

To enforce coherence in hierarchical forecasting, Single-Dimension Hierarchical Reconciliation corrects inconsistencies within either spatial or temporal hierarchies, enhancing forecast reliability and reducing bias. Cross-Temporal Hierarchical Forecast Reconciliation extends this approach by integrating both spatial and temporal constraints, ensuring consistency across all levels. This method improves hourly forecast accuracy by 10% and daily accuracy by 3%, outperforming Variance Scaling by 7% and 9%, respectively.

This dissertation demonstrates that maintaining hierarchical coherence enhances gas demand forecasting. The proposed methods improve accuracy, reduce error, and provide a scalable framework applicable to broader energy forecasting problems. These advancements offer LDCs practical tools to optimize operations and inform decision-making.

ACKNOWLEDGEMENTS

Colin O. Quinn, B.S., M.Sc.

I would like to extend my gratitude to my dissertation committee members Drs. Richard Povinelli, George Corliss, Ronald Brown, Praveen Madiraju, and Debbie Perouli for their support, guidance, and mentorship throughout this journey. Your insights, encouragement, and critical feedback have profoundly shaped my research and helped me refine my thinking at every stage of this dissertation. I am sincerely grateful for the time and effort you have dedicated to my development as a research scientist.

I am especially thankful for Dr. Richard Povinelli and Dr. George Corliss, who played a pivotal role in shaping my research skills and intellectual curiosity. Through their mentorship, I have learned not just how to conduct research, but how to approach problems with clarity, patience, and a deep appreciation for the process of discovery. Their guidance extended far beyond technical expertise; they have imparted lessons in perseverance, integrity, and the importance of always asking the right questions. These lessons will stay with me throughout my professional career and personal life.

I am also deeply appreciative of the support from Marquette Energy Analytics (MEA), whose sponsorship made this research possible. Tom Quinn and Dr. Ron Brown provided me with invaluable opportunities that were instrumental in developing my research skills. The experiences I gained through MEA helped bridge the gap between theoretical research and practical industry applications, deepening my understanding and strengthening my confidence in tackling real-world challenges.

Finally, I want to express my heartfelt thanks to my family—my mom and dad, my brother Sean, my sister Maura, and my girlfriend Kate. Your unwavering support, love, and belief in me have been my greatest source of strength. Mom and Dad, your encouragement and sacrifices laid the foundation for everything I have achieved. Sean and Maura, you have been my lifelong support system, always pushing me forward with humor and perspective. And Kate—your patience, kindness, and steadfast belief in me have meant more than words can express. This journey would not have been possible without you all by my side.

TABLE OF CONTENTS

ABSTRACT i

ACKNOWLEDGEMENTS i

TABLE OF CONTENTS ii

TABLE OF FIGURES vi

NOTATION ix

CHAPTER 1 1

Introduction to Natural Gas Data Hierarchies and Demand Forecasting 1

1.1 Contribution of Dissertation 2

1.2 Problem Statement 3

1.3 Temporal Disaggregation 4

1.4 Single Dimension Hierarchical Data Reconciliation 5

1.5 Cross-temporal Hierarchical Forecast Reconciliation 7

1.6 Outline of Remaining Chapters 8

CHAPTER 2 10

Literature Review 10

2.1 Natural Gas Domain Knowledge 10

2.2 Natural Gas Demand Forecasting Fundamentals 12

2.3 Time Series Disaggregation 16

2.4 Hierarchical Time Series Reconciliation 19

CHAPTER 3 26

Background 26

3.1 Challenges in Natural Gas Demand Forecasting 26

3.2 Temporal Disaggregation and the Need for Higher-Resolution Data 27

3.3 Ensuring Data Consistency through Reconciliation 28

3.4 External Factors Influencing Gas Demand 29

3.5 Summary 30

CHAPTER 4 31

Multi-source Iterative Shifting Disaggregation Algorithm 31

4.1 Objective of Iterative Shifting Disaggregation 31

4.2 Fundamentals of Temporal Disaggregation for Natural Gas Forecasting 33

4.3 Iterative Shifting Disaggregation Algorithm 35

4.3.1 Step 1 - Preparation of Time Series Inputs using Low-frequency Sampling 38

4.3.2 Step 2 - Iteratively Modeling Daily Time Series (High-frequency Sampling) 43

4.3.3 Step 3 - Constrained Load Shifting 46

4.4 Disaggregation Case Studies 51

4.5 Discussion of Disaggregation Results 62

4.6 ILS Disaggregation Algorithm Conclusion 66

CHAPTER 5 68

Single Dimension Hierarchical Reconciliation of Natural Gas Data 68

5.1 Objectives of Hierarchical Data Reconciliation of Natural Gas Data 69

5.2 Single Dimension Hierarchical Reconciliation..... 72

5.2.1 Spatial (Cross-sectional) Hierarchical Reconciliation 73

5.2.2 Temporal Hierarchical Reconciliation 76

5.3 Single Dimension Hierarchical Aggregation Constraints..... 79

5.4 Single Dimension Reconciliation Solution 80

5.5 Single Dimension Reconciliation Case Studies 81

5.6 Single Dimension Reconciliation Conclusion 87

CHAPTER 6 89

Cross-temporal Hierarchical Reconciliation of Natural Gas Forecasts 89

6.1 Cross-temporal Hierarchical Structure and Constraints 90

6.2 Cross-temporal Hierarchical Forecast Solution 93

6.3 Cross-temporal Hierarchical Reconciliation Case Studies 95

6.4 Cross-temporal Hierarchical Reconciliation Conclusion..... 105

CHAPTER 7 107

Discussion, Practical Implications, and Future Work..... 107

7.1 Summary of Contributions..... 107

7.2 Method Summaries and Performance Evaluation..... 109

7.2.1 Multi-Source Iterative Shifting Disaggregation (ISD) Algorithm..... 109

7.2.2 Single-Dimension Hierarchical Reconciliation 109

7.2.3 Cross-Temporal Hierarchical Forecast Reconciliation 110

7.3 Future Work and Research Directions 110

7.4 Dissertation Conclusions 112

BIBLIOGRAPHY 115

TABLE OF FIGURES

Figure 1.1 – Example of temporal incoherence between aggregated hourly, daily, and monthly time series. 6

Figure 2.1 – EIA natural gas data showing typical heat and base loads by plotting the daily consumption with average daily temperature. 14

Figure 4.1 – Low frequency Y and high frequency y time series. 33

Figure 4.2 – Time series A_i, interval durations $\mathcal{T}_i^A$, and daily index d. 39

Figure 4.3 – Time Series A naively disaggregated into daily series a. 40

Figure 4.4 – Three input time series A, B, and C with overlapping time intervals.42

Figure 4.5 – Each day d is aggregated into the y_d daily time series. 43

Figure 4.6 – Plot of y, formed from the daily aggregation of a, b, and c, compared to Step 3's estimated load profile y before any shifting occurs. 47

Figure 4.7 – Newly updated interval $\bar{a}_{d|shifted}$, original $\bar{a}_{d|unshifted}$, constrained interval estimates z_d and error ϵ_d. 49

Figure 4.8 – Three gas consumption load shifting steps carried out by the ISD algorithm. 50

Figure 4.9 – Evaluation of A_1 across all $nLRModels$. Daily series $\bar{a}_{d|unshifted}$. 51

Figure 4.10 – Plot of input time series A, B, and C over the 2021-2023 heating seasons. 54

Figure 4.11 – Plot of the first ten of 100 customers consumption time series from the Residential 100 dataset. 55

Figure 4.12 – Residual comparison of actuals vs forecasts for the NAV, TSR, GLS, and ISD disaggregation algorithm (Operating Area). 59

Figure 4.13 – Residential 100 residuals analysis for the NAV, ISD, TSR and GLS disaggregation algorithms. 61

Figure 5.1 – Illustration of an incoherent natural gas distribution scenario in which gas consumed in operating areas A, B, and C does not aggregate spatially or temporally. 71

Figure 5.2 – Hourly and daily components of hierarchical natural gas consumption series, Y. 73

Figure 5.3 – Two-level LDC hierarchies showing spatial aggregation structure. 74

Figure 5.4 – Structure of , formed by stacking observations from both hourly and daily time series. 75

Figure 5.5 – Two-level LDC hierarchies showing temporal aggregation structure. 77

Figure 5.6 – Structure of $y_{spatial,t}$, formed by stacking observations from both hourly and daily time series. 78

Figure 5.7 – Incoherency between daily and summed hourly gas demands. Similar incoherence is observed in spatial hierarchies. 83

Figure 6.1 – Structure of y_t, formed by stacking observations from each time series. 91

Figure 6.2 – Cross-temporal hierarchy diagram showing the aggregation relationships between hierarchical levels (A – blue, B – orange, C – yellow, $Total$ – purple). 92

Figure 6.3 – Mapping from cross-temporal diagram to summation matrix S. 93

Figure 6.4 – Hourly box plot results for the error distribution across different operating areas, comparing each reconciliation method (variance scaling, minimum trace, and naïve). 99

Figure 6.5 – Daily box plot results for the error distribution across different operating areas, comparing each reconciliation method (variance scaling, minimum trace, and naïve). 100

Figure 6.6 – Hourly reconciled forecasts compared to observed data, with the differences plotted below in corresponding dashed colors. 102

Figure 6.7 – Daily reconciled forecasts compared to observed data, with the differences plotted below in corresponding dashed colors. 103

NOTATION

Table 1: Notation used in this dissertation

Y_i	A low frequency problem time series that is nonuniformly sampled
X	Regression design matrix containing $P+1$ correlated variables x_p
$\hat{\beta}$	Regression weights, $P+1$ total
$x_{1,}\, x_2, \ldots, x_P$	P independent correlated variables measured at the target frequency of y
$\hat{y}$	Estimate of low-frequency time series Y, coherent
P	Total number of independent correlated variables
N_i	Total number of observations in Y_i
h	Forecast horizon
d	A daily index used to denote the days on which each observation occurs
A, B, C	Low frequency input time series with nonuniformly spaced time intervals
$A_{i,d}$, $B_{k,d}$, $C_{l,d}$	Observation of A at interval i, B at k, C at l on day d
N_D	The maximum number of total days each series A, B, or C spans.
$\bar{A}$, $\bar{B}$, $\bar{C}$	Naively disaggregated representations of A, B and C in daily form
$\bar{A}_j$, $\bar{B}_j$, $\bar{C}_j$	Observation of $\bar{A}$, $\bar{B}$, and $\bar{C}$ at day j, where $j = d$
$\hat{Y}_d$	Average daily flow series observation on day d
HDD	Heating Degree Days series
$HDDW$	Wind Adjusted Heating Degree Days
CDD	Cooling Degree Days
m	The total number of nodes in the hierarchy
n	The number of nodes at the lowest hierarchical level available
$S_{spatial}, S_{temporal}$	Aggregation matrices for spatial and geographical hierarchies
G	Reconciliation mapping matrix (incoherent to coherent data)

CHAPTER 1

INTRODUCTION TO NATURAL GAS DATA HIERARCHIES AND DEMAND FORECASTING

This dissertation examines how natural gas data hierarchies can enhance gas demand forecasting. Our goal is to assist Local Distribution Companies (LDCs) in generating accurate and coherent forecasts across temporal and spatial dimensions, thereby enabling more informed decision-making throughout their operations. To address this challenge, we focus on three interrelated but distinct sub-problems, each advancing the current state-of-the-art in gas demand forecasting solutions.

First, temporal disaggregation addresses the challenge of generating high-frequency data (e.g., daily) from low-frequency observations (e.g., monthly), which is crucial when available data do not match forecast needs. Second, single-dimension hierarchical data reconciliation ensures coherency within temporal or spatial hierarchies independently. Temporal reconciliation aligns data across different time levels—such as hourly, daily, and monthly—while spatial reconciliation ensures coherence within the hierarchical structure of LDCs' geographic regions. These independent transformations are applied to historical gas demand data to maintain coherence within each dimension. Finally, cross-temporal hierarchical forecast reconciliation addresses the simultaneous reconciliation of estimated natural gas demand forecasts across both temporal and spatial dimensions. This process aligns gas forecasts to achieve greater accuracy and reliability while ensuring consistency across all hierarchical levels.

Together, these methods enhance natural gas demand forecasting accuracy while maintaining coherence across the hierarchical structures of LDCs.

This chapter introduces the objective and contributions of this dissertation. First, we present the key contributions of this dissertation. Then, we examine temporal disaggregation and how disaggregation techniques are used in LDC operations to improve forecast accuracy. This discussion leads to the definition of data incoherence in the temporal hierarchical dimension, and later, incoherence in the spatial dimension. We then shift from data coherence to forecast coherence, showing how temporal and spatial hierarchical levels can be reconciled cross-temporally to gain forecast coherence across all hierarchical data levels. Finally, we offer a brief survey of this research as a whole and a summary of the remaining chapters.

1.1 Contribution of Dissertation

This dissertation contributes to the field of natural gas demand forecasting by addressing critical challenges through three novel advancements. First, the Multi-Source Iterative Shifting Disaggregation (ISD) algorithm improves the disaggregation of natural gas time series, significantly enhancing accuracy over traditional methods. The ISD algorithm can accommodate multiple input time series with nonuniform and overlapping intervals, ensuring coherent outputs that align with the original data granularity.

Second, single-dimension hierarchical reconciliation constraints tailored for natural gas distribution systems are developed. These constraints reconcile historical consumption data to improve demand forecast models and ensure coherence across temporal or spatial hierarchies independently.

Finally, the dissertation extends single-dimension data reconciliation by implementing Cross-temporal hierarchical forecast reconciliation constraints. Cross-temporal

constraints are used to simultaneously reconcile forecasts across both temporal and spatial dimensions. This novel approach enhances forecast accuracy and coherence, providing a robust tool for decision-making in natural gas distribution operations. These contributions collectively advance the accuracy and reliability of natural gas demand forecasts and support improved decision-making for LDCs.

1.2 Problem Statement

Local natural gas distribution companies (LDC) require accurate demand forecasts across various time periods, geographic regions, and customer class hierarchies. Forecasting challenges emerge when the required data granularity does not align with available data. For example, when only monthly data is available, but daily forecasts are required, the mismatch in temporal granularity affects model accuracy. Temporal disaggregation methods address this issue by transforming low-frequency time series into higher-frequency series.

When disparities arise between hierarchical data levels, LDC companies do not have a ground truth for parameter estimation or model learning, complicating both operational and strategic planning [1]. For example, the natural gas consumption at each of the 24 individual hours within a day should sum to the corresponding daily total. The monthly gas consumption of three adjacent neighborhoods should sum to the monthly gas consumed for the total area. These types of inconsistencies are known as “data incoherence,” and are prevalent in LDC hierarchies.

When forecasts at the hourly level do not aggregate to the forecasts at the daily level, this is referred to as "forecast incoherence." Incoherent forecasts can lead to suboptimal decisions, inaccurate resource allocation, and an increased risk of operational disruptions, thereby undermining the LDC's ability to meet the dynamic demands of their customers and maintain efficient operations [2].

To solve this problem, temporal hierarchical time series reconciliation methods are introduced. These methods constrain forecasts such that each temporal level is coherent. LDCs also forecast across non-temporal hierarchies, such as geographic hierarchies (i.e., neighborhoods, cities, states). Cross-temporal hierarchical forecast reconciliation methods are employed to gain coherence between two-dimensional hierarchies, showing that state-of-the-art gas demand forecast solutions are enhanced with time-series reconciliation techniques.

1.3 Temporal Disaggregation

Given a time series, we may need to forecast a more detailed, higher-frequency series, meaning we generate forecasts at a finer level of granularity. The most accurate forecasts are produced when the frequency of measurement matches the frequency demanded [1]–[3]. Temporal disaggregation addresses this by transforming low-frequency data (e.g., monthly) into higher-frequency series (e.g., daily).

Temporal disaggregation is used to generate additional, higher-frequency historical data to estimate parameters, reintroduce variability into a series that might have been smoothed through aggregation, and analyze the data at a resolution previously unavaila-

ble [4], [5]. Generally, disaggregation is an ill-posed mathematical problem with multiple solutions [6]. Current temporal disaggregation methods can only transform a single low frequency time series with uniform sampling [7]. They cannot disaggregate multiple time series with nonuniform sampling. To address this gap, this dissertation introduces the Multi-source Iterative Shifting Disaggregation (ISD) algorithm in Chapter 4, which addresses this limitation by disaggregating multiple nonuniformly sampled time series into a single high-frequency signal.

1.4 Single Dimension Hierarchical Data Reconciliation

Natural gas Local Distribution Companies (LDCs) measure gas consumption at various temporal frequencies, such as hourly, daily, and monthly. These measurements often result in multiple series that are temporally inconsistent because they come from independent sources, meaning lower frequency measurements do not sum to higher frequency values. This incoherency between hierarchical levels is known as data incoherence, as illustrated in Figure 1.1, which shows discrepancies between aggregated hourly, daily, and monthly time series.

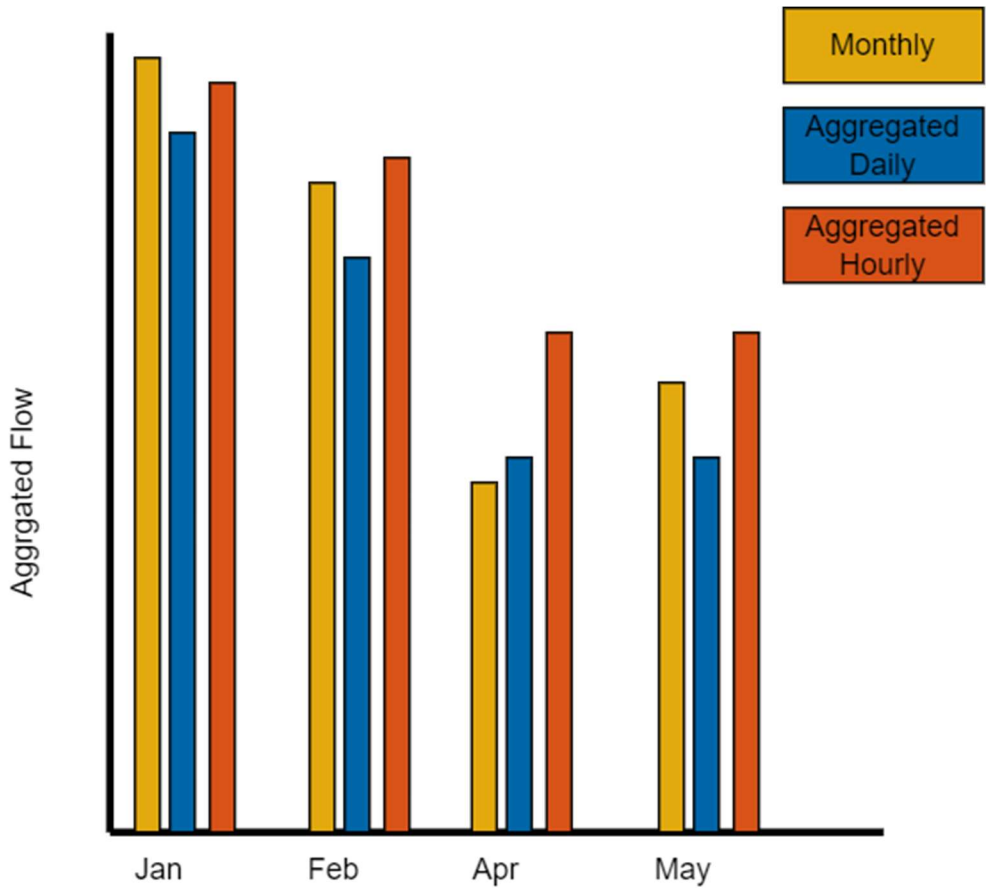

Figure 1.1 – Example of temporal incoherence between aggregated hourly, daily, and monthly time series.

LDCs do not know which time frequency series are most accurate, leading to potential operational risks and inefficiencies. Gas forecasters aim to leverage information available from each time series at each frequency to mitigate these issues. Hierarchical time series reconciliation offers the ability to blend all information available at different parts of the gas hierarchy to produce coherent gas data.

Despite time series reconciliation gaining popularity in demand-forecasting literature in recent years, it has not seen widespread adoption in LDC practice [8], [9]. The primary reason for this is that gas consumption data is inherently temporally incoherent and lacks a reliable basis for reconciliation. However, even with data incoherence, valuable

information can still be extracted from multi-hierarchical level gas consumption data to enhance demand forecast accuracy.

We propose an algorithm for reconciling natural gas data across temporal and spatial levels in Chapter 5. By addressing incoherent consumption data, this approach can produce more accurate and reliable demand forecast models that align with the physical and operational constraints of LDC systems.

1.5 Cross-temporal Hierarchical Forecast Reconciliation

Historically, natural gas forecasting models have been developed at a single level of aggregation [1]. Temporally, models are designed for specific timeframes, e.g., hourly or daily. In the context of gas operations, geographical operating areas are often subdivided into smaller sub-regions or customer classifications and independently modeled. However, it is unlikely that the forecasts at all hierarchical levels are coherent. For example, the underlying hourly and daily forecasts may not aggregate to the corresponding monthly value, and sub-regions gas forecasts may not sum to the total gas forecasted for the whole system.

The motivation for using hierarchical forecasting in this context arises from the complexity of natural gas distribution networks, their dependence on precise timing, and the interdependence of their various components. Achieving aligned decision making is particularly challenging when demand forecasts are incoherent [4]. LDCs rely on hourly, daily, and monthly forecasts for operational and strategic planning, often serving diverse customer bases across large geographic areas [5]. A lack of coherence in gas demand forecasts can disrupt operations, customer relations, finances, and overall sustainability.

Everyday decisions tied to supply procurement, inventory, and distribution scheduling heavily depend on accurate demand forecasts [6]. Hence, incoherent forecasts can lead to poor resource allocation and operational inefficient choices. To address these incoherencies, hierarchical reconciliation methods provide a structured approach to ensuring data consistency across different time scales and spatial dimensions.

Chapter 6 demonstrates how state-of-the-art gas demand forecast solutions can be improved by arranging the forecastable components of gas distribution into a cross-temporal hierarchical structure to produce a new set of reconciled forecasts. Hierarchical time series forecasting involves generating forecasts at every level within the hierarchy. Forecast reconciliation, as implemented in this work, acts as an a posteriori refinement of existing forecasts, where any forecasting method can be used to generate the base forecasts. In Chapter 6, we develop a framework to reconcile forecasts subject to a set of aggregation constraints, ensuring forecast coherence across the hierarchies. The aggregation constraints reflect a realistic spatio-temporal structure of an LDC gas operating area.

1.6 Outline of Remaining Chapters

This dissertation is structured into seven chapters. Each chapter builds upon the previous to progressively develop a comprehensive framework for improving natural gas demand forecasting. Below is an overview of each chapter's contributions. The remaining portion of this dissertation begins with a review in Chapter 2 of current literature and work done in the field of energy forecasting and time series analysis. Chapter 3 introduces the background necessary for understanding natural gas demand forecasting and relevant notion. Chapter 4 presents our temporal disaggregation work known as the

Multi-source Iterative Disaggregation Algorithm. Chapter 5 sets up the hierarchical framework required to reconcile natural gas data across temporal and spatial individually (single-dimension hierarchical data reconciliation). Chapter 6 presents the Cross-temporal Hierarchical Reconciliation structure used to reconcile incoherent forecasts across both the combined spatial and temporal dimensions. Finally, Chapter 7 provides a summary of key findings, discusses future research directions, and offers final remarks.

CHAPTER 2

LITERATURE REVIEW

This chapter presents the history and context of our work. We analyze, interpret, and critically evaluate the existing literature on time series forecasting, disaggregation, and hierarchical reconciliation. We begin with the relevance and motivation for this work and then provide a background on natural gas demand forecasting, time series analysis, temporal disaggregation, and hierarchical reconciliation. Finally, we discuss mathematical advancements in modern time series forecasting techniques and connect them to alternative methodologies introduced in Chapter 3.

2.1 Natural Gas Domain Knowledge

The natural gas industry is a dynamic and competitive business. Employing over 10.3 million people in North America, the gas market continues to grow due to ever-increasing opportunities in the extraction and processing, transportation, and delivery of this natural resource [10], [11]. Following the deregulation of natural gas pricing at the wellhead in 1979, the price at which natural gas is exchanged depends on supply and demand [12]. Using data collected from historical gas consumption and other independent variables, past interactions can be modeled to provide valuable insights on future supply and demand interactions. These insights, referred to in this work as anticipated gas demand, are essential input for decision-making processes in the natural gas industry. This section provides an overview of the natural gas industry and establishes the relevance of this research.

The terms "load," "consumption," and "demand" are commonly used across natural gas demand forecasting literature when describing gas distribution. Gas load refers to the volume of gas that flows through the distribution pipeline. Gas consumption is the amount of gas consumed by the customers and is measured with a gas meter at the customer location. The difference between gas load and consumption is referred to as lost and unaccounted for gas, which can result from fueling compressor stations, leaks, theft, or measurement and accounting errors during transport and delivery [13]. Gas demand, in contrast, represents the consumption that would occur if no service curtailment took place within the distribution network. Service curtailment is a partial or complete temporary suspension of gas deliveries to a customer or customers because of a projected or actual supply or pipeline-capacity constraint [14]. An example of service curtailment was seen in the Texas extreme cold event of February, 2021. During this event, well-heads froze in West Texas, prompting Local Distribution Companies (LDCs) to curtail gas delivery to customers with interruptible service agreements. In a perfect gas delivery system, load would match demand.

Natural gas is consumed in different ways and at varying rates throughout the year. The uses of natural gas vary from sector to sector, but mainly include space and water heating, cooking, clothing dryers, electric power generation, industrial processes, and as a vehicle fuel [3], [15], [16]. As uses become more widely applicable, the task of reliably supplying gas to meet these needs becomes complex.

A distributor organizes, plans, and distributes the gas through a pipeline network. To meet the rising demand for natural gas fueled by population growth and industrial expansion, LDCs adopt new strategies and technologies to ensure reliable service. An LDC

[11] is an energy utility responsible for maintaining a portion of the supply grid in a particular geographical area and delivering gas to their customers. To deliver gas to all customers on their distribution grid successfully, LDCs must predict the gas demand. Forecasts can span horizons ranging from a few minutes to decades and cover multiple geographic areas [17].

Accurate forecasts are critical because both under- and over-prediction come with financial costs. Under-forecasting leaves utilities with insufficient gas on standby, potentially forcing them to curtail usage or purchase additional gas on the expensive spot market [3]. Over-forecasting, on the other hand, can lead to costly storage fees or penalties for excess gas in the distribution pipeline [18]. If an estimate of natural gas load for a particular region is known in advance, strategies can be employed by the LDC to optimize the infrastructure used to meet demand. With accurate forecasts, LDCs can better manage commodity contracts, prepare for peak load events, and streamline daily operations [19]. Additionally, reducing an LDC's operational costs through precise forecasts ultimately lowers costs for customers. Accurate demand forecasts benefit all parties and have driven significant efforts to understand the variables influencing gas consumption.

2.2 Natural Gas Demand Forecasting Fundamentals

Studies throughout the natural gas demand forecasting literature [9], [17] use three categories of variables used in modeling demand: meteorological, economic, and historical. We begin with a description of meteorological variables. Seasonal variability significantly influences natural gas consumption [20]. Key meteorological variables—temperature, wind speed, and humidity—are widely recognized as strong predictors of

natural gas demand [2], [21], [22]. One of the earliest works to identify this correlation was Jager's 1983 review [23] of how climate and energy systems interact. Since that publication, the inclusion of weather inputs in demand modeling has become standard practice. For instance, 22 out of 26 studies published between 2000 and 2018 on distribution-level gas demand forecasting included weather variables as inputs [9].

The consistent use of weather variables in gas demand modeling has led to terminology that captures the seasonal trends in demand. Total load is typically divided into baseload, representing gas used for non-temperature-related purposes, and heat load, which strongly correlates with temperature and defines the typical gas load profile [24]. Public data from the U.S. Energy Information Administration (EIA) [25] demonstrates the nonlinear effects of temperature on consumption, as shown in Figure 2.1. Kaefer explains [3] explains that at warmer temperatures, demand can often be modeled linearly as baseload. However, as temperatures drop, furnaces are activated, driving up consumption and introducing a nonlinear relationship. In Figure 2.1, each data point corresponds to recorded daily consumption (vertical axis) and average daily temperature (horizontal axis).

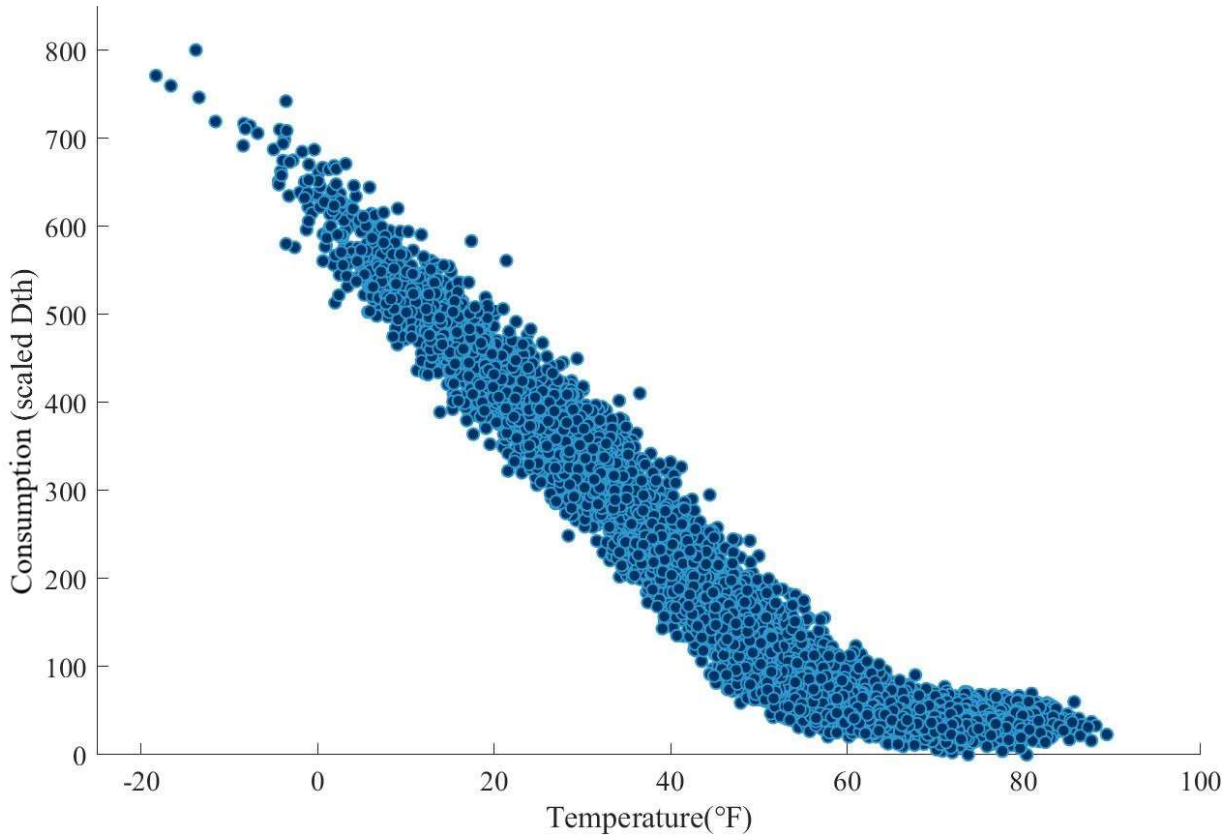

Figure 2.1 – EIA natural gas data showing typical heat and base loads by plotting the daily consumption with average daily temperature.

As shown in Figure 2.1, gas consumption decreases linearly as temperatures rise, leveling off between 55°F and 65°F. At that temperature range, consumption levels off and reveals the baseload. The baseload—gas used for cooking, clothes drying, and industrial processes— persists throughout the year. During cooler months, heat load introduces a periodic trend into long-term series as seasons shift. To model the nonlinear effect of the transition from baseload to heatload, several published works [26]–[28] have found success including transformed-weather variables in demand modeling.

Other weather-related variables have also been employed in modeling natural gas demand. Although temperature, wind speed, and humidity are the most frequently used, studies have also incorporated variables such as atmospheric pressure, cloud cover, wet

bulb temperature, and solar irradiance. Taspinar [29], Hribar [30], and Sailor [20], [31] found that higher wind speeds often lead to increased gas consumption on cold days compared to equally cold but less windy conditions. Wei [32] employs a "daily relative humidity ratio" alongside other weather variables, while Merkel [24] uses the closely related variable "dew point" in his approach. These works take advantage of the fact that higher humidity heating days commonly result in higher consumption rates [3]. Exogenous and econometric variables are discussed next.

Soldo [17] identifies five spatial levels on which consumption of natural gas is measured: the world, state/country, regional or distribution, sector, and individual customer. At the global level, studies like Mohr's [33] long-term forecast of hydrocarbon reserves and Valero's [34] assessment of Earth's remaining mineral resources put the current rates of consumption into perspective when compared to remaining natural reserves. In these works, more weight is placed on econometric variables than short-term weather variables to concentrate on consumption trends effected by population growth, country development, and available infrastructure. Kan's [35] explanation of the critical position of natural gas in global energy supply whether the world's gas infrastructure can scale enough to support projected plans. Nations worldwide are implementing clean energy policies, while scaling up infrastructure and increasing consumption.

At the national level, researchers examine how policy, infrastructure, and resource availability influence natural gas demand. For instance, Fumitaka [36] examines the relationship between China's rates of consumption and economic development. Kilic [37] investigates Russia's increasing energy demand and the infrastructure required to support it. At smaller scales, Sari [38] uses industrial output variables to disaggregate sector-level

consumption time series. Other publications attempt to forecast consumption at the state level [20], while some are only concerned with smaller groups of residential consumers [39]. These varying scales highlight the importance of econometric data, particularly for long-term forecasts where trends outweigh short-term meteorological influences.

The final category of variables commonly seen in the gas demand forecasting literature is categorized as historical [9]. Historical variables capture past trends and provide additional granularity for demand modeling. These variables may overlap with meteorological or econometric categories or represent unique patterns that reflect historical profiles. Sanchez-Ubeda [40] presents a novel statistical decomposition of different historical profiles of industrial end-use natural gas consumption to forecast future demand. In this approach, Sanchez was able to capture and model demand patterns for medium- to long-term forecast horizons. Musilek et al. [41] combined climatic conditions with historical gas consumption profiles to solve the problem of seasonal dependency, identifying the start and end of heating seasons. By combining this information with a statistical mixture model, they successfully predicted the non-stationary behavior typical of gas consumption time series.

2.3 Time Series Disaggregation

Time series disaggregation enables forecasting at finer levels of granularity by breaking down lower-frequency data into higher-frequency components. Forecasts are most accurate when the measurement frequency matches the required forecast frequency [1]–[3]. Temporal disaggregation generates additional, high-frequency historical data to better estimate parameters, reintroduces variability into a series that might have been

smoothed through aggregation, and enables analysis at a previously unavailable resolution [4], [5]. There are two general approaches to distributive disaggregation: pure time series dynamic models, which are less common, and methods relying on independent correlated series at the desired frequency, which are more widely used [42]. This section begins with single time series dynamic models.

Pure time series dynamic disaggregation algorithms evaluate the data without incorporating exogenous information. These methods assign proportions of low-frequency measurements to higher-frequency components. Chan's naïve disaggregation approach [43] provides a simple but effective method. Chan's method disaggregates yearly time series into quarterly values using historical proportions to equally split annual values among quarters. This technique can be generalized to disaggregate any low-frequency interval into higher-frequency intervals. Chan's method disaggregates a yearly time series into quarterly values by using historical proportions to split annual values equally among quarters. This technique can be generalized to disaggregate any low frequency interval to a higher frequency interval. However, this method assumes uniform data variation and fails to account for variability across higher-frequency intervals. As a result, naïve disaggregation is often treated as a baseline method in the literature and is typically applied to single-series cases.

Wei and Stram address the challenge of recreating the absent variability in the naïve approach [5]. They develop an aggregate ARIMA model [44] that leverages high-frequency data. This model uses its autocovariance structure to estimate the unknown autocovariance structure of unobserved high-frequency series. Once acquired, this structure

reintroduces variability to the unobserved series. Wei and Stram show that the autocovariance structure of a model at high frequency can be uniquely derived from an aggregated model if the aggregation order (number of summed intervals constituting the aggregated series) evenly divides the total number of observations in the aggregated series. They conclude that disaggregation methods cannot achieve optimality unless they incorporate intrinsic time series correlation characteristics [5]. Although their work examines autocorrelation within a single time series, many studies incorporate correlated exogenous series into disaggregation methods.

Chow and Lin's Generalized Least Squares (GLS) disaggregation method uses exogenous variables [45]. When auxiliary data is available, it is advantageous to use disaggregation methods that integrate aggregated and disaggregated information rather than relying solely on aggregated data [46]. Similar to Chan [43], Chow and Lin disaggregate a yearly time series into quarterly values, a common practice in econometrics [47]. However, rather than using historical proportions or lagged versions of the original series, they estimate GLS coefficients with correlated quarterly series, which generates variation in the disaggregated time series. To avoid linear bias, their method constrains the disaggregated data so that it sums to the aggregated observations. Leveraging multiple related high-frequency series has since become a popular approach for disaggregation [47].

Moauro introduces a state-space disaggregation method that uses seemingly unrelated structural time series (SUTSE) to estimate a higher frequency time series [4]. The SUTSE model is particularly flexible, supporting a variety of temporal disaggregation types (e.g., yearly to monthly or daily to hourly) on raw or seasonally adjusted data, making it a more robust choice than univariate regression-based approaches. This approach

effectively handles the constraint that all disaggregated intra-period data must sum to the original aggregated period. Unlike traditional methods, Moauro's approach simultaneously disaggregates the time series and extracts seasonal trends, enhancing its efficiency and robustness. However, in cases with high noise, other approaches may achieve better accuracy [9].

The Time Series Reconstruction (TSR) algorithm [6] uses multiple related high frequency time series in its disaggregation and was developed for the specific purpose of disaggregating natural gas consumption time series. TSR is a regression-based approach utilizing Generalized Least Squares as outlined by Chow and Lin [45]. The TSR algorithm's objective is to accurately disaggregate natural gas billing cycle data, especially during extreme cold events. Unlike other disaggregation methods, TSR does not impose forecast coherence constraints. This flexibility enables utilities to prioritize forecast accuracy during peak demand scenarios. The TSR disaggregation algorithm inspired the development of our novel Multi-source Iterative Shifting (ISD) disaggregation algorithm presented in Chapter 4. The ISD algorithm extends TSR by integrating principles of pure time series dynamic models and exogenous-variable approaches with hierarchical constraints to obtain a final set of coherent estimates.

2.4 Hierarchical Time Series Reconciliation

A hierarchical time series is a collection of time series organized across multiple levels of aggregation. These hierarchies can be structured along dimensions such as time (e.g., hourly, daily, monthly), space (e.g., neighborhoods, cities, states) or a grouping characteristic that supports hierarchical relationships within the data. At the most aggregated

level, the series represents the total value (e.g., total gas demand), while lower levels capture increasingly granular data or forecasts (e.g., demand for specific regions or hours). Incorporating hierarchical structures facilitates a comprehensive understanding of time series data by capturing both the behaviors of individual series and their interactions within larger aggregates [48].

Hierarchical data reconciliation ensures that observed data within a hierarchy are consistent, or coherent, across all levels. Similarly, hierarchical forecast reconciliation ensures that forecasts at all levels of a hierarchy align. Coherence is achieved when either the sums of the forecasts or the observed data across all levels equals the value at the most aggregated level [49]. For example, in coherent temporal hierarchies, the total hourly gas consumption for a day sums to the daily total. In coherent spatial hierarchies, gas consumption for individual neighborhoods must sum to the total consumption for the city. These reconciliation processes are mathematically identical; the difference lies in whether observed data or estimates are being reconciled and in the specific structure of the hierarchy.

Hierarchical reconciliation can be applied to any data hierarchy, whether reconciling observed historical data (addressed in Chapter 5) or reconciling forecasts (discussed in Chapter 6). Both processes rely on aggregation constraints defined by the hierarchy's structure to enforce coherence and produce consistent results. This alignment is critical for accurate analysis and decision-making.

The gas distribution problem fits naturally into a hierarchical framework - with the total gas demanded being the aggregate-most level in the hierarchy. To achieve coherence, hierarchical reconciliation employs aggregation constraints tailored to the structure of the hierarchy. These constraints are critical for reconciling observed discrepancies in data or

ensuring forecast alignment. Various hierarchical aggregation methods have been developed. These approaches, generally categorized as cross-sectional, temporal, and cross-temporal [50], [51], offer distinct strategies for reconciling forecasts across hierarchy levels while improving accuracy within each level. The gas demanded is subdivided into temporal and spatial partitions of gas consumption uniquely nested below. Early hierarchical methods employed a form of directional-structural scaling, which involves generating estimates for a single gas demand series and linearly combining the forecasts to obtain demand estimate series for the other levels in the LDC hierarchy [52].

The bottom-up (BU) reconciliation framework [53] reconciles gas data at the lowest level of the hierarchy and sums them to produce estimates for the higher levels, e.g., hourly forecasts aggregated to coherent daily forecasts. The top-down (TD) approach operates in the opposite direction. The top level of the hierarchy is disaggregated into the lower, more granular levels, e.g., total gas demanded disaggregated into underlying operating areas [54]. The middle-out (MO) approach takes a time series from the middle of the hierarchy and both aggregates and disaggregates the series into a coherent structure [55]. The BU, TD, and MO approaches originate from decomposition and smoothing techniques in spatial econometrics [56]. While these single-series methods are computationally efficient and produce coherent results, they do not consider the inherent correlation structure of the hierarchy.

The factors affecting gas demand at one hierarchical level are correlated with the gas demanded at other levels. Sánchez-Úbeda presents the first algorithm capable of balancing gas data using this correlation in their multi-horizon gas demand forecast decomposition [40]. Hierarchical works progressed from single-level (BU, TD, MO) to combination

(COM) approaches to leverage dependencies between different levels [57]. COM approaches [8], [51], [58] are implemented such that they independently model and produce forecasts for all gas demand series in the hierarchy. The forecasts made at these levels are likely to be incoherent, but much more likely to be accurate than forecasts obtained via aggregation or disaggregation of a single forecasted series [59]. Numerous studies have concentrated on generating accurate gas forecasts at a single-level (ignoring aggregation constraints) [9], [60], [61]. As a result, COM methods use these specialized forecasts, considering both time and space, and reconcile all levels within the hierarchy simultaneously to attain overall coherence.

Hyndman et al. pioneered early work in optimal reconciliation methods by describing the statistical quantities of the BU, TD, and MO methods and identifying a minimum variance unbiased estimator to produce coherent forecasts [8], [58]. Wang successfully applied the insights gained from optimal reconciliation to a grouped time series, finding an optimal weighted-least-squares solution (now known as variance-scaling) [49]. These methods summarize the correlations and interactions among the hierarchical levels linearly to optimally combine and reconcile the forecasts. Wickramasuriya [8] and Athanasopoulos [52] show how ad hoc adjustments can be incorporated into the optimal combination process using important covariates, sub-series trends, and domain knowledge. While ample empirical results exist to support the use of COM methods over structural-scaling approaches [51], [57], early optimal reconciliation methods were limited to applications with smaller-hierarchies either bound to the unit of analysis, such as geographical region and customer type, or to the unit of time [62]. The gas distribution problem depends on adapta-

bility across various hierarchical variables such as time, geography, customer class, and average yearly usage. Consequently, it is crucial for gas demand forecasts to be coherent across the various hierarchical variables, including temporal and spatial dimensions.

Given that the physical process of natural gas distribution has both temporal and spatial aggregation constraints, the forecasts should also adhere to these constraints. Cross-sectional hierarchies are constructed from multiple contemporaneous time series, such as geographical divisions or customer groupings [63]. Cross-sectional methods focus on achieving coherence among various spatial elements or series within a specific context, reconciling data across different sections or units at a single point in time [50]. van Erven & Cugliari were among the first to apply a cross-sectional forecast reconciliation method on energy data [57], [64]. Specifically, they implemented a Game-Theoretically Optimal (GTOP) reconciliation method for electricity demand data disaggregated into 17 tariff groupings. Bai and Pinson also propose a distributed reconciliation method based on GTOP in their application of Day-Ahead Wind Power Forecasting [65]. Gawel implements a cross-sectional global and local approach for gas consumption, specifically focusing the distribution infrastructure in Poland [66]. However, they chose to reconcile cross-sectionally across the spatial dimension of their data. The cross-sectional reconciliation methods focus on achieving coherence between forecasts at different aggregates, but not across different frequencies.

Temporal hierarchies are constructed from one or more time series by means of non-overlapping temporal aggregation [53]. Early examples of temporal aggregation were implemented to overcome limited memory availability by smoothing daily stock time se-

ries into weekly [67]. Jeon, Panagiotelis, and Petropoulos concentrate on producing temporally coherent probabilistic electricity demand forecasts [68]. Temporal reconciliation methods leverage time series modeling techniques to refine forecasts at different frequencies [62]. Theodosiou investigates how to combine independent, temporally incoherent forecasts using different deep learning architectures in their forecast refinement. Di Fonzo implements the closest industrial process to use hierarchical forecasting with photovoltaic power generation [11]. Cross-temporal methods combine elements of both cross-sectional and temporal approaches, emphasizing the incorporation and evolution of temporal patterns in a coherent hierarchy [69].

Cross-temporal reconciliation harmonizes and ensures coherence across temporal and spatial hierarchies [58]. Kourentzes et al. demonstrate how to improve forecast accuracy by exploiting relationships in both the cross-sectional and temporal hierarchies in forecasting Australian Tourism data [50]. The tourism forecasting problem resembles the geographical and temporal constrains of natural gas distribution [50]. Spiliotis offers a non-linear perspective of the problem of hierarchical reconciliation, incorporating the constraints of hierarchical time series forecasting with machine learning forecasting techniques [51].

Natural gas distribution represents a real-world forecasting problem that theoretically fits the hierarchical forecasting model but has not yet been explored. While theoretical works showing the uses of HTS are common, particularly using the Australian Tourism data set [50], no published works, to our knowledge, demonstrate the effectiveness of applying hierarchical constraints to more dynamic problems such as natural gas distribution. Our study addresses this gap by explicitly outlining how the cross-temporal hierarchy is

formed for gas distribution, demonstrating how to effectively combine the spatial and temporal dimensions of the demand forecasting problem.

Hierarchical time series reconciliation offers a robust framework for handling complex hierarchies and taking advantage of dependencies between different aggregates, thereby providing more accurate and coherent forecasts, contributing to better planning, resource allocation, and decision-making [50].

CHAPTER 3

BACKGROUND

Natural gas demand forecasting presents unique challenges due to the irregular nature of data collection, the hierarchical structure of consumption records, and the operational constraints faced by local distribution companies (LDCs). Effective forecasting requires addressing inconsistencies in historical data, ensuring coherence across different levels of forecast aggregation, and incorporating external influences such as weather variability. This chapter provides the necessary background to contextualize the methods introduced later in this dissertation, with a focus on the key challenges in natural gas demand modeling, the role of temporal disaggregation, and the importance of maintaining hierarchical coherence in both historical data and forecasted values.

3.1 Challenges in Natural Gas Demand Forecasting

Gas consumption data is collected through customer billing cycles, which do not follow a uniform schedule. Some meters are read monthly, while others are read every two months or quarterly, leading to inconsistently sampled data. This nonuniformity presents a challenge for forecasting models that require regularly spaced observations to make accurate predictions. Additionally, gas consumption fluctuates significantly based on weather conditions and customer usage patterns, making short-term and long-term forecasting highly dependent on external factors.

In addition to irregular meter readings, LDCs often lack a structured approach for organizing demand data at different levels of aggregation. While it is possible to arrange

natural gas consumption records into hierarchical structures, this is not common practice. Consumption data is typically analyzed at separate levels such as total system demand or customer-level demand without explicitly maintaining relationships between them. As a result, forecasts generated at different resolutions may not align, leading to discrepancies in operational planning.

LDCs must generate forecasts for multiple purposes, including daily load planning, monthly procurement, and long-term infrastructure investment. These different time horizons require demand estimates at varying levels of granularity. A forecasting model using monthly data may fail to capture the variability required for daily predictions, while one relying on hourly data may not provide reliable long-term projections. The need to bridge gaps between different levels of data aggregation is a recurring challenge in natural gas forecasting.

3.2 Temporal Disaggregation and the Need for Higher-Resolution Data

Forecasting models require data at a resolution that matches operational needs. However, gas consumption data is often available at lower frequencies, such as monthly meter readings, while decision-making processes require estimates at daily or hourly intervals. Temporal disaggregation bridges this gap by transforming coarse data into a finer resolution, allowing for more detailed analysis and improved forecasting accuracy.

Early approaches to temporal disaggregation relied on uniform distribution methods, assuming that demand remained constant within an aggregation window. While these methods provide a straightforward solution, they fail to capture variations in consumption that occur within each period. More advanced techniques incorporate external

information, such as weather patterns, to generate more realistic estimates that better reflect actual gas consumption.

Later chapters introduce the Multi-Source Iterative Shifting Disaggregation (ISD) algorithm, which improves upon existing methods by accommodating multiple overlapping time series with nonuniform sampling intervals. Unlike traditional approaches that treat disaggregation as a static process, ISD dynamically adjusts estimates based on observed patterns, improving the alignment of disaggregated data with real-world consumption behavior.

3.3 Ensuring Data Consistency through Reconciliation

Recorded gas consumption values are often inconsistent across different levels of aggregation, leading to discrepancies that can impact forecasting accuracy [6]. Ideally, hourly demand should sum to daily totals, and daily demand should sum to monthly consumption. Similarly, demand recorded at the neighborhood level should align with citywide and total system consumption. However, in practice, these relationships are not always maintained due to differences in data collection methods, reporting errors, and operational adjustments.

Forecasts generated from inconsistent data can propagate these discrepancies, leading to misaligned projections that affect procurement, storage, and distribution decisions. For example, if regional demand estimates do not align with total system consumption, LDCs may make inefficient supply adjustments that increase operational costs. Similarly, inconsistencies in historical records can introduce biases into forecasting models, distorting demand estimates and reducing predictive reliability.

Reconciliation is the process of adjusting recorded data to ensure that it remains internally consistent across different levels of aggregation. This dissertation introduces Single-Dimension Hierarchical Reconciliation, which enforces coherence within either a temporal or spatial hierarchy, and Cross-Temporal Hierarchical Forecast Reconciliation, which extends this framework to simultaneously reconcile both dimensions. By applying these reconciliation techniques, inconsistencies can be corrected before forecasts are generated, improving the overall reliability of demand estimation and decision-making processes.

3.4 External Factors Influencing Gas Demand

Weather conditions play a crucial role in shaping natural gas consumption patterns [20]. Temperature fluctuations drive heating and cooling demand, while factors such as wind speed and humidity influence the rate at which buildings lose or retain heat. Cold weather increases the demand for residential and commercial heating, leading to higher gas consumption. Conversely, mild temperatures reduce demand, as less energy is required for heating. Incorporating these variables into forecasting models improves predictive accuracy by accounting for seasonal and short-term variations in usage [44].

Heating Degree Days (HDD) and Cooling Degree Days (CDD) are commonly used metrics to quantify temperature-driven demand. HDD measures the extent to which daily temperatures fall below a given threshold, indicating the need for heating, while CDD captures the cooling demand for temperatures above the same baseline. Wind-Adjusted Heating Degree Days (HDDW) refine this approach by incorporating wind speed effects, improving the accuracy of heating load estimates.

Beyond temperature effects, seasonal variability impacts demand at different times of the year. During winter, natural gas consumption spikes due to heating requirements, whereas in summer, demand is often lower except for regions where gas-fired power plants supply electricity for air conditioning. Wind speed and humidity also contribute to variations in consumption, as wind can accelerate heat loss from buildings, increasing the need for heating, while higher humidity levels may enhance insulation effects.

By integrating these external factors, forecasting models can better reflect real-world consumption trends, enhancing their ability to predict demand fluctuations and optimize resource allocation.

3.5 Summary

Accurate natural gas demand forecasting requires addressing inconsistencies in recorded data, improving the resolution of available consumption records, and incorporating external influences such as weather variability. Temporal disaggregation provides a method for generating high-resolution demand estimates, while reconciliation techniques ensure that recorded data is internally consistent. The next chapters build on this foundation: Chapter 4 introduces the Multi-Source Iterative Shifting Disaggregation (ISD) algorithm, Chapter 5 covers single-dimension hierarchical reconciliation, and Chapter 6 addresses cross-temporal hierarchical forecast reconciliation to enhance forecast reliability.

CHAPTER 4

Multi-source Iterative Shifting Disaggregation Algorithm

Accurate time series forecasting often requires higher temporal resolution than that provided by available data, such as when daily forecasts are needed from monthly data. Existing temporal disaggregation techniques, which typically handle only single, uniformly sampled time series, have limited applicability in real-world, multi-source scenarios. This chapter introduces the Iterative Shifting Disaggregation (ISD) algorithm, designed to disaggregate multiple, nonuniformly sampled time series into a single, coherent high-frequency signal. ISD operates in an iterative, two-phase process: a prediction phase that uses multiple linear regression to generate high-frequency series from low-frequency data and correlated variables; followed by an update phase that redistributes low-frequency observations across high-frequency periods. This process repeats, refining estimates with each iteration cycle. The ISD algorithm's key contribution is its ability to disaggregate multiple, nonuniformly spaced time series with overlapping intervals into a single daily representation. In two case studies using natural gas data, ISD successfully disaggregates billing cycle and grouped residential customer data into daily time series, achieving a 1.4–4.3% WMAPE improvement for billing cycle data and a 4.6–10.4% improvement for residential data over existing methods.

4.1 Objective of Iterative Shifting Disaggregation

A common challenge in forecasting is that the required forecast granularity often does not match the granularity of the observed data. For example, a daily forecast may be needed, but only historical monthly data is available. To address this, low-frequency time

series data can be temporally disaggregated into higher frequency series. Current temporal disaggregation methods, however, typically transform only a single, uniformly sampled time series [7]. These approaches do not accommodate the disaggregation of multiple time series with nonuniform sampling. This chapter introduces the Iterative Shifting Disaggregation (ISD) algorithm, which can disaggregate multiple nonuniformly sampled time series into a single, high frequency signal. The ISD algorithm maintains coherence between aggregation levels, i.e., the disaggregated time series can be re-aggregated to the original low frequency time series.

We evaluate the effectiveness of the ISD algorithm in two case studies using natural gas data. The first study focuses on disaggregating low-frequency billing cycle time series from approximately monthly to daily frequency. The second case study focuses on disaggregating the time series of 100 residential customers' monthly bills to obtain a single daily time series representing the aggregated consumption of all residents (i.e., neighborhood consumption). The primary goal of each evaluation is to transform disparate billing cycle time series into consistent daily time series, enabling more precise analysis and forecasting.

The ISD algorithm is an iterative two-phase process. The prediction phase uses multiple linear regression to generate high frequency time series using the observable low frequency time series and independent correlated high frequency signals. The shifting phase, in a piecewise linear manner, redistributes low-frequency observations among high-frequency periods. This two-phase process continues iteratively, with each subsequent prediction phase modeling the updated estimates produced by the shifting phase.

4.2 Fundamentals of Temporal Disaggregation for Natural Gas Forecasting

The temporal disaggregation problem arises when a low frequency time series does not meet the granularity required for accurate forecasting or analysis. This section presents the single-series disaggregation and introduces the notation used throughput Chapter 4. Let Y be a low frequency problem time series that is nonuniformly sampled,

$$Y = \{ Y_i, \ \ i = 1, \dots, m\}, \tag{4.1}$$

and let y be the corresponding higher frequency time series that is uniformly sampled.

$$y = \{ y_j, \ \ j = 1, \dots, n\}. \tag{4.2}$$

Figure 4.1 illustrates an example temporal relationship between these two series, where $\mathcal{T}_i$ is the interval spanned by Y_i.

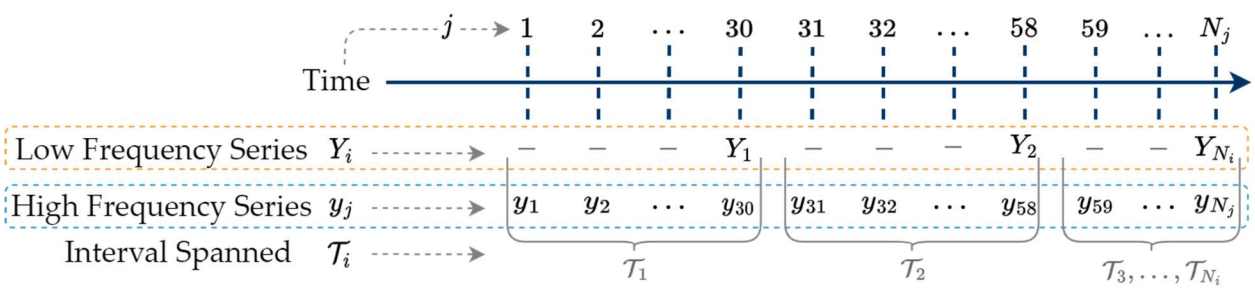

Figure 4.1 – Low frequency Y and high frequency y time series.

Figure 4.1 shows time interval $\mathcal{T}_1$ containing the first low frequency observation Y_1 and the first thirty high frequency observations $[y_1, \ \dots, y_{30}]$. Similarly, time interval $\mathcal{T}_2$ contains low frequency observation Y_2 and high frequency observations $[y_{31}, \ \dots, y_{58}]$.

Let $\mathcal{T} = [\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_m]$ represent all m intervals of Y. Let $\mathcal{A}$ be an aggregation operator such that

$$Y = \mathcal{A}(y, \mathcal{T}), \qquad \text{where} \qquad Y_i = \sum_{j \in \mathcal{T}_i} y_j. \tag{4.3}$$

Aggregation operator $\mathcal{A}$ sums the high frequency time series y for each interval $\mathcal{T}_i$ to yield low frequency time series $Y = \{ Y_i \}$. The disaggregation operator, $\mathcal{A}^{-1}$, takes a low frequency series and creates a corresponding high frequency time series.

$$y = \mathcal{A}^{-1}(Y, \mathcal{T}). \tag{4.4}$$

The ISD algorithm processes nonuniformly sampled data with irregular intervals ranging from 4 to 40 days. It also handles gaps in data by initializing missing values with zeros during the preparation step. This ensures alignment of time series for subsequent disaggregation and shifting phases. An example disaggregation would take a monthly time series (e.g., gas meter readings) and create the corresponding ~30 daily consumption estimates. The disaggregation operator $\mathcal{A}^{-1}$ is not uniquely determined, so we rewrite (4) as

$$\hat{y} = \mathcal{A}^{-1}(Y, \mathcal{T}) + \epsilon. \tag{4.5}$$

To address the under-constrained nature of $\mathcal{A}^{-1}$ [6] [70], additional constraints are applied to provide a near-unique definition of $\mathcal{A}^{-1}$ [47]. Distributive disaggregation techniques constrain $\mathcal{A}^{-1}$ such that sum of $\hat{y}_{j \in \mathcal{T}_i}$ equals Y_i [4], [47]. This constrained relationship is referred to as forecast coherence and ensures that the high frequency time series $\hat{y}$ maintains the same properties of the low frequency observable Y [71]. Let J be the

sum of square errors between y and estimate $\hat{y}$. Then the forecast reconciliation constraint is

$$\text{minimize} \qquad J(y,\hat{y}): e_j = \left(y_j - \hat{y}_j\right)^2 \tag{4.6}$$

$$\text{with respect to} \qquad \hat{y} \in \mathcal{A}^{-1}(Y,\mathcal{T})$$

$$\text{subject to} \qquad Y_i = \sum_{j \in \mathcal{T}_i} \hat{y}_j \,, \hat{y}_j \geq 0. \tag{4.7}$$

The disaggregation problem estimates high frequency $\hat{y}$ to minimize the objective function $J(y,\hat{y})$ subject to the temporal aggregation constraints in (4.7) for all m intervals.

The Multi-source Iterative Shifting Disaggregation Algorithm presented in this Chapter mimics a constrained optimization problem by iteratively minimizing J over $\mathcal{A}^{-1}$, subject to the linear constraint (4.7). While the disaggregation operators described in Chapter 2's literature review maintain coherency between the sum of the high frequency estimates and the low frequency actual values, traditional operators do not recreate the variability within the high frequency estimates. This is especially true when working with datasets that contain measurements with significant temporal variability. Maintaining the coherence between temporal hierarchies is a way of ensuring the total of $\hat{y}$ agrees with Y and limits the solution space of $\mathcal{A}^{-1}(Y)$. However, this constraint can limit the magnitude of $\hat{y}$, even though accurately estimating high-variance intervals is critical in many applications.

4.3 Iterative Shifting Disaggregation Algorithm

This section presents the Iterative Shifting Disaggregation (ISD) algorithm. The ISD algorithm disaggregates multiple data series structured at nonuniform, low levels of

aggregation while maintaining the reconciliation constraint introduced in our constrained load shifting step (4.3.3). ISD iteratively generates consumption load profiles based on the relationship between observable low frequency series and high-frequency, independent correlated variables. The resulting high frequency series are a product of "shifting," or redistributing, the low frequency observations among the high frequency intervals. ISD is presented in three steps, summarized in Algorithm 1

ALGORITHM 1: ITERATIVE SHIFTING DISAGGREGATION

Inputs: **Multiple low frequency**, inconsistently spaced, time series for disaggregation
$\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_P} - P$ independent correlated variables measured at the target frequency of y
$\boldsymbol{nLRModels}$ –Hyperparameter (integer) used to control the number of linear regression models trained in the disaggregation process
$\boldsymbol{nDisaggCycles}$ – Hyperparameter (integer) used to control the number of times load is shifted among underlying intervals
$\boldsymbol{\alpha}$ – Parameter (float) used to control how error is redistributed in disaggregation process

Output: $\hat{\boldsymbol{y}}$ – An estimate of the high frequency, consistently spaced, disaggregated time series

```
Naively disaggregate multiple input series and aggregate into daily time series ŷ
Form design n × P matrix X = [x_1, x_2, ..., x_P] of transformed independent correlated variables
for (nLRModels) do
    Prediction Phase: Model daily ŷ using exogenous variables X, produce β̂_p
    Evaluate X · β̂_p to produce ỹ

    for (nDisaggCycles) do // Each iteration "shifts" interval load among underlying days
        for (mIntervals) do
            Update Phase: Remove kth interval from aggregate ŷ, calculate error ϵ = ỹ − ŷ
            Distribute error α · ϵ error into kth interval of ŷ// Constraints must be considered, original interval still removed
            Add kth interval back aggregate ŷ
        end // mIntervals
    end // nDisaggCycles
end // nLRModels
```

Figure 4.8 shows the individual steps of Algorithm 1. The iterative shifting process consists of multiple nested loops, each serving a specific function in refining the disaggregated time series.

- The outer loop iterates over multiple regression models trained at different stages of disaggregation, allowing progressive refinement.
- The middle loop performs several disaggregation cycles, incrementally re-distributing the error signal across high-frequency periods.
- The inner loop iterates over all input series (e.g., A, B, C, …), ensuring that each interval's contribution is updated independently before restoring the full series

The ISD Algorithm accepts a list of one or more low frequency time series measuring the same dependent variable. A design matrix X is formed from the high frequency time series. The term 'design matrix' follows standard usage in statistical learning and regression modeling. It refers to a structured representation of explanatory variables, where each row corresponds to an observation (e.g., a day in the disaggregated series), and each column represents a predictor variable. By organizing the data in this matrix form, the regression framework efficiently estimates coefficients and generates high-frequency predictions. The P high frequency time series x_p are the exogenous, correlated time series.

ISD depends on the correlations between the exogenous variables and the dependent high frequency time series. Control parameters $nLRModels$, $nDisaggCycles,$ and α are user-defined inputs used in the prediction and update phases. $nLRModels$ specifies the number of gas consumption profiles generated in the prediction phase, while $nDissagCycles$ controls how many permutations occur over the overlapping intervals of the multiple low frequency time series per $LRModel$. Parameter α is used to determine

how much weight should be given to the current-state estimate when the shifting occurs in the update phase. Lastly, control loop parameter $mIntervals$ depends on the number of intervals in the multiple low frequency time series and is calculated in Step 1. These parameters were selected based on empirical validation and best practices in time series disaggregation. Their specific values are tuned for each dataset based on historical error analysis and model performance assessments.

For the remainder of the ISD description, and without loss of generality, we assume the desired high frequency is daily, the low frequency consists of nonuniformly sampled intervals of multiple days, and there are three low frequency time series: A, B, and C. For example, suppose energy is supplied to a set of consumers through three sources. Which sources supply energy on a given day is unknown, but the total energy supplied over intervals of days is measured. These intervals are not synchronized across the sources.

4.3.1 Step 1 - Preparation of Time Series Inputs using Low-frequency Sampling

We extend the notation introduced in Section 4.2 to accommodate the multiple time series. For illustration, we will use time series A, B, and C. The objective is to estimate coherent $\hat{y}$, which is the daily sum of the multiple time series. Let A be a time series with index i and N_A observations

$$A = \{ A_i, i = 1, \dots, N_A\}. \tag{4.8}$$

The time steps of A may span long, nonuniform durations, for example, anywhere between 4 – 40 days. Let $\mathcal{T}_i^A$, also indexed by i, represent the number of days each A_i observation spans (similar to $\mathcal{T}_i$ in Figure 4.1, however now a single integer of days). Let d be a daily index used to denote the days on which each observation A_i occurs. Index d later serves to synchronize multiple time series to a common time step. Figure 4.2 illustrates A_i, d, and $\mathcal{T}_i^A$.

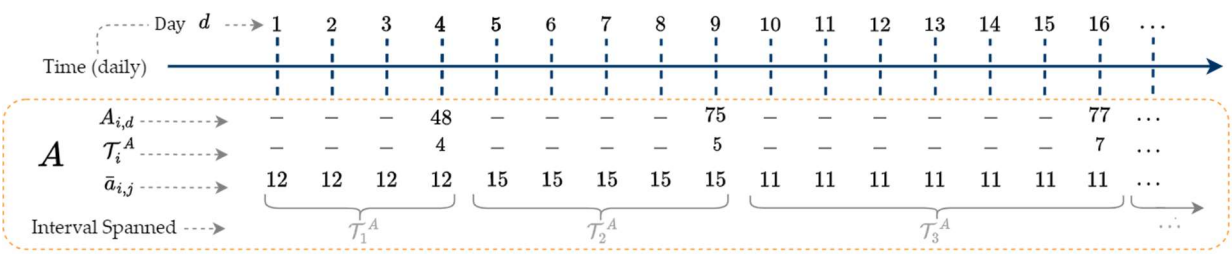

Figure 4.2 – Time series A_i, interval durations $\mathcal{T}_i^A$, and daily index d.

Figure 4.2 shows the first observation A_1= 48 occurs on the fourth day (d = 4) and is measured over a four-day interval ($\mathcal{T}_1^A$ = 4). Thus, A_1 is defined on the time interval [1,4] as illustrated by the gray bracket. The general time interval for each A_i is $[d - \mathcal{T}_i^A + 1, d]$, where d is the day number over which A_i is aggregated. The second gas consumption observation A_2= 75 spans five days ($\mathcal{T}_2^A = 5$) and is on the time interval [5,9].

To enhance reproducibility, this section provides details on how input time series are structured and preprocessed before applying ISD. Each time series consists of irregularly spaced interval measurements, which are first aligned to a common time domain using the naive disaggregation method. The transformation ensures that each interval's total consumption remains preserved. The first step of the ISD algorithm naively disaggregates

A into a daily series [43]. Naïve disaggregation creates a daily representation of an interval series, while maintaining the coherence constraints that the underlying days of the high frequency disaggregated series sum to the interval observation from which they originate. This is achieved by dividing each A_i by $\mathcal{T}_i^A$ to get an average value for all days within the ith interval. Missing data in low-frequency time series, such as a missing monthly bill, are estimated using a low-frequency regression model based on similar intervals or correlated variables to avoid bias introduced by assuming zero values. Once estimated, the interval is divided evenly across its days, maintaining alignment during iterative updates. For example, if a missing monthly value is estimated as 100 and spans 5 days, 20 units are allocated per day. We augment A_i with the day d on which it occurred, yielding $A_{i,d}$. The naively disaggregated series, $\bar{a}$ indexed by i, j is

$$\bar{a}_{i,j} = \frac{A_{i,d}}{\mathcal{T}_i^A}, \text{where } j = d - \mathcal{T}_i^A + 1, \dots, d \text{ and } i = 1, \dots N_A. \tag{4.9}$$

Figure 4.3 shows an example series $\bar{a}$ created from the A shown in Figure 4.2.

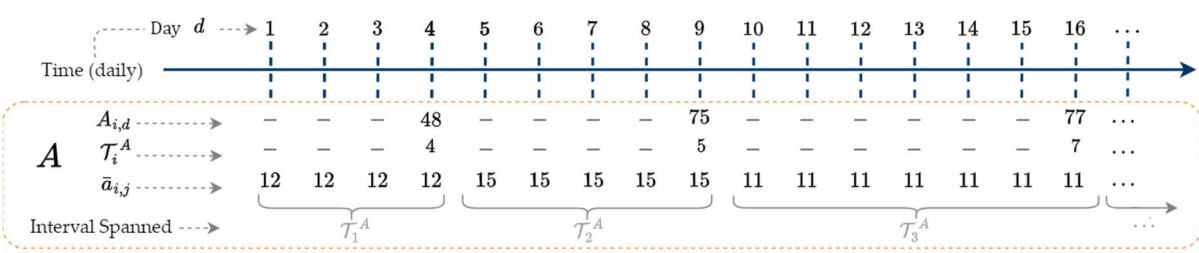

Figure 4.3 – Time Series A naively disaggregated into daily series $\bar{a}$.

Index j in Figure 4.3 is used to distribute each observation $A_{i,d}$ equally among the $\mathcal{T}_i^A$ days spanned by interval i. The index i is used to keep track of the interval of $\bar{A}_{i,d}$. For example, the first interval observation $A_{1,4}$ creates disaggregated $\bar{a}_{1,j} = \frac{48}{4} = 12$, for $j \in$

$[1, 2, 3, 4]$. The next interval observation $A_{2,9}$ creates disaggregated $\bar{a}_{2,j} = \frac{75}{5} = 15$, for $j \in [5, 6, 7, 8, 9]$.

We now consider the case when multiple low frequency time series are used as inputs to ISD. Given an additional time series B, indexed by k with N_B observations, and C, indexed by l with N_C observations, then

$$B = \{ B_k, k = 1, \dots, N_B \}, \tag{4.10}$$

$$C = \{ C_l, l = 1, \dots, N_C \}. \tag{4.11}$$

Repeating the process of creating $\bar{a}$, we create $\bar{b}$ and $\bar{c}$.

A, B, and C are defined on the same time domain and may have overlapping intervals. Figure 4.4 illustrates the overlapping intervals of each series and presents example series $\bar{b}$ and $\bar{c}$. In practice, the interval lengths of B and C span longer durations than pictured.

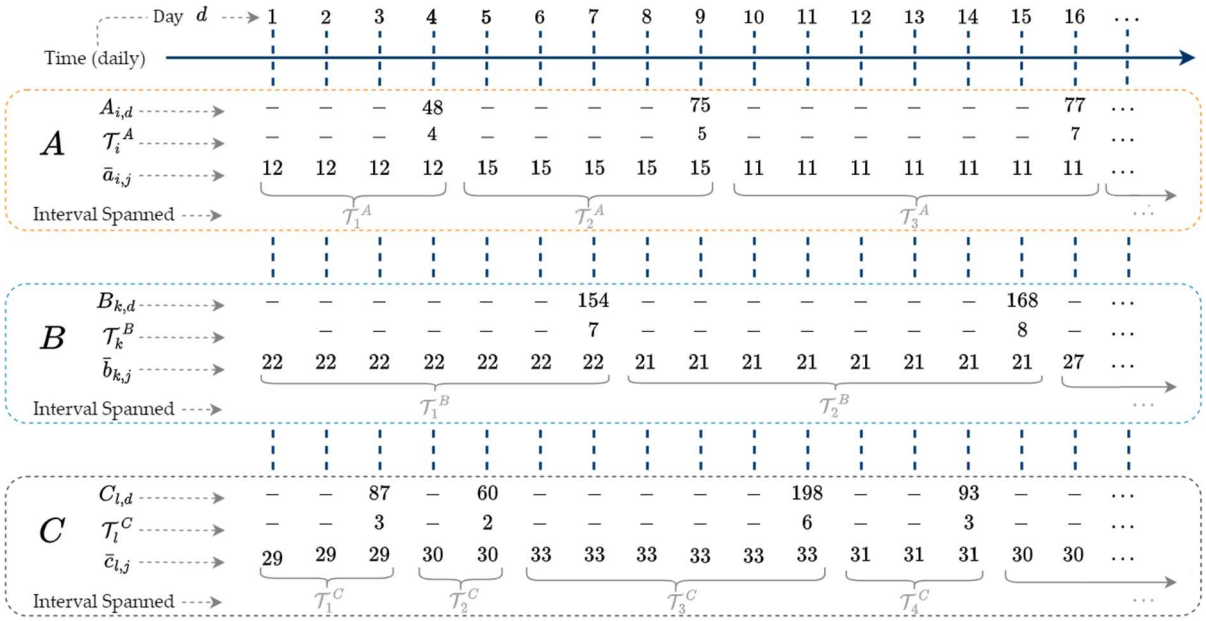

Figure 4.4 – Three input time series A, B, and C with overlapping time intervals.

The example series in Figure 4.4 have nonuniform time steps and overlapping intervals. We replace all missing/unaccounted for observations throughout $\bar{a}$, $\bar{b}$, and $\bar{c}$ with a value of zero, forcing each series to have an equal number of days. We assume each time series starts and ends on the same day.

We aggregate each day's average within $\bar{a}$, $\bar{b}$, and $\bar{c}$ to create an initial estimate of the daily time series $\hat{y}$ indexed by days d. Aggregate $\hat{y}$ as

$$\hat{y} = \bar{a} + \bar{b} + \bar{c}. \tag{4.12}$$

The aggregation in (4.12) is shown in Figure 4.5.

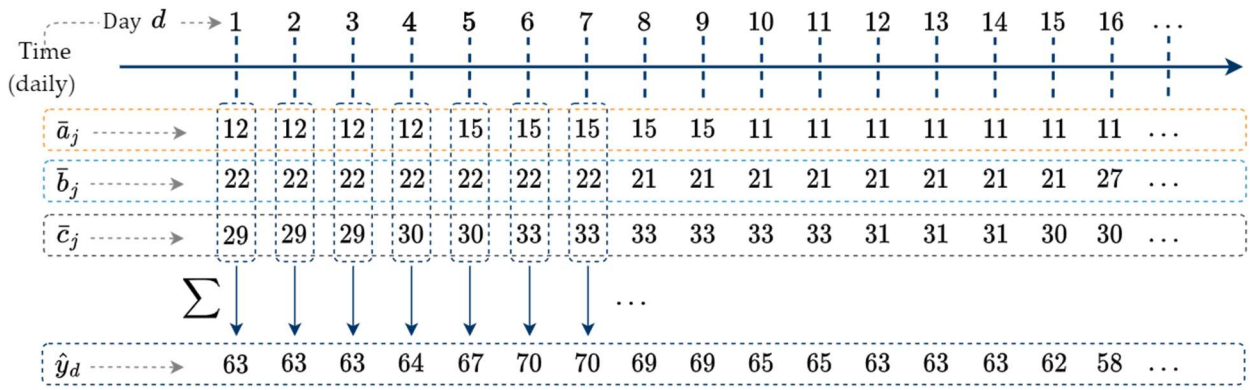

Figure 4.5 – Each day d is aggregated into the $\hat{y}_d$ daily time series.

Figure 4.5 shows how each day d is aggregated into the $\hat{y}_d$ daily time series. For example, on day one ($d = 1$), $\bar{a}_1 = 12$, $\bar{b}_1 = 22$, and $\bar{c}_1 = 29$ are aggregated to form $\hat{y}_1 = 63$. On $d = 7, \bar{a}_7 = 15$, $\bar{b}_7 = 22$, and $\bar{c}_7 = 33$ are aggregated to form $\hat{y}_7$ = 70. Naïve disaggregation maintains coherence in the construction of $\hat{y}$, meaning that the coherency constraint of $\bar{a}$, $\bar{b}$, and $\bar{c}$ is met in $\hat{y}$. We now consider the independent correlated variables used to generate gas consumption profile estimates.

4.3.2 Step 2 - Iteratively Modeling Daily Time Series (High-frequency Sampling)

This section introduces common exogenous variables used to model natural gas demand. The exogenous variables are sampled daily and on the same time domain as $\hat{y}$. The performance of the gas demand forecasting models depends on a high statistical correlation between these observed exogenous variables and dependent consumption variables. Natural gas consumption is affected by seasonal variability [20]. Weather variables, specifically temperature, wind, and humidity, have proven to be the good indicators of future natural gas demand [2], [21], [22]. Variable transformation techniques are applied to

account for the nonlinearity of these variable's time series trends. The exogenous variables are:

- Heating Degree Days (*HDD*): A temperature-based metric that quantifies heating demand by measuring how much daily temperatures fall below a base threshold.
- Cooling Degree Days (*CDD*): A complementary metric that captures cooling demand based on how much daily temperature exceeds a base threshold.
- Wind Adjusted Heating Degree Days (*HDDW*): An adjusted metric that accounts for wind speed effects on perceived temperature, improving heating load estimates.

The selection of exogenous variables is guided by two key principles: relevance to the underlying consumption patterns and statistical robustness. Variables such as Heating Degree Days (*HDD*) and Cooling Degree Days (*CDD*) are chosen based on their established relationships with gas consumption in temperature-sensitive regions. Correlation analysis is conducted to confirm strong linear relationships between candidate variables and the target time series. Additionally, cross-validation techniques are applied to ensure that the inclusion of these variables improves predictive accuracy without introducing spurious correlations. For non-temperature-sensitive datasets, alternative variables such as socioeconomic indicators, day-of-week effects, or calendar-based seasonality may be employed, guided by domain-specific insights.

Given daily temperature series t and wind speed series w, both indexed by d with the same number of days (N_D) as $\hat{y}$

$$t = \{ t_d, d = 1, \dots, N_D \}, \tag{4.13}$$

$$w = \{ w_d, d = 1, \dots, N_D \}, \tag{4.14}$$

a nonlinear transformation is applied to these variables to obtain Heating Degree Days (HDD), Wind Adjusted Heating Degree Days (HDDW), and Cooling Degree Days

(CDD). To transform these weather time series, a reference temperature T_{ref} typically between 55- and 65-degrees Fahrenheit is used such that

$$HDD_d = \max\,(T_{ref} - t_d, 0), \tag{4.15}$$

$$HDDW_d = \begin{cases} HDD_{ref}\dfrac{152\ +\ w_d}{160}\ w_d \leq 8 \\ HDD_{ref}\dfrac{72\ +\ w_d}{80}\ w_d > 8 \end{cases} \tag{4.16}$$

$$CDD_d = \max\big(0, t_d - T_{ref}\big), \tag{4.17}$$

for all N_D days. This work uses a T_{ref} = 65-degrees Fahrenheit. Transpose each resulting series from (4.15), (4.16), and (4.17) using T_{ref} = 65 to column vectors and form design matrix X

$$X = \begin{bmatrix} 1 & HDD_1 & HDDW_1 & CDD_1 \\ 1 & \vdots & \vdots & \vdots \\ \vdots & HDD_{N_D} & HDDW_{N_D} & CDD_{N_D} \end{bmatrix}. \tag{4.18}$$

Design matrix X is an N_d by $P + 1$ matrix, where N_d is the number of days in the daily aggregate series $\hat{Y}$, and P is the number of independent correlated daily variables. We include the bias vector of 1's in the first column X to capture the base load factor, or the natural gas load that is not dependent on temperature [24]. Design matrix X holds all exogenous variables used throughout the remainder of this work.

The naïve disaggregation described in Step 1 fails to recreate the variability seen in a actual daily series. The remaining steps of the ISD algorithm iteratively adjust $\hat{y}$ to capture this variability. Shifting occurs in two phases: a prediction phase generates the *a priori* state estimate $\tilde{y}$, and an *a posteriori* update phase shifts $\hat{y}$ based on the *a priori* state estimate. The prediction phase is the first iterative step of the ISD algorithm seen in line 2 in Algorithm 1. Let β be a vector of regression coefficients. In each iteration, a

multiple linear regression model is fit using daily gas consumption estimates $\hat{y}$ and design matrix X to estimate $\hat{\beta}$. Estimate β using least squares, minimizing the mean square of residuals of all N_D days

$$\min_{\beta} \|\hat{y} - X\beta\|. \tag{4.19}$$

Use the estimated coefficients $\hat{\beta}$ to calculate daily load profile

$$\tilde{y} = X\beta. \tag{4.20}$$

Series $\tilde{y}$ is this iteration's *a priori* state estimate. It does not meet the coherency constraint [72].

Design matrix X remains constant across all iterations. Once the constrained shifting in $\hat{y}$ is carried out in Step 3, a new version of $\hat{y}$ is modeled with X in (4.20) to summarize the updated relationship between the daily gas consumption estimates and correlated variables in future iterations.

4.3.3 Step 3 - Constrained Load Shifting

Step 3 iteratively adjusts daily series $\hat{y}$ towards an estimated daily series $\tilde{y}$. The update in $\hat{y}$ maintains the temporal disaggregation reconciliation constraint that the disaggregated days $\hat{y}_d$ sum to the observed interval observation. In the case of multi-source data, ISD applies shifting corrections independently to each source before merging their refined estimates into the final disaggregated time series. This ensures that discrepancies in one source do not disproportionately affect another, maintaining the integrity of distinct load profiles while enforcing coherence in the overall disaggregation process. A, B, and C are the original interval observations. Recall that $\bar{a}$, $\bar{b}$, and $\bar{c}$ are initially naively

disaggregated daily counterparts, respectively. In each iteration of this adjustment, the daily values are "shifted" within each interval to reflect the variability existing in correlated variables X.

Daily time series $\hat{y}$ and $\tilde{y}$ are plotted in Figure 4.6 along with the values of naively disaggregated $\bar{a}$, $\bar{b}$ and $\bar{c}$ shown beneath to show all available daily information used in the ISD constrained shifting step.

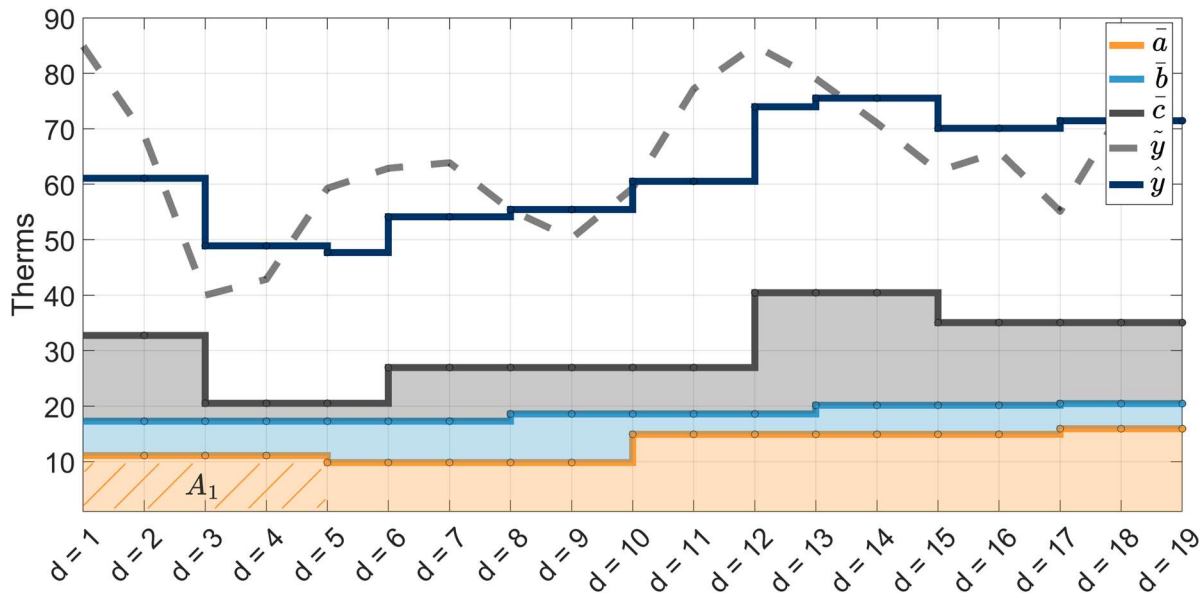

Figure 4.6 – Plot of $\hat{y}$, formed from the daily aggregation of $\bar{a}$, $\bar{b}$, and $\bar{c}$, compared to Step 3's estimated load profile $\tilde{y}$ before any shifting occurs.

The constrained shifting step of ISD iterates over all intervals of original input series A, B, and C. The following example of ISD's constrained shifting is carried out using the first interval $A_1 = 48$, which spans four days over the time interval $[1, 2, 3, 4]$. This information is directly reflected in the same four days of $\bar{a}$, where the sum of $\sum_{d=1}^{4} \bar{a}_d$ is constrained to equal $48 = A_1$ (Figure 4.6– gold hash marks). The value of A_1 is a reconciliation constraint. The shifted values are then added into $\hat{y}$ at the end of each iteration.

Begin the shifting by removing the contribution of A_1 from $\hat{y}$ over the d days A_1 spans

$$\hat{y}_d = \hat{y}_d - \bar{a}_d \text{ for } d = 1, \dots, 4. \tag{4.21}$$

The remaining values in $\hat{y}_d$ for days $d \in [1, \dots, 4]$ represent the overlapping interval contributions of $\bar{b}$ and $\bar{c}$. Calculate the error ϵ_d between $\tilde{y}$ and $\hat{y}$ over the same days

$$\epsilon_d = \tilde{y}_d - \hat{y}_d \text{ for } d = 1, \dots, 4. \tag{4.22}$$

Error ϵ_d is the difference between profiles of $\tilde{y}$ and $\hat{y}$ with the contribution of A_1 removed. It can be thought of as the estimated variability missing from naively produced interval $\bar{a}_d$. We constrain ϵ_d to be a nonnegative value by replacing negative errors with zero. The profile of $\bar{a}_d$ with error ϵ_d incorporates variability into new interval

$$z_d = \frac{\epsilon_d * \sum_{j=1}^{4} \bar{a}_j}{\sum_{j=1}^{4} \epsilon_j}, \text{for } d = 1, \dots, 4. \tag{4.23}$$

Temporary interval z can be thought of as A_1 shifted by the day-to-day variability extracted from the system as a whole. Some of the variability captured in temporary interval z may originate from components independent of A (e.g., overlapping intervals of B and C).

The exact day-to-day variability of $\bar{a}$ in relation to $\bar{b}$ and $\bar{c}$ is unknown. Therefore, ISD updates $\bar{a}$ in small increments. Update $\bar{a}$ over the days spanned in A_1 as a weighted combination of $\bar{a}$'s original state and z. Let α be the weight used in this combination

$$\bar{a}_d = (1 - \alpha)(\bar{a}_d) + (\alpha)(z_d) \text{ for } d = 1, \dots, 4. \tag{4.24}$$

The newly updated interval within $\bar{a}_d$ is plotted in Figure 4.7 as $\bar{a}_{d|shifted}$. Also plotted are the values of $\bar{a}_d$ before any shifting occurs ($\bar{a}_{d|unshifte}$), error ϵ_d (4.22), and interval estimate z_d (4.23).

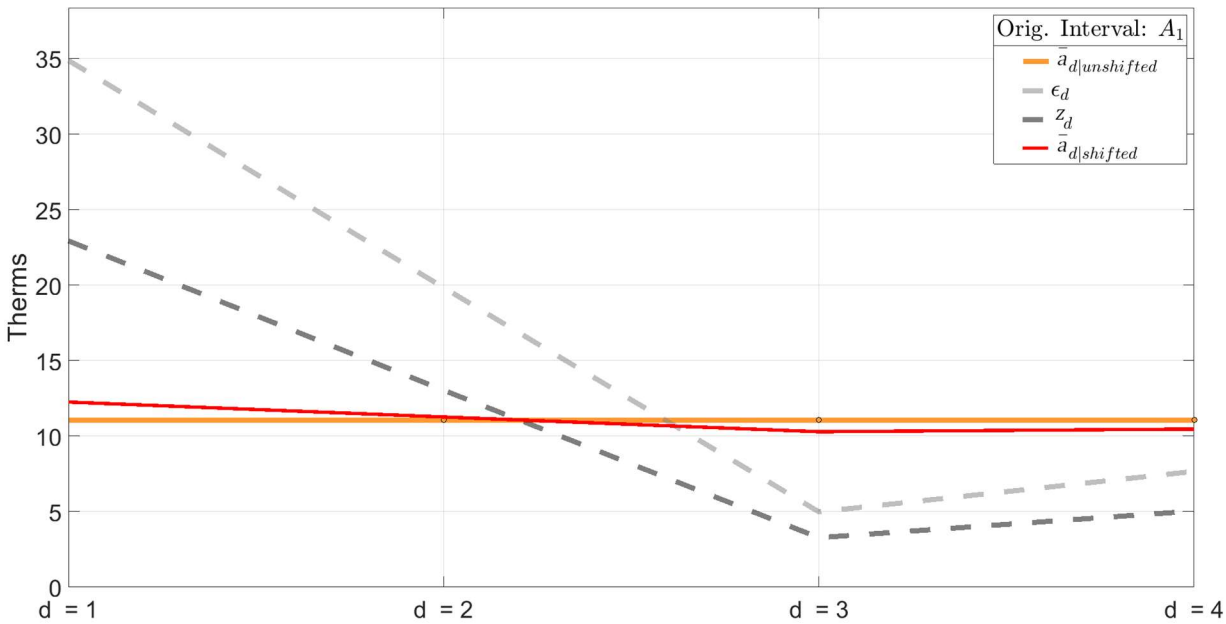

Figure 4.7 – Newly updated interval $\bar{a}_{d|shifted}$, original $\bar{a}_{d|unshifted}$, constrained interval estimates z_d and error ϵ_d.

Step 3 carries out the shifting of a single interval example A_1 for a single disaggregation cycle. After all $nDisaggCycles$ disaggregation cycles complete, a new relationship is modeled in the outermost for-loop of Figure 4.8

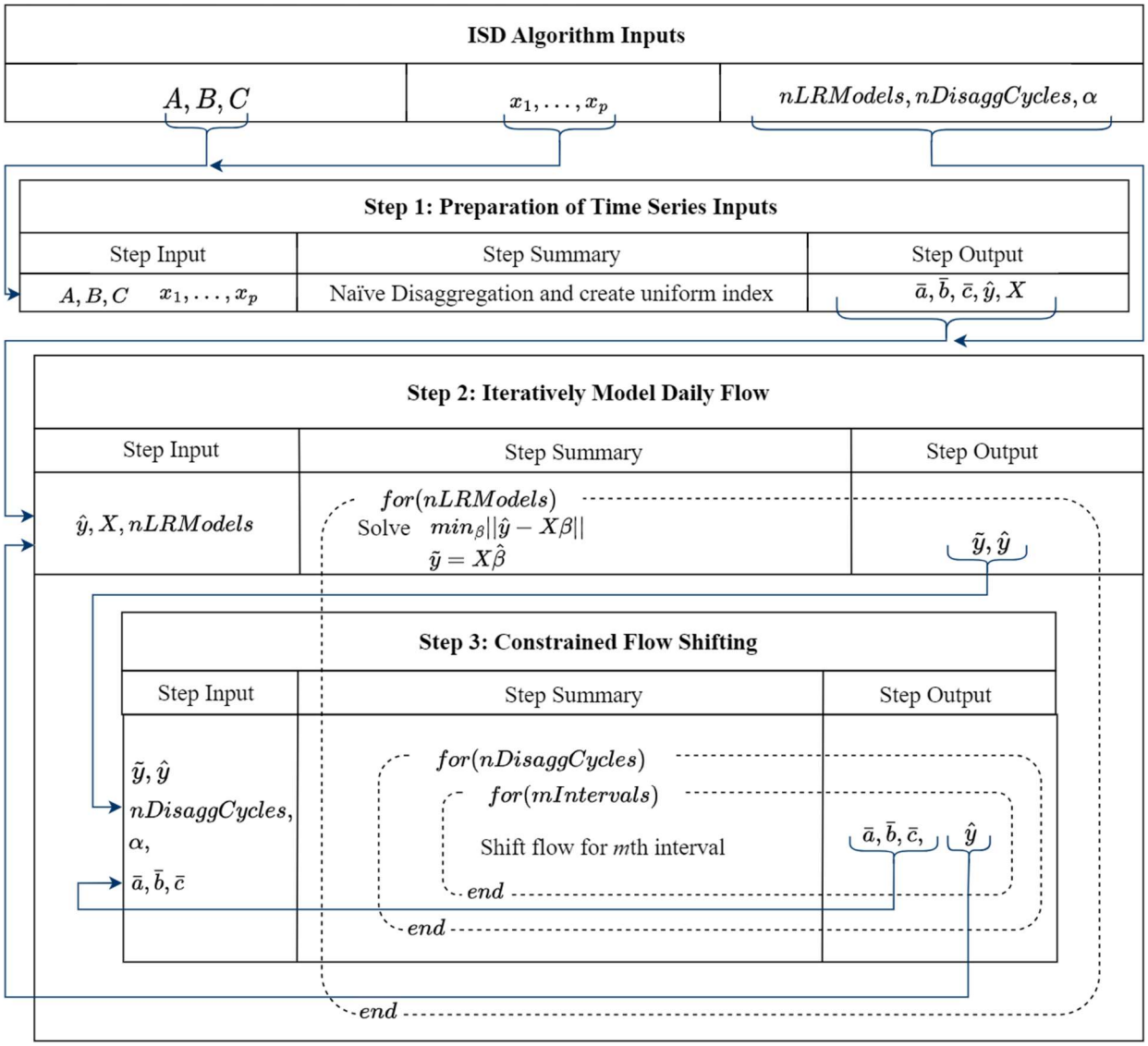

Figure 4.8 – Three gas consumption load shifting steps carried out by the ISD algorithm.

The control-loop and gain variables $nDisaggCycles$, $nLRModels$, and α in Figure 4.8 are empirically determined and based on the number of interval series input to the ISD algorithm. If there are three or more input time series, $nDisaggCycles = nLRModels = 10$, and $\alpha = 0.05$. If there are fewer than three inputs, more iterations are required. These variables may be adjusted for different applications.

Figure 4.9 shows how example interval A_1 shifts over each regression evaluation as the ISD algorithm determines the proportion of gas consumption to assign in each underlying day.

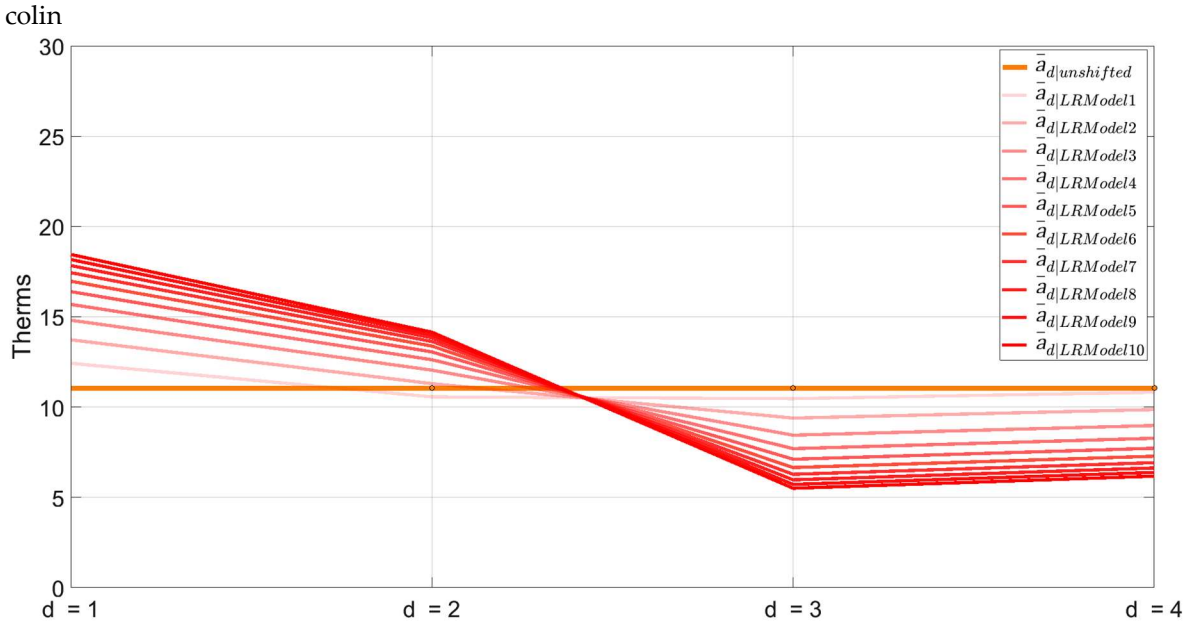

Figure 4.9 – Evaluation of A_1 across all $nLRModels$. Daily series $\bar{a}_{d|unshifted}$.

Daily series $\bar{a}_{d|unshifted}$ is the naively disaggregated version of A_1 before any shifting occurs. $\bar{a}_{d|LRModel1}$ is the first permutation made and the same as $\bar{a}_{d|shifted}$ in Figure 4.7. This shifting process is repeated for all remaining intervals in A, B, and C, then repeated for a total of $nLRModel$ iterations.

4.4 Disaggregation Case Studies

Four temporal disaggregation methods are compared in the following case studies: Iterative Shifting Disaggregation algorithm (ISD - 4.2), Naïve (D-NAV), Time Series Reconstruction (TSR– [73]) and Generalized Least Squares (GLS – [5]). We present an analysis comparing each disaggregation technique with the others in the following section.

Two case studies are examined. The first study, which we refer to as the Operating Area case study, focuses on disaggregating low frequency operating area billing cycle

time series from an approximately monthly cycle frequency to a daily frequency. The Operating Area dataset consists of aggregated gas flow measurements from multiple distribution zones within a utility's network. These data are sampled at irregular intervals based on billing cycles and operational reporting schedules. The dataset structure aligns with typical industry datasets used for forecasting and load analysis.

The second case study, referred to as Residential 100, focuses on disaggregating 100 residential customers series to daily time series. The Residential 100 dataset used in this study consists of anonymized, aggregated natural gas consumption data collected from 100 residential customers. The Residential 100 dataset reflects industry-standard customer billing structures and is representative of natural gas distribution data encountered by local distribution companies (LDCs).

We use three low frequency time series (A, B, C) with overlapping and varying interval lengths as inputs to each disaggregation algorithm. This study follows the description given in Section 3 – Iterative Shifting Disaggregation Algorithm. The observed values of Operating Area times series A, B, and C from 2021 – 2024 are shown in the top plot of Figure 4.10. These series exhibit seasonal patterns, peaking during winter, indicating higher gas demand in colder months. The naively disaggregated time series $\bar{a}$, $\bar{b}$, $\bar{c}$ and their aggregate $\hat{y}$ (result of Step 1) are shown in the middle plot. These disaggregated series also show seasonal fluctuations, mirroring the patterns seen in the observed data. The bottom plot shows daily temperature ($F°$), revealing an annual cycle of higher summer temperatures and lower winter temperatures. The inverse relationship between temperature and gas demand is evident: gas consumption increases in winter as temperatures drop and decreases in summer when temperatures rise.

Figure 4.10 shows gas consumption from the three independently metered areas displaying periodic and cyclical characteristics. In the Residential 100 case study, we add a layer of granularity to the disaggregation problem by focusing on individual residential gas consumers. Similar patterns are observed in this dataset; however, instead of analyzing gas consumption across different geographical levels (e.g., neighborhoods like A, B, and C), we disaggregate the data for 100 individual consumers in a less temperature-sensitive area. The first ten customer gas consumption time series are plotted in Figure 4.11. The Residential 100 data set consists of overlapping and incoherent customer time series.

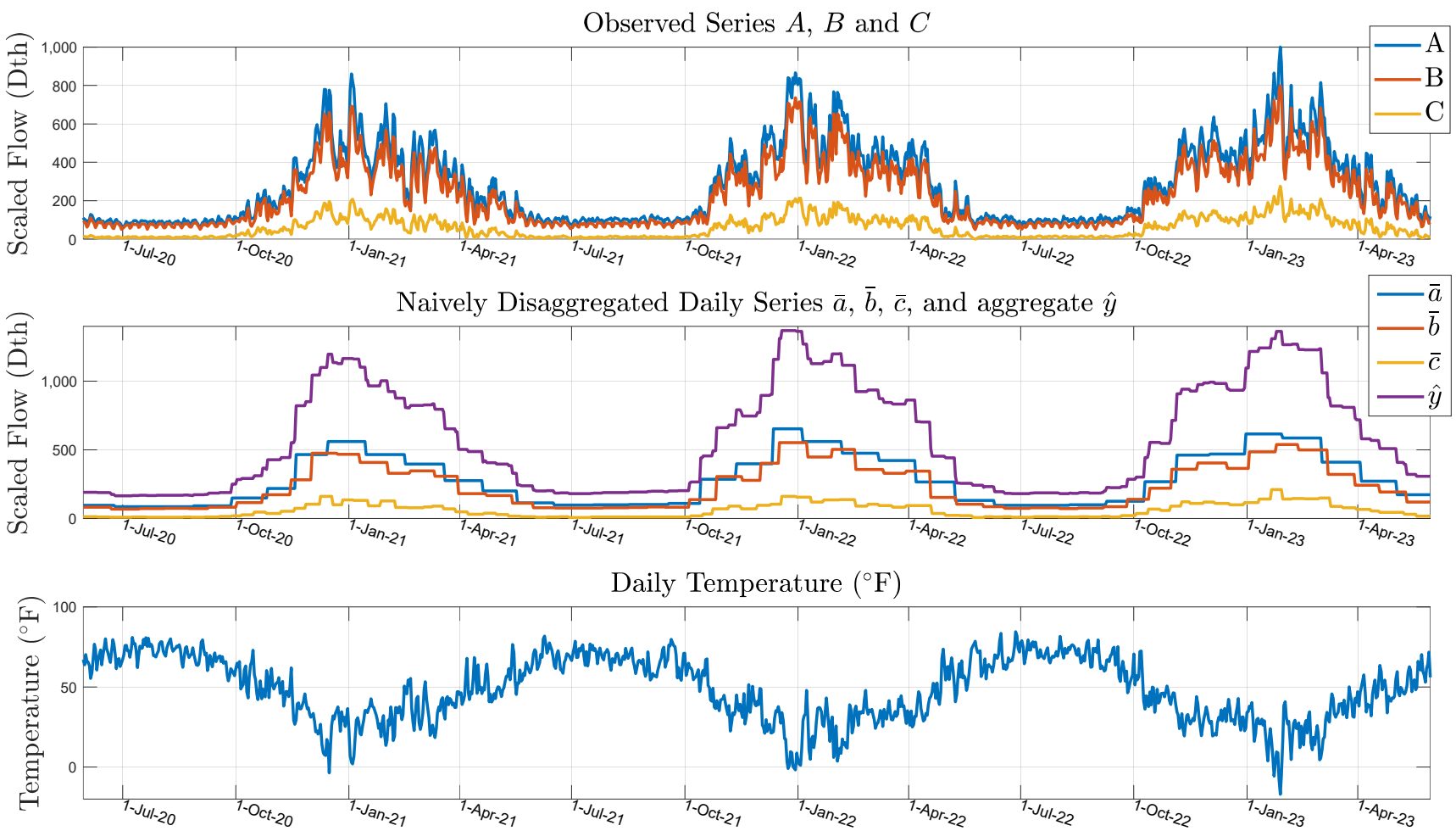

Figure 4.10 – Plot of input time series A, B, and C over the 2021-2023 heating seasons.

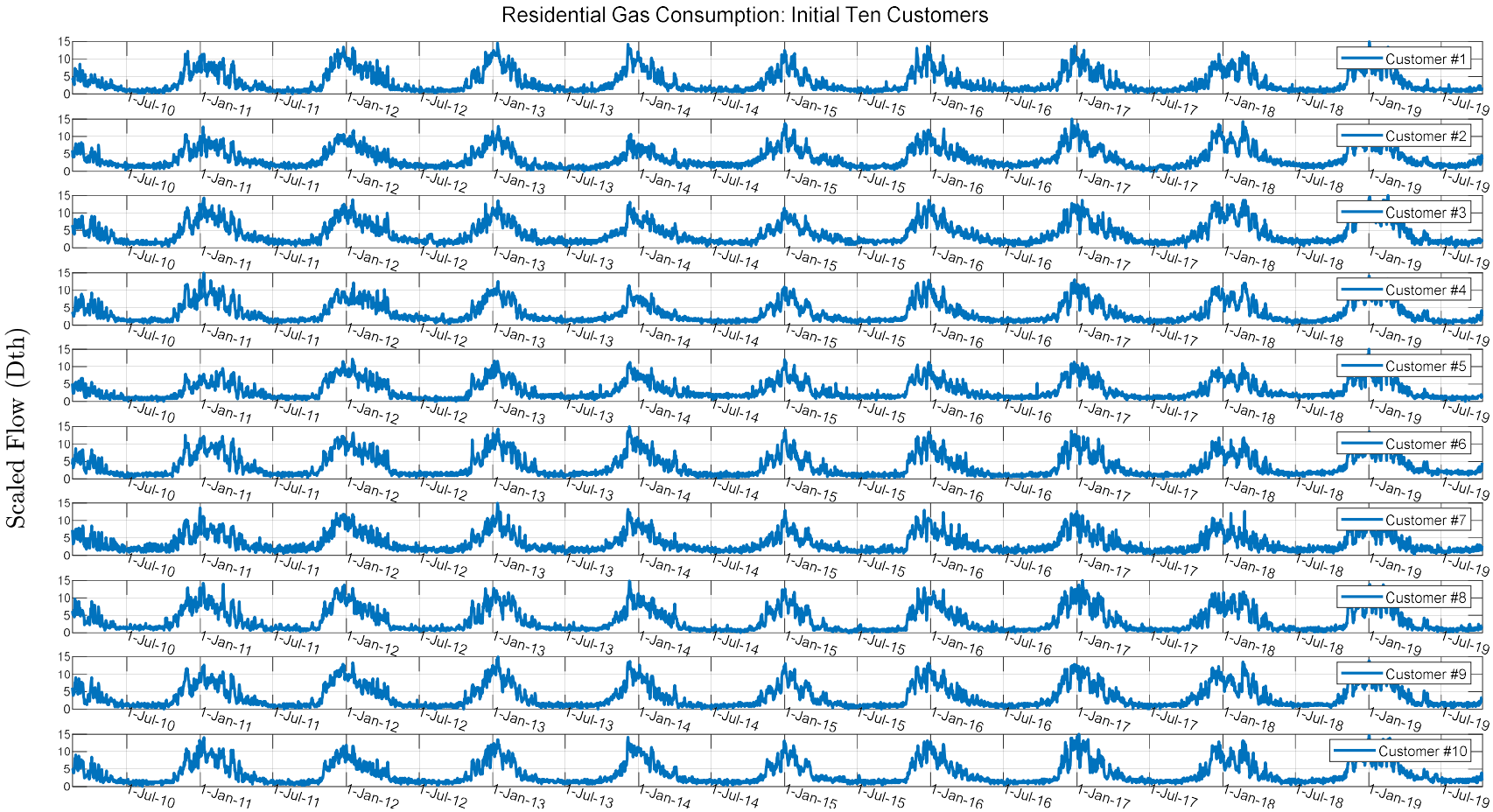

Figure 4.11 – Plot of the first ten of 100 customers consumption time series from the Residential 100 dataset.

We use a regression model with variables commonly used in gas demand forecasting, calculated similarly to those in [2], [21], [22] for incorporating exogenous variables. Based on these works, domain knowledge, and best practices from utilities' forecasting in Vitullo et al. [6], eight correlated variables are used in model training: a linear trend, heating degree days $HDD65$ and $HDD55$, wind-adjusted heating degree day $HDDW65$, cooling degree day $CDD65$, and four seasonal variables representing the first two harmonics of a Fourier series. Daily temperature and wind are transformed into HDD^* and CDD variables using (4.15), (4.16), and (4.17). The four seasonal variables are yearly periodic functions representing the first two harmonics of a Fourier series [61]. We format these exogenous variables into design matrix X as described in Section 4.3.1. The ISD algorithm can accept any number of exogenous variables, with the first iteration initializing these variables into design matrix X.

Three measures are used to summarize the performance of each algorithm: root-mean-square error ($RMSE$), mean absolute percentage error ($MAPE$), and weighed mean absolute percentage error ($WMAPE$). $RMSE$ is a measure of the average residual. Let y_d be the value of signal y at time d, $\hat{y}_d$ be the disaggregated quantity on day d, and N_d be the total number of underlying days, then

$$RMSE = \sqrt{\frac{\sum_{d=1}^{N_D}(\hat{y}_d - y_d)^2}{N_D}}. \tag{4.25}$$

$RMSE$ is measured in dekatherms. $MAPE$ is the average absolute error. It is unitless and represents the percent of error.

$$MAPE = \frac{1}{N_D}\sum_{d=1}^{N_D}\left|\frac{y_d - \hat{y}_d}{y_d}\right| \times 100. \quad 4.26$$

Lastly, weighted means absolute error ($WMAPE$) is commonly used to evaluate gas forecasting performance and weights the heating season more than the summers.

$$WMAPE = \frac{\sum_{d=1}^{N_D}|y_d - \hat{y}_d|}{\sum_{d=1}^{N_D}|y_d|} \times 100. \quad 4.27$$

The NAV, TSR, and GLS disaggregation methods do not share the ISD algorithm's advantage of being able to simultaneously disaggregate multiple time series with overlapping intervals. Therefore, the low frequency operating-area time series A, B, and C as well as the 100 residential customer billing cycle time series are disaggregated individually and aggregated post-process for all days d.

Table 2 presents the $RMSE$, $MAPE$, and $WMAPE$ for each disaggregation algorithm applied to the Operating Area dataset.

Table 2. RMSE, MAPE, and WMAPE, and coherency indicator for NAV, TSR, GLS, and ISD disaggregation algorithms applied to the Operating Area dataset.

Method	Coherent	RMSE [Dth]	MAPE [%]	WMAPE [%]
NAV	Yes	153.9	18.8	16.9
TSR	No	71.5	11.4	8.6
GLS	No	101.24	16.9	12.9
ISD	**Yes**	**53.3**	**9.7**	**7.2**

Table 2 shows that the ISD algorithm has the lowest $MAPE$, $RMSE$, and $WMAPE$, making it the most accurate disaggregation method when applied to the low frequency Operating Area dataset. As expected, the NAV benchmark model is the least

accurate. The TSR algorithm is the second most accurate, outperforming GLS, but still has higher error metrics than ISD. Using the Diebold-Mariano test at the 0.05 significance level, the improvements in predictive accuracy by ISD over the other methods are statistically significant [74].We examine the estimated daily series and residual errors for each disaggregation algorithm during the 2021–2023 heating seasons in Figure 4.12.

Figure 4.12 compares the different disaggregation methods ($\hat{y}_d$- orange) with the actual observed values (y_d- blue). The residual difference between these series is plotted beneath (yellow) for each day d. The plot highlights how each method captures seasonal variability, showing higher values during peak periods and lower values during off-peak periods. Differences between the forecast methods and the actual values are more noticeable during periods of high variability, indicating differences in accuracy. The top NAV plot of Figure 4.12 illustrates this as the naïve disaggregation smooths out consumption across each heating season. Accurate forecasts are especially critical during high variability periods, such as winter [61]. The GLS and ISD plots of Figure 4.12 show better performance over these heating seasons, while the TSR algorithm shows the smallest residuals. This is examined further in the Discussion Section.

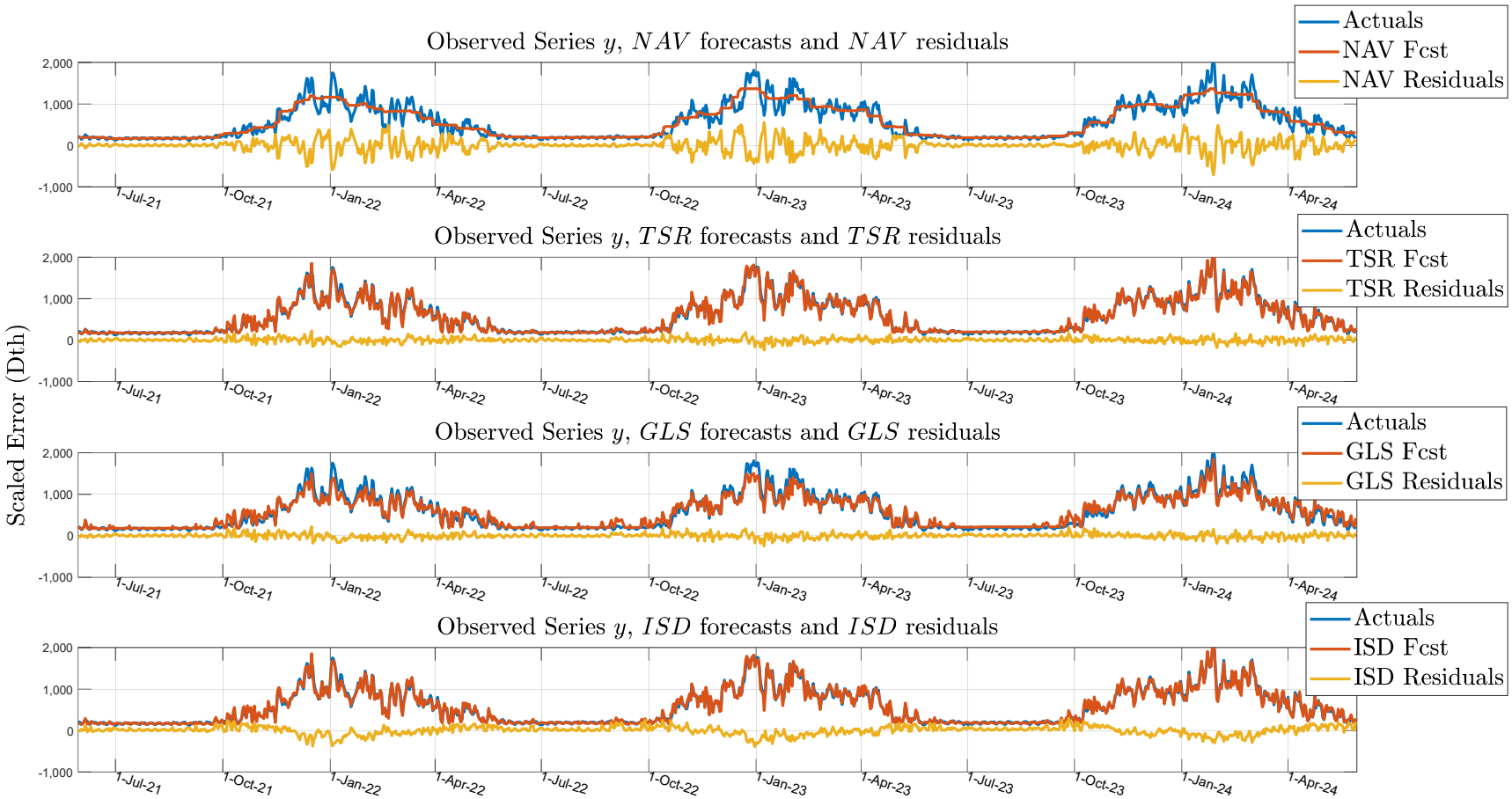

Figure 4.12 – Residual comparison of actuals vs forecasts for the NAV, TSR, GLS, and ISD disaggregation algorithm (Operating Area).

The $RMSE$, $MAPE$, and $WMAPE$ for each algorithm for the Residential 100 case study are presented in Table 3.

Table 3. RMSE, MAPE, WMAPE, and coherency indicator for NAV, TSR, GLS, and ISD disaggregation algorithms applied to the Residential 100 dataset.

Method	Coherent	RMSE [Dth]	MAPE [%]	WMAPE [%]
NAV	Yes	20.32	37.14	33.82
TSR	No	11.62	18.45	13.73
GLS	No	15.89	23.96	19.43
ISD	**Yes**	**6.61**	**11.13**	**9.07**

Table 3 shows that the ISD method outperforms the other disaggregation algorithms across all error metrics, achieving the lowest $RMSE$ (6.61 Dth), $MAPE$ (11.13%), and $WMAPE$ (9.07%). This suggests that ISD's coherent structure likely enhances its overall accuracy, especially in relative error terms. Despite its coherent structure, NAV shows the highest $RMSE$ (20.32 Dth), indicating that coherence alone does not ensure minimal error in absolute terms. The high error values for NAV, with $MAPE$ at 37.14% and $WMAPE$ at 33.82%, are reasonable for a benchmarking naive baseline, which serves as a reference point rather than an optimized solution. By contrast, TSR and GLS, though non-coherent, demonstrate moderate error rates, with TSR achieving lower $WMAPE$ (13.3%) than GLS (19.43%), suggesting that TSR may better manage specific types of errors. Overall, these results underscore ISD as the most reliable disaggregation method.

The data used in Residential 100 is considerably less temperature sensitive than the A, B, and C Operating Area dataset. We examine residential gas forecasting residual distribution for each method on the Residential 100 data in Figure 4.13.

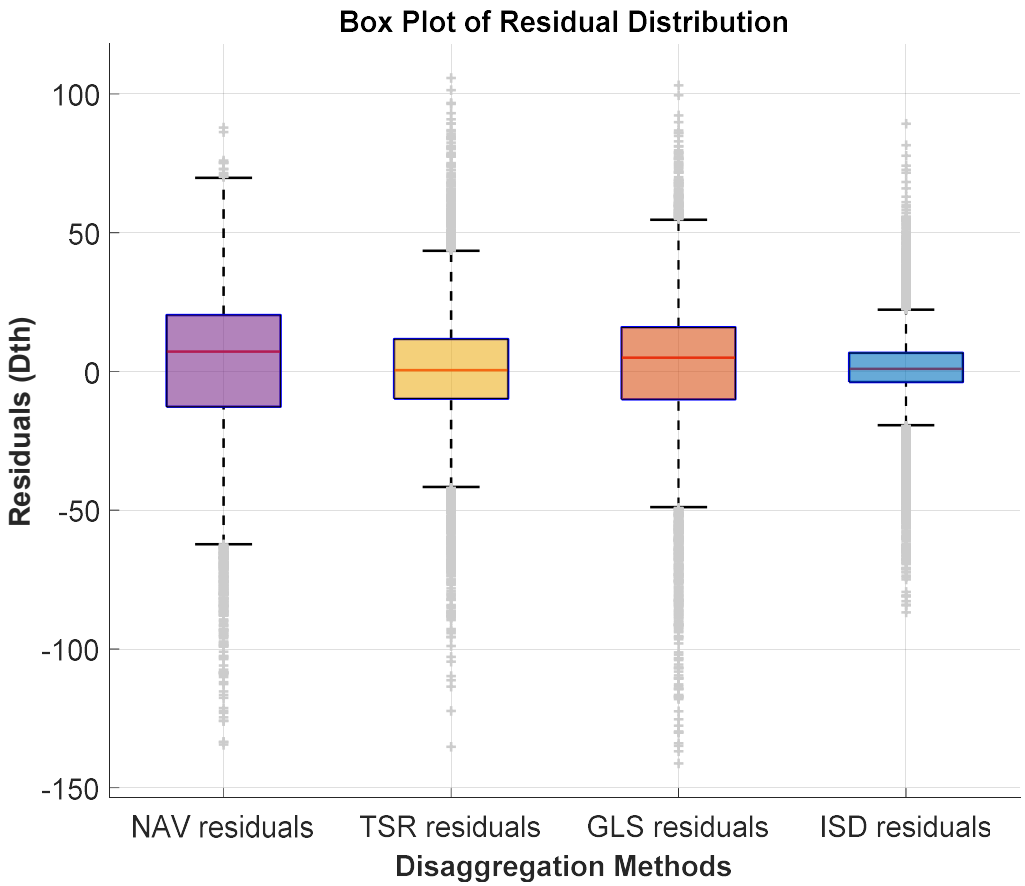

Figure 4.13 – Residential 100 residuals analysis for the NAV, ISD, TSR and GLS disaggregation algorithms.

Figure 4.13 reflects the results seen in Table 3, with the NAV persistence model having the largest spread of residual error in purple. The second largest interquartile range is shown in orange and is the GLS residual distribution. The TSR (yellow) and ISD (blue) have narrower residual spreads, with ISD algorithm performing best overall in the Residential 100 data disaggregation. Further conclusions are drawn between the temperature dependence between the A, B, and C Operating Area and Residential 100 data sets, and how they relate to each set of residuals in the following discussion section.

4.5 Discussion of Disaggregation Results

The Iterative Shifting Disaggregation (ISD) algorithm is carried out in an iterative two-phase process; where first a prediction phase uses multiple linear regression to generate high frequency time series based on the relationship between the observable low frequency series and high frequency exogenous correlated variables. Then second, load profiles are used in a piecewise linear update phase to "shift" or redistribute the low frequency consumption observations among the underlying high frequency periods within each low frequency measured interval. This two-step process repeats, with each subsequent prediction phase modeling the updated estimates produced in the preceding update phase.

The ISD algorithm relies on the use of highly correlated exogenous variables measured at the desired target frequency of $\hat{y}$ in the prediction phase. Without such variables, the multiple linear regression fails to generate accurate high-frequency estimates, leading to incorrect distribution of low frequency consumption observations in the ISD shifting step. It is the careful selection of these exogenous variables using domain knowledge that make the ISD algorithm perform well.

We used eight daily weather-based time series as our exogenous variables. All forecasting methods show a strong correlation with the actual values, indicating they are reasonable models. The eight exogenous variables in Table 4 correspond to: one constant ($Var1$), a linear trend ($Var2$), heating degree days $HDD65$ and $HDD55$ ($Var3-4$),

cooling degree day $CDD65$ ($Var5$), and four seasonal variables ($Var6 - 9$) used to capture yearly seasonality [61]. The same variables were used as inputs to all other disaggregation algorithms except for NAV.

The ISD algorithm does not have more information than TSR and GLS disaggregation methods yet distinguishes itself as the best disaggregation approach in the case studies presented in Section 4. This success is attributed to the ISD constrained shifting to iteratively redistribute low frequency load observations among underlying high-frequency periods. The proportion of load shifted is dependent on the correlation between all independent exogenous variables and the underlying gas consumption series per-ISD-iteration. In the case studies in Section 4, which use extremely simple natural gas demand forecasting models, the effectiveness of distributed shifting is dependent on each dataset's temperature sensitivity. The correlation between the temperature sensitive exogenous variables and the underlying gas consumption series per-ISD-iteration are shown in Table 4.

Table 4. Correlation between independent exogenous variables and underlying gas consumption series.

Iteration no.	Exogenous Variable							
	$Var2$	$\boldsymbol{Var3}$	$\boldsymbol{Var4}$	$\boldsymbol{Var5}$	$Var6$	$Var7$	$Var8$	$Var9$
Temperature Sensitive LDC Operational Area (Operating Area)								
1	0.29	0.92	0.90	-0.53	0.61	0.09	0.21	0.14
2	0.25	0.93	0.92	-0.37	0.75	0.19	0.35	0.16
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
10	0.21	0.98	0.97	-0.44	0.68	0.17	0.41	0.37
Non-Temperature Sensitive LDC Operational Area (Residential 100)								
1	0.03	0.72	0.75	-0.26	0.59	0.17	0.41	0.27
2	-0.06	0.76	0.61	-0.19	0.74	0.22	0.35	0.34
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
10	-0.11	0.81	0.83	-0.33	0.43	0.27	0.26	0.31

The $Var1$ constant variable is not included in the correlation analysis of Table 4. For the more temperature-sensitive Operating Area dataset, the strongest correlations are observed with temperature-related variables ($Var3$ and $Var4$), which consistently exceed 0.9 as the iterations progress. The cooling degree day variable ($Var5$) shows a moderate negative correlation, while the seasonal variables ($Var6 - 9$) exhibit varying levels of positive correlation, indicating their influence on gas consumption patterns.

In contrast, the non-temperature-sensitive Residential 100 dataset shows weaker correlations overall, especially for $Var2$ and $Var5$, which reflect the lesser impact of temperature on gas consumption in this area. However, temperature-related variables still maintain moderate to strong correlations, indicating their relevance even in non-temperature-sensitive contexts as seen in Table 4. Since we are fitting with this model as opposed to forecasting, the expected insignificant t-statistics are not as important in this context. The seasonal variables again show variable influence, particularly in later iterations.

These results support the effectiveness of the ISD algorithm in refining the correlation between the exogenous variables and gas consumption over successive iterations, particularly in temperature-sensitive regions. This is especially evident when comparing ISD estimates against GLS and TSR algorithm output (Figure 4.12) and numerical results (Table 2). Figure 4.12 shows the GLS algorithm underestimates consumption during peak heating seasons, and that the TSR algorithm more accurately follows the actual values during these periods of high volatility. Table 2 further confirms this with 4.3% $WMAPE$ improvement of TSR over GLS, as well as 1.4% $WMAPE$ improvement of ISD over TSR disaggregation results. Both GLS and TSR minimize the same cost function, but TSR,

tailored for natural gas disaggregation, performs better than the generic GLS implementation.

The TSR method operates over a single series at a time and is effective in disaggregating consumption in temperature-sensitive areas. However, the TSR algorithm fails to maintain the reconciliation constraint (4.7) and the sum of the intramonthly (disaggregated) days does not sum to the original monthly totals. This failure to meet the reconciliation constraint is intentionally violated by the authors so accuracy on days with extreme weather events may be more accurate [61]. This places how important accurate daily estimates are to the Local Distribution Companies distribution system in extreme winter months. This is again observed in Table 3, where we observe a smaller spread of $RMSE$ metrics for each algorithm but significant differences in $WMAPE$, indicating more accurate results were calculated for the ISD winter months.

Our results show that the ISD algorithm similarly can extract exogenous variable volatility as TSR and accurately redistribute gas consumption while maintaining the coherence constraint. This redistribution of load is based on the inter- and intra-series interactions of each dependent variable and the rate of gas consumption. The stronger the correlation between the low frequency observed gas intervals and high frequency independent correlated variables, the more accurate ISD disaggregate estimates become.

Under certain conditions, the ISD disaggregation algorithm will not perform well. As seen in the Table 3 results, the disaggregating the Residential 100 dataset was too dynamic of a problem to iteratively redistribute the gas consumption of 100 customers simultaneously and more accurately than the NAV persistence model. The variability introduced by considering individual residential consumption patterns outweighed and

smoothed the effects of the ISD constrained load shifting. Still, ISD is the second most accurate algorithm disaggregating the Residential 100 dataset despite the large number of time series having lower correlation with exogenous variables.

4.6 ILS Disaggregation Algorithm Conclusion

The Iterative Load Shifting (ISD) disaggregation algorithm demonstrates superior accuracy compared to Naïve, Generalized Least Squares (GLS), and Time Series Reconstruction (TSR) methods when evaluated on both operating area and individual residential gas consumption datasets. A key contribution of this work is ISD's ability to disaggregate multiple, nonuniformly spaced time series with overlapping intervals into a single daily representation while maintaining coherence constraints. By iteratively redistributing load across high-frequency periods, the algorithm effectively captures underlying consumption patterns and improves forecast reliability.

Future refinements to the ISD algorithm should focus on enhancing the initial disaggregation process in Step 1 by incorporating time series reconciliation techniques. By aligning the disaggregation process with hierarchical forecast constraints, ISD could achieve greater accuracy while preserving coherence across multiple levels of aggregation. The integration of reconciliation techniques may also improve computational efficiency by reducing the number of iterations required to reach convergence. Additionally, exploring the trade-offs between iterative-update disaggregation algorithms like ISD and optimal forecast reconciliation methods could provide insight into balancing coherence and forecast precision. While ISD ensures strict aggregation consistency, reconciliation techniques optimize forecasts by allowing minor adjustments across hierarchical levels.

A hybrid approach combining these strengths could improve both accuracy and efficiency, particularly in applications with complex hierarchies or high variability.

Beyond natural gas consumption, the ISD algorithm has potential applications in other domains where irregularly sampled data must be transformed into high-frequency estimates. Electricity usage forecasting could incorporate exogenous factors such as solar irradiance and grid load, while water demand forecasting may leverage precipitation and temperature data. Similarly, traffic flow modeling could integrate vehicle counts, road capacity, and urban development metrics. Tailoring the ISD framework to domain-specific constraints would enable its application in a wide range of time series disaggregation challenges, provided coherence constraints align with each field's requirements.

The iterative nature of ISD makes it particularly suited for applications requiring strict consistency, such as energy distribution and resource planning. In contrast, reconciliation techniques offer greater flexibility by optimizing forecasts across hierarchical structures [58]. A combined methodology that integrates ISD's coherence-preserving structure with reconciliation's forecast refinement capabilities could further enhance predictive performance. Future work should examine how reconciliation-based load-shifting can improve ISD's speed and accuracy, particularly in large-scale forecasting environments where both coherence and precision are critical.

CHAPTER 5

Single Dimension Hierarchical Reconciliation of Natural Gas Data

Local natural gas distribution companies (LDCs) rely on accurate historical gas consumption data to facilitate decision-making across various functions within their organization. Historically, natural gas models have been developed at a single level of aggregation, either for specific timeframes (e.g., hourly or daily) or across different geographic operating areas [17]. However, inconsistencies of historical data within these hierarchies frequently arise, where observed gas consumption at lower levels fails to aggregate correctly to total consumption at higher levels. Such historical data incoherence, whether temporal (misalignment between hourly, daily, and monthly demand) or spatial (discrepancies across regions or customer classifications), undermine the integrity of historical datasets used for forecasting.

This chapter introduces Single Dimension Hierarchical Reconciliation, a process that refines structural inconsistencies in historical gas consumption data by ensuring alignment with known physical and operational constraints within a single hierarchy - either temporal or spatial. Unlike forecast reconciliation, which adjusts predictive estimates, this chapter focuses exclusively on reconciling historical data to remove aggregation discrepancies that could distort model learning.

By eliminating statistical inconsistencies in recorded consumption, reconciliation ensures that forecasting models learn from an internally coherent dataset rather than from erroneous data relationships. This process prevents the introduction of biases caused by misaligned historical records, improving the reliability of forecast model training. The structured reconciliation methods presented here serve as an essential preprocessing step,

allowing predictive models to be trained on coherent and operationally meaningful historical data. The impact of this process is further explored in the historical data reconciliation case studies, where improvements in forecasting performance are quantified.

5.1 Objectives of Hierarchical Data Reconciliation of Natural Gas Data

Natural gas consumption data is inherently structured within hierarchical relationships that reflect how measurements are recorded and aggregated by an LDC. These hierarchies exist along both temporal and spatial dimensions. Temporal hierarchies define relationships between hourly, daily, and monthly consumption, ensuring that higher-level aggregates reflect the sum of lower-level measurements. Spatial hierarchies reflect the organization of gas consumption across regions, operating areas, or customer classes, where localized measurements should sum to total system demand.

Despite these structured relationships, real-world data often exhibits incoherencies, where recorded consumption at finer levels fails to aggregate correctly to broader totals [75]. These inconsistencies, referred to as data incoherence, may arise from data collection errors, missing values, or operational discrepancies commonly encountered in gas distribution. Left unaddressed, such discrepancies distort statistical relationships within the historical data, weakening the integrity of historical datasets used for model training and ultimately reducing the reliability of demand forecasting in later stages [76].

Hierarchical data reconciliation is introduced as a corrective measure to resolve these incoherencies by enforcing internal coherence within a single historical hierarchical structure [62]. The primary objective is to ensure that historical gas consumption data correctly aggregates within either a temporal or spatial hierarchy. By ensuring coherency

before the historical data is used in forecasting models, the reconciliation process prevents distortions that could introduce biases during model training. This results in higher-quality input data, which improves the performance of forecasting models trained on reconciled datasets, as we will show in Section 5.4.

In a temporal hierarchy, gas data must be structured to reflect consumption patterns at each frequency level. This process involves segmenting data into hourly, daily, and monthly observations to maintain alignment within the reconciliation framework. Structuring temporal variables in this way enables bottom-up reconciliation, where lower-frequency forecasts are adjusted to align with total demand at each subsequent level. Techniques such as proportional adjustments and scaling factors help distribute discrepancies across time levels, allowing forecasts to maintain a coherent progression from the finest to the coarsest temporal resolution. Similarly, spatial reconciliation applies these principles across geographic hierarchies, where consumption data is structured at various regional levels. Localized measurements, such as gas use at the city or district level, must align with broader regional aggregates. Mapping matrices integrate local consumption data into regional totals, reducing discrepancies that could otherwise distort operational decision-making.

Processing hierarchical variables in both temporal and spatial dimensions ensures that historical consumption data remains consistent at each level of aggregation. This structured approach not only aligns hierarchical data levels across dimensions but also reinforces the integrity of higher-level consumption records, providing a solid foundation for more complex cross-temporal reconciliation covered in the following chapter.

For illustration, Figure 5.1 present scenarios in which temporal and spatial hierarchical levels in a natural gas distribution system within a single U.S. state exhibit incoherence. The scenario is incoherent spatially as the gas is consumed in the three different geographic sub-regions, A, B, and C do not aggregate to the $Total$ gas burned throughout the state. Temporal incoherence is evident in hourly values A, B, and C do not aggregating across 24-hour periods to equal their daily counterparts.

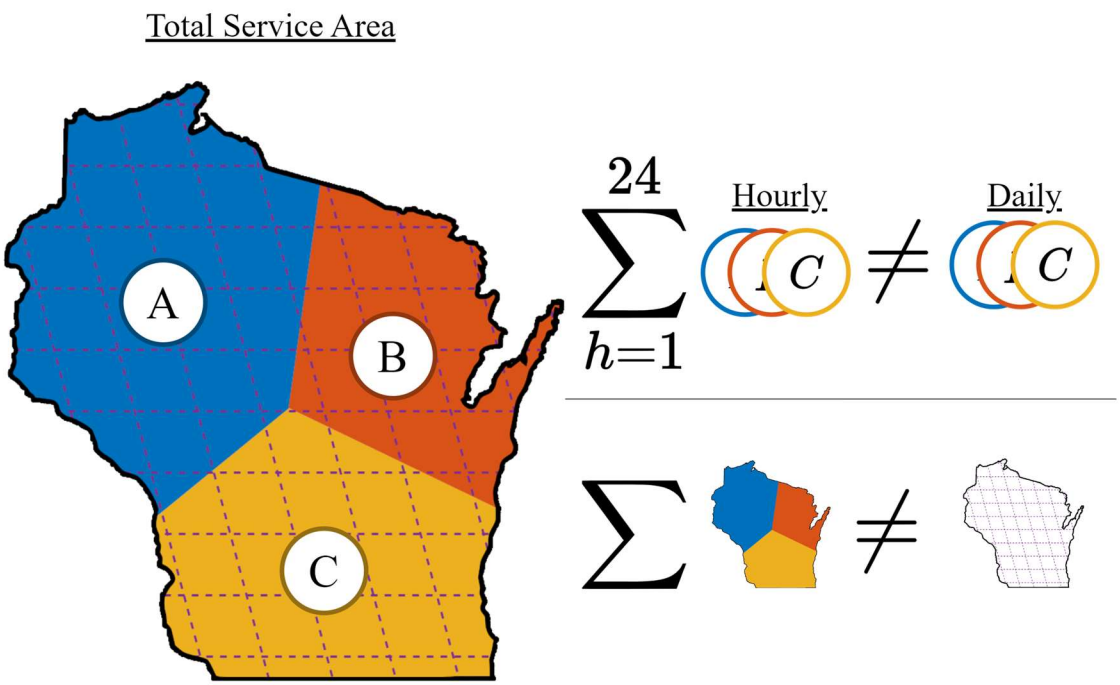

Figure 5.1 – Illustration of an incoherent natural gas distribution scenario in which gas consumed in operating areas A, B, and C does not aggregate spatially or temporally.

When hierarchical incoherencies persist, confidence in model reliability diminishes, complicating both operational and strategic planning [62]. Such incoherence in historical data may result in suboptimal decisions, misallocation of resources, and an increased risk of operational disruptions, ultimately affecting an LDC's ability to meet customer demand efficiently [57]. Hierarchical reconciliation of historical data, as presented in this chapter, is a necessary step to correct these incoherencies before they propagate into forecasting models. By addressing historical data inconsistencies at the preprocessing stage, LDCs can ensure their forecasting models are trained on coherent and physically meaningful datasets, leading to improved demand predictions in later chapters.

5.2 Single Dimension Hierarchical Reconciliation

Hierarchical reconciliation is a corrective approach that resolves incoherencies in historical gas consumption data by ensuring coherence within a single hierarchy, either temporal or spatial. This section focuses on cases where reconciliation is performed along a single dimension, aligning historical consumption data either across time intervals (temporal reconciliation) or across geographic levels (spatial reconciliation).

Using the scenario presented in Figure 5.1, an LDC was partitioned into three operating areas, A, B, and C, and into hourly and daily forecast frequencies. Figure 5.2 shows the observed series at each of these frequencies, including the aggregated hourly and daily LDC total series Y_{daily} and Y_{hourly}.

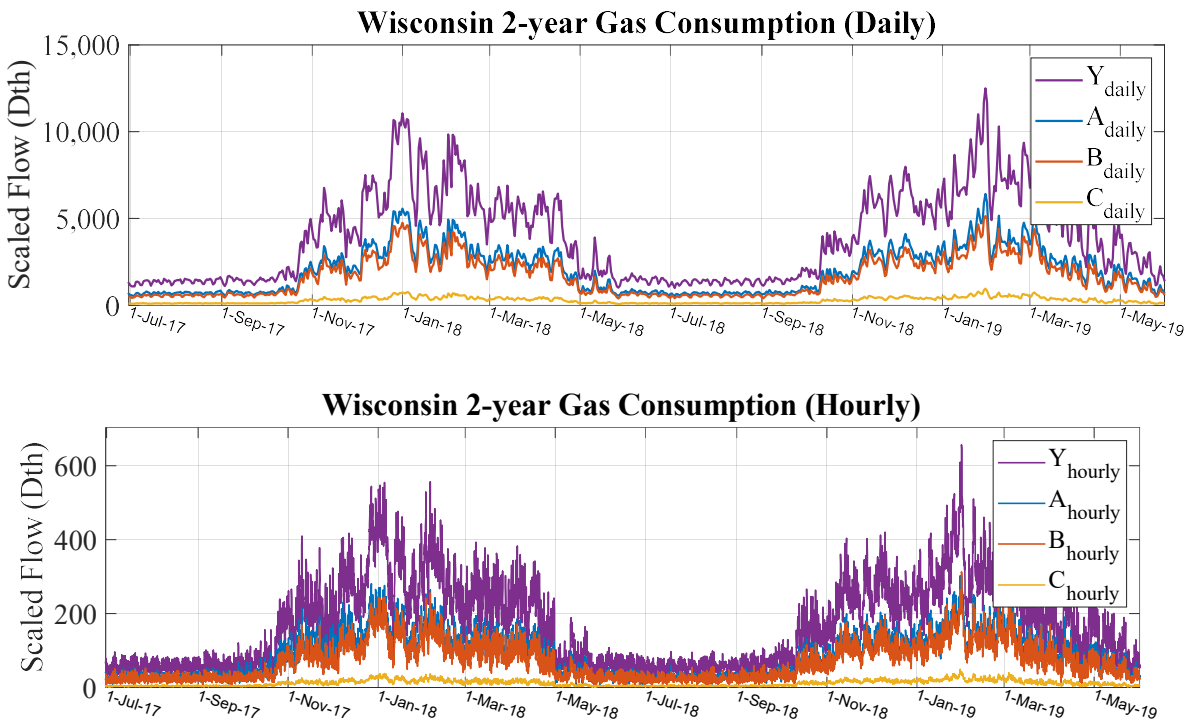

Figure 5.2 – Hourly and daily components of hierarchical natural gas consumption series, Y.

The series shown in Figure 5.2 can be organized into a hierarchy according to the LDC's spatial and temporal constraints. We propose individual spatial and temporal hierarchies to reconcile natural gas demand data next. The hierarchies introduced in Section 5.2.2 and 5.2.1 are based on [57].

5.2.1 Spatial (Cross-sectional) Hierarchical Reconciliation

Spatial reconciliation aligns data across different geographic levels, typically aligning local estimates with larger aggregate regions to ensure coherence. For example, a city-level data would need to integrate coherently within the state-level data, ensuring that the spatially distributed forecasts at lower levels align with higher-level geographic

aggregates. Spatial reconciliation techniques often use distribution keys, historical allocation patterns, and proportional adjustments to resolve any discrepancies among different geographic layers.

Maintaining spatial coherence is essential for distribution companies, as it impacts regional resource allocation and allows for efficient gas distribution. By aligning spatial data, LDCs can respond more effectively to local demand fluctuations, making accurate resource adjustments that optimize operational outcomes. Furthermore, coherent spatial data supports regulatory and planning requirements, which often necessitate coherency across jurisdictions.

Figure 5.3 shows the spatial structure of a two-level LDC hierarchy.

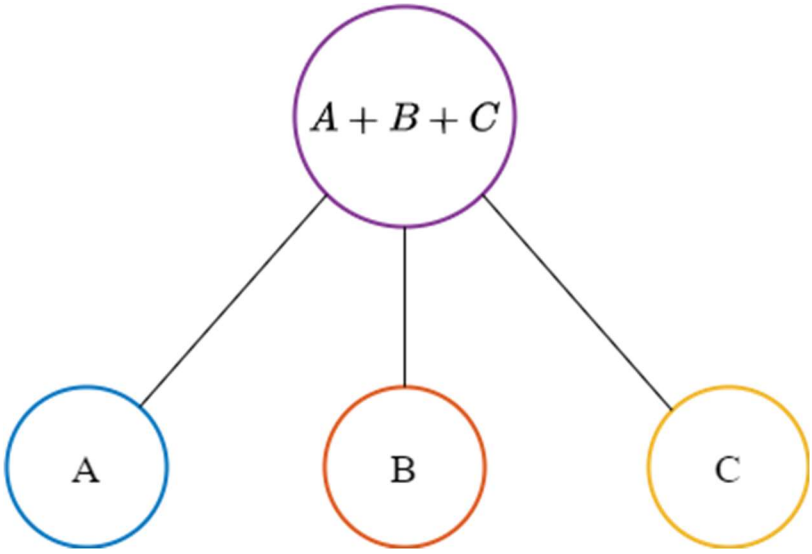

Figure 5.3 – Two-level LDC hierarchies showing spatial aggregation structure.

Figure 5.3 illustrates the spatial structure for operating areas A, B, C, and $Total$. Let the number of nodes in the hierarchy be represented by scalar m, and the number of nodes at the lowest hierarchical level available be n. We convert the hierarchy in Figure 5.3 into

its summation matrix to form S of size $m \times n$. Let $S_{spatial}$ be the summation matrices for the spatial hierarchy illustrated in Figure 5.3, respectively:

$$S_{spatial} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5.1}$$

$S_{spatial}$ is a 4 × 3 matrix, with the number of rows, $m = 4$, corresponding to the total number of spatial nodes in Figure 5.3, and number of columns, n = 3, corresponding to the number of nodes at the lowest hierarchical level (in this case, the level containing series A, B, and C). The observations shown in Figure 5.3 are also measured at times $t = 1, 2, \ldots\ N$, with N representing the total number of observations in a single series at the lowest hierarchical level. The hierarchical time series, $y_{spatial,t}$, is an m-vector containing all observations made at time t.

$$y_{spatial,t} = [Total_t, A_t, B_t, C_t]. \tag{5.2}$$

Figure 5.4 illustrates the structure of $y_{spatial,t}$.

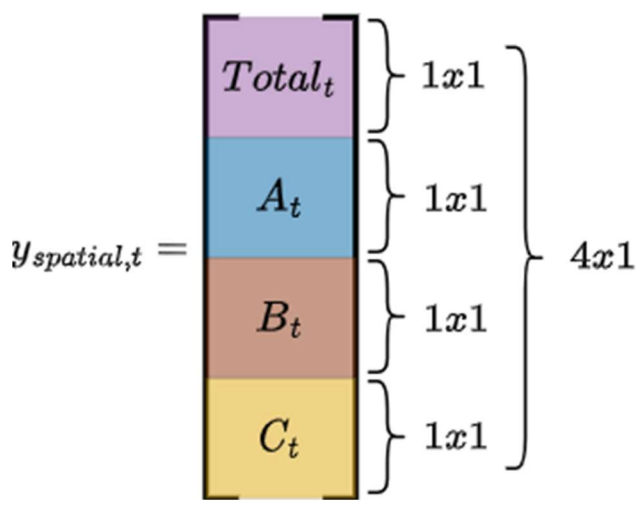

Figure 5.4 – Structure of $y_{spatial,t}$, formed by stacking observations from both hourly and daily time series.

The next section presents the temporal version of Figure 5.1. Variables $y_{spatial,t}$ and $S_{spatial}$ are used in Section 5.3.

5.2.2 Temporal Hierarchical Reconciliation

Temporal reconciliation focuses on achieving data coherence across different time scales within the hierarchy—such as hourly, daily, and monthly aggregations. In practice, this means ensuring that the aggregated data values across various frequencies remain consistent with the total demand. The reconciliation process begins by identifying the data at each temporal level, and applying techniques that adjust discrepancies to maintain alignment.

For instance, in natural gas forecasting, daily forecasts must align with monthly values, ensuring that the sum of daily demand matches expected monthly totals. This alignment is crucial, as errors in any level of temporal hierarchy can propagate, distorting overall predictions and impacting supply chain decisions. Techniques such as bottom-up, top-down, and middle-out approaches are employed to adjust values across temporal layers, providing a structured method to refine forecast accuracy. Temporal reconciliation methods not only refine predictions but also support LDCs in aligning resources to meet fluctuating demand with a high degree of reliability.

Figure 5.5 shows the temporal structure of a two-level LDC hierarchy.

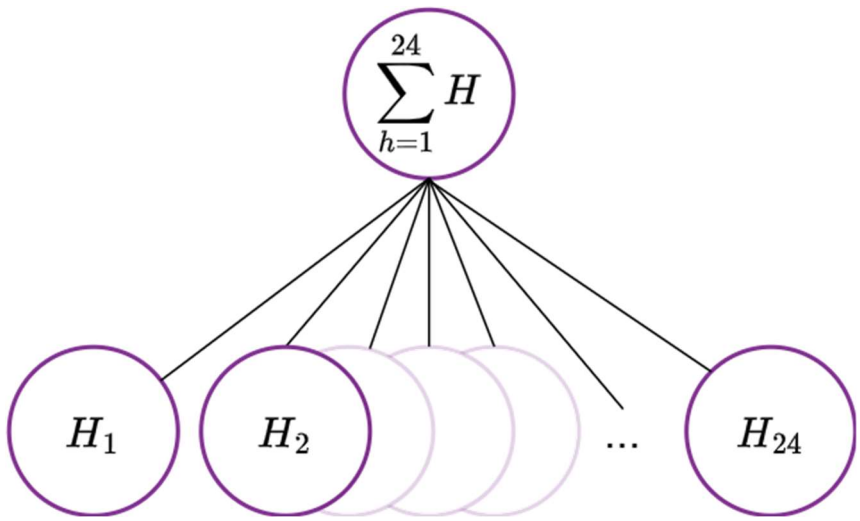

Figure 5.5 – Two-level LDC hierarchies showing temporal aggregation structure.

Figure 5.5 shows the hierarchical mapping of a temporal structure with daily demand (top level) aggregated from hourly demand (bottom-level). Let the number of nodes in the hierarchy be represented by scalar m and the number of nodes at the lowest hierarchical level available be n. We convert the hierarchy in Figure 5.5 into its summation matrix to form S of size $m \times n$. Let $S_{temporal}$ be the summation matrices for the temporal hierarchy illustrated in Figure 5.5, respectively:

$$S_{temporal} = \begin{bmatrix} 1 & 1 & 1 & & 1 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{5.3}$$

In this example, the top row of $S_{temporal}$ is a unit vector of length n. The summation matrix, $S_{temporal}$, is 25 × 24. The rows correspond to the 24 h and daily time series. The columns correspond to the signals at the lowest level of the hierarchy, i.e., 24 h. [9]. The observations shown in Figure 5.2 are measured at times $t = 1, 2, \ldots\ N$, with N

representing the total number of observations in a single series at the lowest hierarchical level. The hierarchical time series, $y_{temporal,t}$, is an m-vector containing all observations made at time t,

$$y_{temporal,t} = [Y_{daily,t}, Y_{hourly,t}]. \tag{5.4}$$

Not all time series necessarily have an observation at time t (i.e., no midday readings in daily time series, $Y_{daily,t}$). Let the 24 h spanned in $Y_{daily,t}$ coincide temporally with $Y_{hourly,t}$. Figure 5.6 illustrates the structure of $y_{temporal,t}$.

$$y_{temporal,t} = \left[\begin{array}{c} y_{daily,t} \\ \hline y_{hourly,t} \end{array}\right] \begin{array}{l} \}\ 1x1 \\ \}\ 24x1 \end{array} \Bigg\}\ 25x1$$

Figure 5.6 – Structure of $y_{temporal,t}$, formed by stacking observations from both hourly and daily time series.

The top level of the hierarchy, level 0, is the LDC's total daily consumption $Y_{daily,t}$. Variables $y_{temporal,t}$ and $S_{temporal}$ are used in Section 5.3.

5.3 Single Dimension Hierarchical Aggregation Constraints

This section describes the methods used to reconcile natural gas data across spatial and temporal levels. The summation matrices $S_{spatial}$ and $S_{temporal}$ represent common gas operation hierarchies [9]. Comparing (5.1) with (5.3) to the aggregation constraints shown in Figure 5.1, it is clear how each matrix, S, can be partitioned by levels of the hierarchy [17]. Despite these summation matrices representing common gas operation hierarchies, current LDC practice is to forecast these levels independently without coherence constraints [9].

Let b_t be an n-vector of the lowest level of the hierarchy. For example, in the spatial hierarchy, $b_t = [A_t, B_t, C_t]$. In the temporal hierarchy, $b_t = [H_1, H_2, \dots H_{24}]$, or the 24 hours within the day. To describe the intra-hierarchy relationships (constraints), we generalize $S_{spatial}$ and $S_{temporal}$ to interchangeable notation S. Similarly, we shorten hierarchical time series $y_{temporal,t}$ and $y_{spatial,t}$ to y_t. In the single dimension hierarchical reconciliation framework, we construct a summation matrix, S, to reflect the hierarchical structure seen in Figure 5.1, such that all levels in y_t aggregate in

$$y_t = S \times b_t. \tag{5.5}$$

The mathematical formulation defined in (5.5) represents the core aggregation constraints that LDCs must adhere to when analyzing gas consumption data. However, in practice, LDCs often face significant data incoherence, where recorded gas consumption at lower levels fails to sum correctly to higher levels. This incoherence has traditionally hindered

the adoption of formal reconciliation methods in LDC operations, despite growing interest in hierarchical reconciliation techniques within the broader demand-forecasting literature [8], [9].

Even with these challenges, multi-hierarchical gas consumption data still contains valuable information that can enhance forecast accuracy—provided that systematic reconciliation techniques are applied. In the remaining sections of Chapter 5, we introduce several single-dimension reconciliation algorithms designed to correct incoherencies across spatial and temporal levels. By addressing data incoherence, we aim to establish a more reliable historical dataset, improving model training and overall forecast accuracy.

5.4 Single Dimension Reconciliation Solution

In practice, LDCs often collect gas consumption data at different temporal and spatial resolutions, leading to incoherent records that must be adjusted to satisfy operational constraints. This section describes how hierarchical data reconciliation adjusts an incoherent set of historical gas consumption records into a coherent and reconciled dataset [16]. The adjusted data is then compared with the original historical consumption data at each level of aggregation to assess the impact of reconciliation. The hierarchical time series reconciliation framework proposed by Hyndman et al. [58] forms the basis of our method. Here, we assume familiarity with [51], [62], [68].

Hierarchical time series y_t is not constrained to the LDC hierarchy structures introduced in Sections 5.2.2 and 5.2.1. Thus, y_t is likely to be incoherent [50], [65]. To resolve the incoherence, we use the variance-shrinking (VS) and minimum trace (MinT) [53] reconciliation methods constrained by the spatial or temporal hierarchies illustrated

in Figure 5.3 and Figure 5.5 to produce coherent forecasts. The hierarchical reconciliation tool is the hts R package (hts 4/4.3.0) [77].

Let G be an $n\ x\ m$ mapping matrix that maps historical data, y_t, to a set of reconciled estimates, $\tilde{y}_t$. In general, the reconciliation methods according to [71] are defined as

$$\tilde{y}_t = SGy_t. \tag{5.6}$$

G linearly projects the incoherent historical data, y_t, into bottom-level disaggregated estimates that are then summed by summation matrix S [71]. For example, bottom-up reconciled estimates are obtained by setting $G = [0_{n\ x\ (m-n)} \mid I_n]$, where $0_{n\ x\ (m-n)}$ is a null matrix. The top-down reconciliation method uses proportions $p = [p_1, p_2, \dots p_n]$ in forming projection matrix $G = [p \mid 0_{n\ x\ (m-n)}]$. The bottom-up and top-down methods (Chapter 2) focus on a particular level of the hierarchy and do not share information between hierarchical levels.

5.5 Single Dimension Reconciliation Case Studies

A main contribution of this chapter is demonstrating the utility of hierarchical reconciliation of historical data in the absence of a coherent ground truth. Unlike traditional hierarchical reconciliation approaches, which assume underlying hierarchical data coherence, we apply reconciliation techniques to raw historical data that is already known to be incoherent. As such, we refer to y_t as unreconciled training data. While the reconciliation matrix G can be estimated following methodologies described by Hyndman et al. [57], it does not retain the same statistical properties typically assumed in the hierarchical reconciliation literature.

To illustrate the practical implications of single-dimension hierarchical reconciliation, we conduct case studies examining both spatial and temporal reconciliation techniques. These case studies evaluate how reconciling historical data used for model training impacts forecast accuracy and reduces incoherencies in operational decision-making. By enforcing coherence in a single dimension, either spatially or temporally, these methods mitigate discrepancies that would otherwise distort analytical insights. The ability to enhance the coherency of historical data through hierarchical reconciliation has direct applications in natural gas supply chain management, where data incoherencies can lead to inefficiencies in procurement, storage, and delivery.

Most hierarchical data processing methods assume that training data is already coherent. As reviewed in Chapter 2, several researchers have successfully reconciled forecasts using optimal combination methods derived from coherence theory. However, a key challenge in this study is that the raw historical data used is inherently incoherent. This discrepancy is illustrated in Figure 5.7, which displays incoherencies between the summed hourly gas demand and the recorded daily demand for January 2019. Similar patterns of incoherence are observed in the spatial hierarchy. These incoherencies may introduce biases into model training, potentially impairing the ability of forecasting models to generalize effectively.

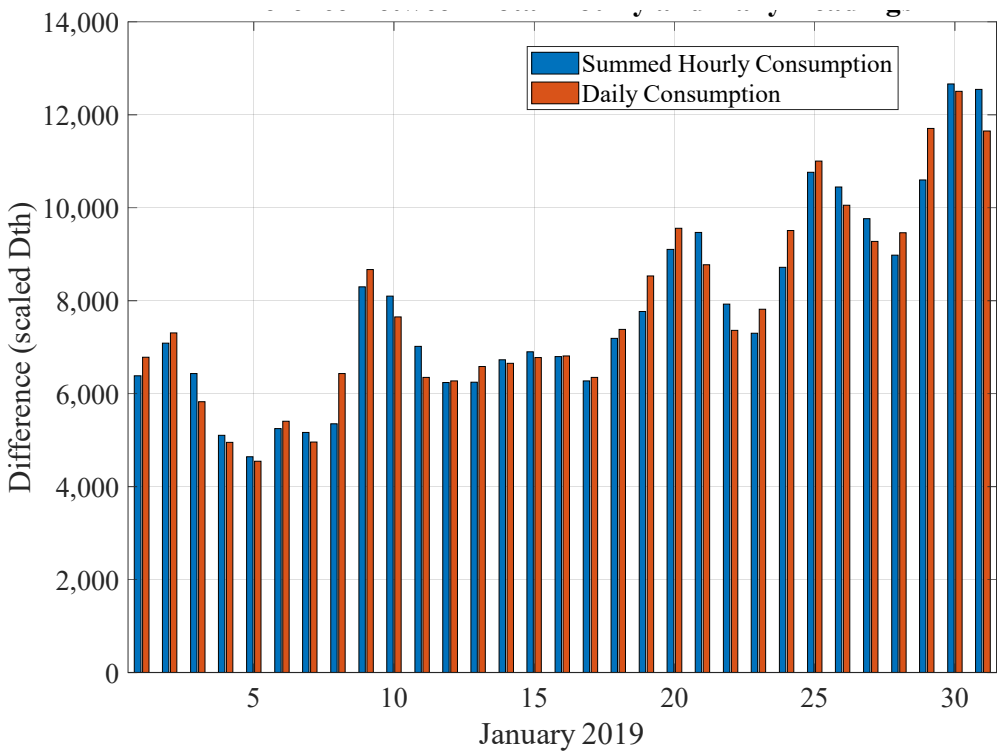

Figure 5.7 – Incoherency between daily and summed hourly gas demands. Similar incoherence is observed in spatial hierarchies.

Despite the incoherence observed in Figure 5.7, we demonstrate that applying reconciliation to historical gas consumption data before model training results in more accurate out-of-sample forecasts than training a model directly on incoherent data. The reconciliation matrix G is learned dynamically over time $t = 1, \dots, N_t$, enforcing coherence between hierarchical levels to align with both physical and operational constraints within the gas distribution system.

We evaluate the effectiveness of hierarchical reconciliation by comparing reconciled historical data against traditional, incoherent historical LDC data. The results are presented in Table 5, which shows the impact of spatial reconciliation on gas consumption data across operating areas A, B, C, and the $Total$ system level.

Table 5: Spatial Reconciliation Results for Operating Areas A, B, C, and $Total$ Hierarchy ($1\ MDth = 1{,}000\ Dth$).

Operating Area	Initial Recorded Consumption (MDth)	Reconciled Consumption (MDth)	Adjustment (%)	Discrepancy Before Reconciliation (MDth)	Discrepancy After Reconciliation (MDth)
$\boldsymbol{A}$	1,732.87	1,638.72	-5.43 %	-94.14	0.00
$\boldsymbol{B}$	1,467.36	1,504.15	2.51 %	36.79	0.00
$\boldsymbol{C}$	378.29	392.37	3.72 %	14.07	0.00
$\boldsymbol{Total}$	3,672.42	3,535.24	-3.73 %	-137.17	0.00

Following reconciliation, the total system consumption was adjusted from 3,672.42 MDth to 3,535.24 MDth, representing a -3.73% overall increase. The adjustments within individual operating areas vary, with Area A experiencing a -5.43% increase, while areas B and C saw increases of 2.51% and 3.72%, respectively. The discrepancy before reconciliation was most pronounced in Area A and $Total$, where recorded consumption exceeded its reconciled value by 94.15 and 137.17 MDth, suggesting an overestimation of demand. Conversely, Areas B and C required upward adjustments of 36.79 MDth and 14.08 MDth, respectively, correcting for initial underreporting.

These results confirm that the recorded gas consumption at the operating area level did not originally sum to the total system consumption, violating the hierarchical relationship that should exist between these levels. By applying reconciliation, the dataset is adjusted to enforce internal coherence, ensuring that total consumption across the LDC accurately reflects the sum of its component regions. These corrections are critical for forecast model training, as they ensure that the historical data used to develop forecasting models is internally consistent with operational constraints.

In the subsequent analysis, we assess whether models trained on reconciled historical data lead to more accurate out-of-sample forecasts compared to those trained on the original incoherent dataset. The effectiveness of hierarchical reconciliation in improving model performance is evaluated through temporal reconciliation, which is presented in Table 6.

Table 6: Temporal Reconciliation Results for Daily and Hourly Hierarchy (1 $MDth = 1{,}000\ Dth$).

Time Interval	Initial Recorded Consumption (MDth)	Reconciled Consumption (MDth)	Adjustment (%)	Discrepancy Before Reconciliation (MDth)	Discrepancy After Reconciliation (MDth)
Hour 1 (GasDay 9:00)	31.43	38.57	**-22.72** %	7.14	0.00
Hour 2 (GasDay 10:00)	48.59	51.32	**-5.62** %	2.73	0.00
Hour 3 (GasDay 11:00)	72.89	73.51	-0.85 %	0.62	0.00
Hour 4 (GasDay 12:00)	121.49	122.43	-0.77 %	0.94	0.00
Hour 5 (GasDay 13:00)	195.17	195.76	-0.30 %	0.59	0.00
Hour 6 (GasDay 14:00)	242.97	243.48	-0.21 %	0.51	0.00
Hour 7 (GasDay 15:00)	291.57	294.04	-0.85 %	2.47	0.00
Hour 8 (GasDay 16:00)	246.92	243.85	1.24 %	-3.07	0.00
Hour 9 (GasDay 17:00)	194.38	195.21	-0.43 %	0.83	0.00
Hour 10 (GasDay 18:00)	145.78	147.05	-0.87 %	1.27	0.00
Hour 11 (GasDay 19:00)	121.49	126.37	-4.02 %	4.88	0.00
Hour 12 (GasDay 20:00)	97.19	97.66	-0.48 %	0.47	0.00
Hour 13 (GasDay 21:00)	121.49	126.45	**-4.08** %	4.96	0.00
Hour 14 (GasDay 22:00)	196.41	203.38	**-3.55** %	6.97	0.00
Hour 15 (GasDay 23:00)	242.97	253.49	**-4.33** %	10.52	0.00

Hour 16 (GasDay 24:00)	291.57	306.82	**-5.23** %	15.25	0.00
Hour 17 (GasDay 1:00)	252.27	252.67	-0.16 %	0.4	0.00
Hour 18 (GasDay 2:00)	194.55	195.58	-0.53 %	1.03	0.00
Hour 19 (GasDay 3:00)	127.29	122.36	3.87 %	-4.93	0.00
Hour 20 (GasDay 4:00)	95.29	98.01	-2.85 %	2.72	0.00
Hour 21 (GasDay 5:00)	72.89	76.53	**-4.99** %	3.64	0.00
Hour 22 (GasDay 6:00)	60.74	64.19	**-5.68** %	3.45	0.00
Hour 23 (GasDay 7:00)	48.59	50.06	**-3.03** %	1.47	0.00
Hour 24 (GasDay 8:00)	36.45	37.68	**-3.37** %	1.23	0.00
Total Daily	3,672.42	3,612.48	-2.18 %	77.24	0.00

Temporal reconciliation addresses the incoherence present in hourly natural gas data by ensuring that the sum of hourly gas usage aligns with daily recorded totals. The improvement in forecast accuracy is statistically significant at the 0.05 level, based on the Diebold-Mariano test [74]. As shown in Table 6, the total daily gas consumption was adjusted from 3,672.24 MDth to 3,612.48 MDth, representing an reduction of 2.18% after reconciliation. This correction ensures that the hourly consumption values reflect a physically consistent distribution of demand within each day, eliminating aggregation discrepancies that could misrepresent system-wide gas usage**.**

The adjustments are not uniform across all hours and reflect corrections that are most impactful during key gas consumption periods. Notably, the morning ramp-up (Hours 21-24, 1-2) and evening peak (Hours 13-16) received the largest adjustments, aligning the reconciled hourly demand profile with the expected diurnal gas consumption pattern. The

most significant correction occurred at Hour 1 (GasDay 9:00), where consumption was adjusted by +22.72% (7.14 MDth increase), indicating a major discrepancy between originally recorded and expected demand at this hour. Similarly, Hours 15 and 16 (GasDay 23:00-24:00) saw upward adjustments of +4.33% and +5.23%, respectively, reinforcing that discrepancies were concentrated in peak demand periods.

Unlike the spatial reconciliation results, which had both upward and downward corrections across operating areas, the temporal reconciliation primarily resulted in net demand reductions across most hours. This suggests that the original dataset systematically overestimated total daily consumption while misallocating demand across certain hourly periods. Such misallocation can degrade intraday operational planning, particularly in scheduling gas procurement and adjusting supply to meet peak-hour demand.

While the overall accuracy gains appear modest, these findings confirm that data incoherence degrades forecast model learning. By enforcing hierarchical coherence in historical data, reconciliation reduces bias and prevents structural inconsistencies from propagating into out-of-sample forecasts. The next section summarizes these findings and transitions to Chapter 6, which introduces Cross-Temporal Forecast Reconciliation to address combined spatial and temporal forecast coherence.

5.6 Single Dimension Reconciliation Conclusion

This chapter demonstrated the importance of enforcing hierarchical coherency in historical natural gas consumption data before using it for forecast model training. Unlike traditional hierarchical reconciliation approaches that adjust forecasts post hoc, this study

applied reconciliation directly to historical consumption records, ensuring that operational decisions are based on coherent data.

Through spatial and temporal reconciliation case studies, we showed that historical data often contains incoherencies that violate fundamental hierarchical relationships, distorting model training and forecast performance. Figure 5.7 highlighted discrepancies in recorded gas consumption across different temporal aggregation levels, while Table 5 and Table 6 quantified corrections applied to spatial and temporal inconsistencies, respectively. The results demonstrated that reconciliation aligns historical data with known operational constraints, reducing statistical distortions in training datasets.

These adjustments are critical for ensuring that gas consumption records accurately reflect system operations. Prior to reconciliation, inconsistencies in recorded values suggested systematic over- or under-reporting across different levels of the hierarchy. By applying reconciliation techniques, we eliminated these discrepancies, providing a more reliable foundation for decision-making in natural gas distribution.

While this study focused on reconciling historical data, ensuring coherence at the data level does not inherently guarantee coherence in derived forecasts. Single-dimension reconciliation ensures coherence within either a spatial or temporal hierarchy, but it does not address inconsistencies that emerge when forecasts are generated across multiple hierarchical dimensions. The next chapter expands on this foundation by introducing Cross-Temporal Hierarchical Forecast Reconciliation, a methodology designed to simultaneously enforce coherence across both spatial and temporal levels

CHAPTER 6

Cross-temporal Hierarchical Reconciliation of Natural Gas Forecasts

Local natural gas distribution companies (LDCs) require accurate demand forecasts across various time periods, geographic regions, and customer class hierarchies. While Chapter 5 focuses on reconciling historical natural gas consumption data to enforce hierarchical coherence within training datasets, Chapter 6 extends this concept cross-temporally to out-of-sample forecasts. Instead of correcting past consumption records, the focus here is on ensuring that demand forecasts remain coherent across spatial and temporal levels. This is the usual domain of the forecast reconciliation of Hyndman and colleagues [77], [78].

This chapter introduces a cross-temporal hierarchical reconciliation framework that unifies the spatial and temporal hierarchies from Chapter 5 into a single structure. This approach ensures that forecasts at finer resolutions (e.g., hourly or regional demand forecasts) remain aligned with broader aggregate forecasts (e.g., daily or system total). Cross-temporal coherence is critical for LDCs, as misaligned forecasts can lead to inefficiencies in gas supply management, infrastructure planning, and regulatory compliance.

We apply our method to a case study involving three operational regions, forecasting at different geographical levels and analyzing both hourly and daily frequencies. Trained on five years of historical data and tested on one year, our model achieves a 10% reduction in hourly mean absolute scaled error and a 3% reduction in daily mean absolute scaled error.

In the remainder of Chapter 6, we introduce the cross-temporal hierarchy, extending the reconciliation framework introduced in Chapter 5 to natural gas demand forecasting. The contributions of this Chapter are a novel cross-temporal forecast reconciliation framework tailored for natural gas demand forecasting. State-of-the-art gas demand forecast solutions are enhanced with cross-temporal reconciliation techniques**.** Forecast reconciliation is an a posteriori refinement of existing forecasts—any technique or preferred style of estimation can be used to generate the set of initial base forecasts. Therefore, this work, intends to pick up where many leave off, after forecast generation.

6.1 Cross-temporal Hierarchical Structure and Constraints

Until recently, hierarchical structures used for forecast reconciliation were typically constrained to a single dimension of analysis, either spatial or temporal. In cross-sectional hierarchies, such as $S_{spatial}$ from Equation (5.1), constraints are imposed based on the unit of analysis, typically geographic regions, while temporal hierarchies, $S_{temporal}$ enforce aggregation constraints based on time units, such as hourly or daily periods. The introduction of cross-temporal hierarchies allows both dimensions to be incorporated simultaneously, providing a more comprehensive reconciliation of natural gas forecasts [57].

We built a new cross-temporal hierarchy using the same data shown in Figure 5.2, incorporating both spatial and temporal constraints in a unified framework. This hierarchical extension ensures that forecasts at finer resolutions, such as hourly and regional levels, remain coherent with broader aggregate forecasts, such as daily and system-wide totals.

For our new cross-temporal hierarchy, let $n = 72$ be the total number of series at the most disaggregated level (24 h × 3 operating areas) and $m = 100$ be the total number of nodes within our cross-temporal hierarchy. The hierarchical time series, y_t, is an m-vector containing all observations made at time t.

$$y_t = \left[Y_{daily,t},\ A_{daily,t}, B_{daily,t}, C_{daily,t}\ Y_{hourly,t}, A_{hourly,t}, B_{hourly,t}, C_{hourly,t}\right]. \quad (6.1)$$

Not all time series necessarily have an observation at time, t (i.e., no midday readings in daily time series, $Y_{daily,t}$). Let the 24 h spanned in $Y_{daily,t}$ coincide temporally with $Y_{hourl\ \ ,t}$. Figure 6.1 illustrates the cross-temporal structure of y_t.

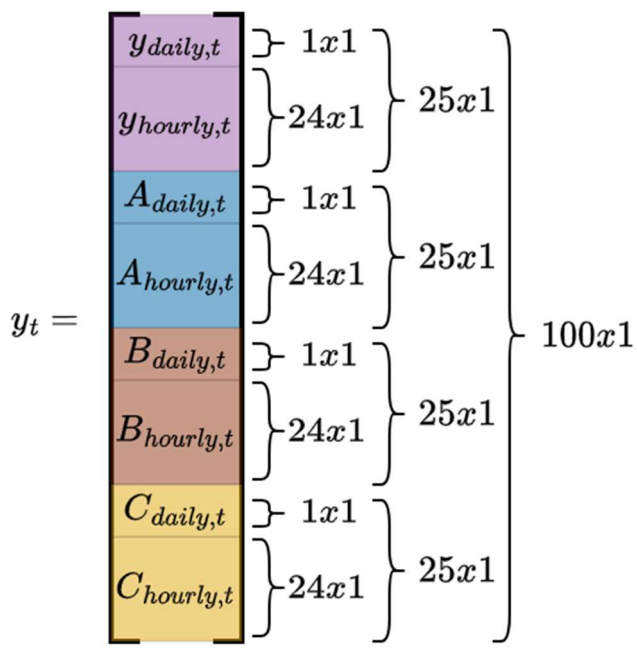

Figure 6.1 – Structure of y_t, formed by stacking observations from each time series.

The top level of the hierarchy, level 0, is the LDC total daily demand, Y_{daily}. Let b_t be an n-vector of the lowest level of the hierarchy $[A_{hour\ \ \ ,t}, B_{hourly,t}, C_{hourly,t}]$. We construct a summation matrix, S, to reflect the hierarchical relationship

$y_t = S \times b_t$. This relationship is same as (5.5). However, the summation matrix S introduced in this chapter enforces coherence across both time and space and is later used to reconcile hierarchical forecasts $\hat{y}$. An illustrated version of such a cross-temporal hierarchical structure is shown in Figure 6.2.

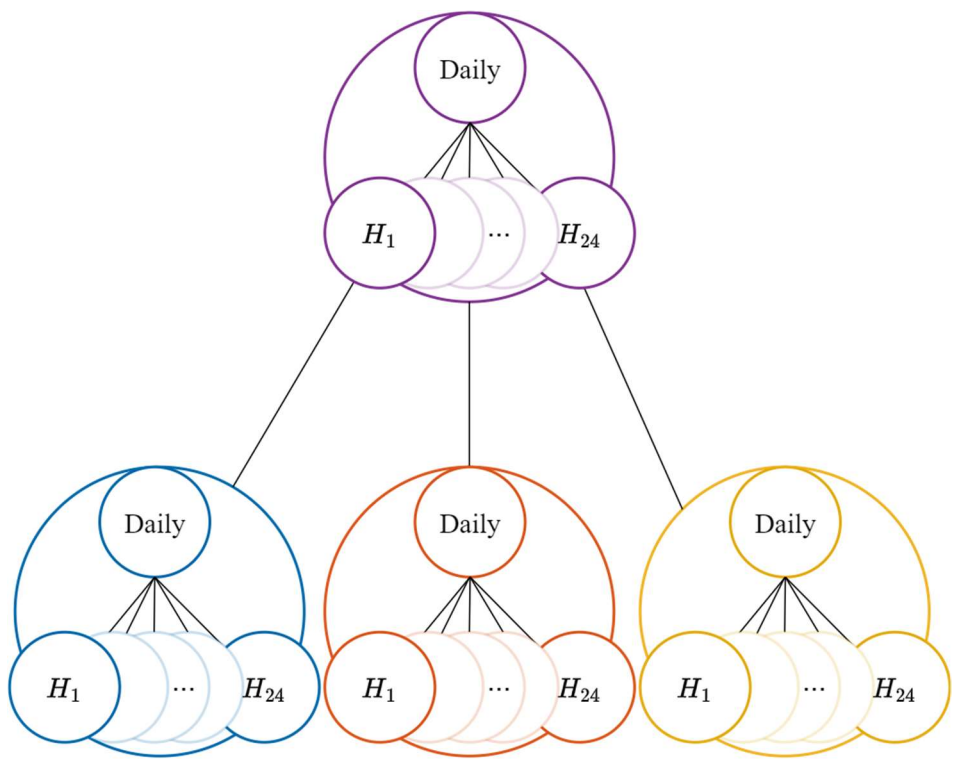

Figure 6.2 – Cross-temporal hierarchy diagram showing the aggregation relationships between hierarchical levels (A – blue, B – orange, C – yellow, $Total$ – purple).

This diagram highlights the aggregation relationships between hierarchical levels, where regional forecasts for operating areas A. B,and C aggregate into the system $Total$.

Our example cross-temporal summation matrix, S, has a size of 100 × 72, which is shown on the right side of Figure 6.3.

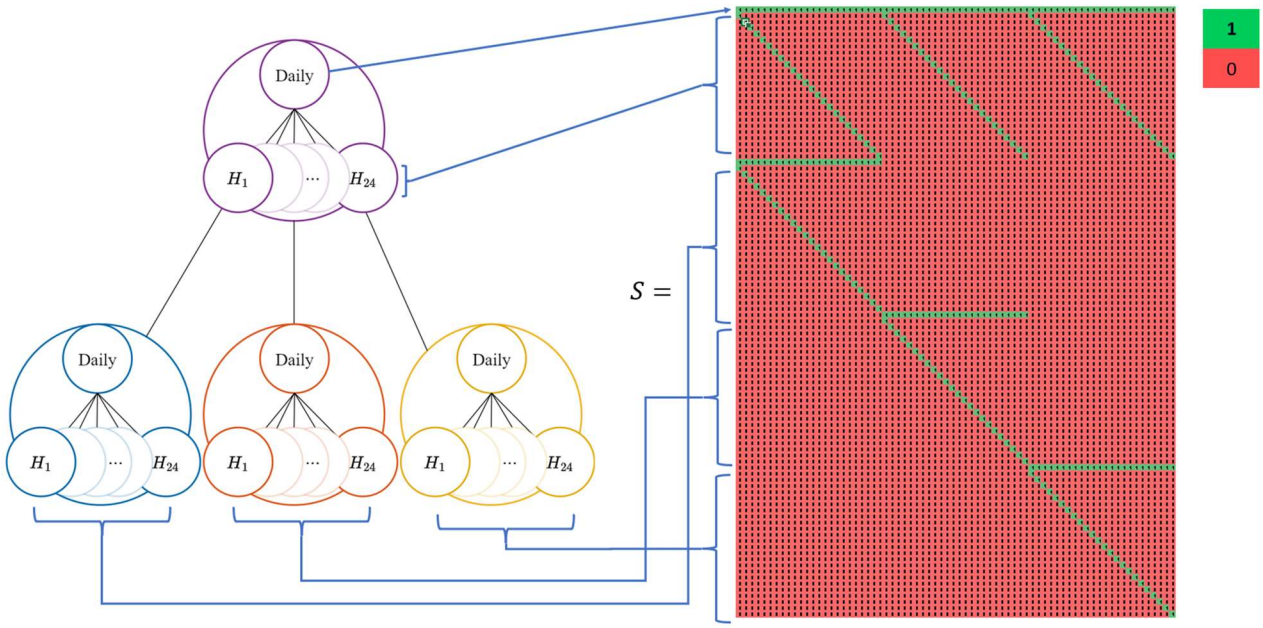

Figure 6.3 – Mapping from cross-temporal diagram to summation matrix S.

The blue lines in Figure 6.3 help visually map where the spatial and temporal aggregation constraints from the single dimension hierarchies are present. In the next section, we apply the relationship defined in Equation (5.5) and illustrated in Figure 6.3 to enforce cross-temporal hierarchical coherence, generating reconciled forecasts that align across both spatial and temporal dimensions.

6.2 Cross-temporal Hierarchical Forecast Solution

Cross-temporal forecast reconciliation begins by first defining the base forecasts, which represent the initial, incoherent set of natural gas demand projections generated independently for each hierarchical level. Base forecasts are the incoherent set of individualized gas forecasts, which may be generated using any forecasting method [9]. The hierarchical relationship defined in Equation (5.5) is applied here to establish the aggregation

relationships across both spatial and temporal dimensions, ensuring that forecast reconciliation adheres to the underlying structural constraints of the system.

Let h be the forecast horizon and $\hat{y}_h$ be h-step-ahead base forecasts for each gas demand series in the hierarchy. The base forecast vector follows the same hierarchical structure as y_t and is given by

$$\hat{y}_h = \begin{matrix} [\hat{Y}_{daily,h}, \hat{A}_{daily,h}, \hat{B}_{daily,h}, \hat{C}_{daily,h}, \ldots \\ \hat{Y}_{hour\ \ ,h}, \hat{A}_{hourly,h}, \hat{B}_{hourly,h}, \hat{C}_{hourly,h}]' \end{matrix} \tag{6.2}$$

The standard practice for gas demand forecasting is to use exogenous variables that match the temporal frequencies in our gas hierarchy (daily and hourly) during the model training of the multivariate time series, $\hat{y}_h$ [5].

The cross-temporal hierarchical reconciliation solution is similar to Equation (5.6) but indexed by out-of-sample forecast time horizon h

$$\tilde{y}_h = SG\hat{y}_h. \tag{6.3}$$

G linearly projects the base forecasts, $\hat{y}_h$, into bottom-level disaggregated estimates that are then summed by summation matrix S [71]. The MinT reconciliation matrix, G, is

$$G = (S'W_h^{-1}S)^{-1}S'W_h^{-1}. \tag{6.4}$$

Given an appropriately defined summation matrix S (Figure 6.3) that represents the structural constraints of the gas distribution hierarchy, solving for the reconciliation matrix G ensures that the resulting set of forecasts $\tilde{y}_h$ remains coherent across both spatial and temporal dimensions. This reconciled forecast retains internal coherence between all levels of aggregation, mitigating the inconsistencies that typically arise in independently estimated forecasts.

In the next section, we apply the cross-temporal hierarchical reconciliation framework to real-world natural gas demand forecasts, evaluating its effectiveness in improving forecast coherence across multiple spatial and temporal resolutions.

6.3 Cross-temporal Hierarchical Reconciliation Case Studies

While Chapter 5 demonstrated the effectiveness of hierarchical reconciliation in correcting historical consumption inconsistencies, the case studies in this section extend these benefits to out-of-sample forecasts, ensuring that future demand projections maintain coherence across temporal and spatial hierarchies. If cross-temporal reconciliation successfully enforces coherence, we expect reductions in forecast error metrics, particularly at finer resolutions where discrepancies are most pronounced. We use scale-independent error metrics accounting for the differences in scales across hierarchical levels to evaluate performance [79]. Let the naïve model (NAV) produce incoherent forecasts

$$\hat{y}_{t+h}^{NAV} = y_t. \tag{6.5}$$

for both daily and hourly series. The naïve method produces forecasts equal to the last observed value and is frequently used as a benchmark against more sophisticated reconciliation techniques [51]. The base forecasts are included in the following performance analysis to compare the reconciled forecasts performed against the incoherent estimates. We evaluate the forecasting performance of the naïve, base (BF), variance-shrinking (VS) and minimum trace (MinT) [53] reconciliation methods in terms of accuracy and bias.

Forecast skill is measured as the mean absolute scaled error ($MASE$) and root mean squared scaled error ($RMSSE$). The mean absolute scaled error is:

$$MASE = \frac{\frac{1}{h}\sum_{t=n+1}^{n+h} |y_t - \hat{y}_t|}{\frac{1}{n-1}\sum_{t=2}^{n} |y_t - y_{t-1}|}. \tag{6.6}$$

The in-sample MAE is favored by Hyndman and Koehler for its consistent availability and effective scaling of errors [79]. Root mean squared scaled error

$$RMSSE = \sqrt{\frac{\frac{1}{h}\sum_{t=n+1}^{n+h} (y_t - \hat{y}_t)^2}{\frac{1}{n-1}\sum_{t=2}^{n} (y_t - y_{t-1})^2}}. \tag{6.7}$$

$RMSSE$ serves as a standardized measure, providing an indication of the relative magnitude of errors by normalizing them based on the in-sample MSE. Absolute mean scaled error ($AMSE$) is

$$AMSE = \frac{\frac{1}{h}|\sum_{t=n+1}^{n+h} y_t - \hat{y}_t|}{\frac{1}{n-1}\sum_{t=2}^{n} |y_t - y_{t-1}|} \tag{6.8}$$

and minimizes error using the median.

All measures are scale-independent, meaning averaging across series is possible [51]. Reducing forecast error through hierarchical reconciliation improves the reliability of scheduling and infrastructure planning for LDCs. Lower RMSE values ensure that system-wide forecasts align with regional operational realities, mitigating risks in supply chain management. Table 7 compares these metrics and their averages across the cross-temporal hierarchy.

Table 7: Forecasting accuracy metrics $MASE$, $RMSSE$, and $AMSE$ for base forecasts (BF), VS, and MinT across hourly and daily resolutions and operating areas A, B, C, and $Total$

	A		**B**		**C**		**Total**		**Average**	
Resolution:	*Hourly*	*Daily*	*Hourly*	*Daily*	*Hourly*	*Daily*	*Hourly*	*Daily*	*Hourly*	*Daily*
Method	**MASE**									
BF	3.14	0.69	2.34	0.68	3.11	0.75	2.95	0.68	2.92	0.70
VS	3.16	0.74	2.35	0.75	3.12	0.83	2.92	0.72	2.89	0.76
MinT	**3.10**	**0.66**	**2.29**	**0.65**	**3.06**	**0.73**	**2.84**	**0.63**	**2.82**	**0.67**
	RMSSE									
BF	2.73	0.60	**2.04**	0.60	2.81	0.66	2.63	0.60	2.58	0.61
VS	2.75	0.68	2.06	0.68	2.82	0.75	2.61	0.67	2.56	0.69
MinT	**2.70**	**0.58**	**2.04**	**0.58**	**2.79**	**0.64**	**2.60**	**0.57**	**2.55**	**0.59**
	AMSE									
BF	**0.99**	**0.42**	**0.26**	**0.17**	**0.71**	0.31	**0.71**	0.33	0.67	0.31
VS	1.01	0.42	0.28	0.16	0.72	**0.30**	0.68	**0.30**	**0.67**	**0.29**
MinT	1.09	0.45	0.33	0.18	0.86	0.36	0.75	0.33	0.76	0.33

In Table 7, lower values indicate better results. The base forecasts have inherent incoherence, while VS and MinT rows display coherent results. Any value in Table 7 exceeding 1.00 indicates that the naïve model outperformed the base forecast model. Since the base forecast models have no autoregressive terms, it is not surprising that the naïve 1-step-ahead-persistence model out-performs all hourly forecasts.

The MinT is the most accurate method in this study, with average hourly $MASE$, $RMSSE$ of $\{2.82, 2.55\}$ and daily $\{0.67, 0.59\}$ across all operating areas and temporal frequencies. These improvements are statistically significant at the 0.05 level based on the Diebold-Mariano test. Comparing these results to the base forecasts and VS reconciliation $MASE$ and $RMSSE$, Table 7 shows a 10% hourly and 3% daily improvement when

compared to incoherent base forecasts and 7% hourly and 9% daily improvement when compared to coherent VS forecasts. Examining $MASE$ and $RMSSE$ of each hierarchical level, MinT consistently demonstrates superior forecasting accuracy compared to base forecast and VS methods. This consistency shows the robustness of MinT in reconciling forecasts across different temporal and hierarchical levels.

In measures of bias, MinT does not outperform the VS technique, with average hourly and daily $AMSE$ $\{0.67, 0.29\}$. We found this observation to be interesting considering VS is closely related to the implementation of MinT, except for the full covariance matrix of forecast errors produced in (6.4) is used in MinT and only the diagonal in VS. With more context of hierarchical interactions and effects, we expected the bias to decrease with the additional information used in MinT.

Figure 6.4 provides additional insights into the relative accuracy of each reconciliation method across gas operating areas A, B, C, and $Total$.

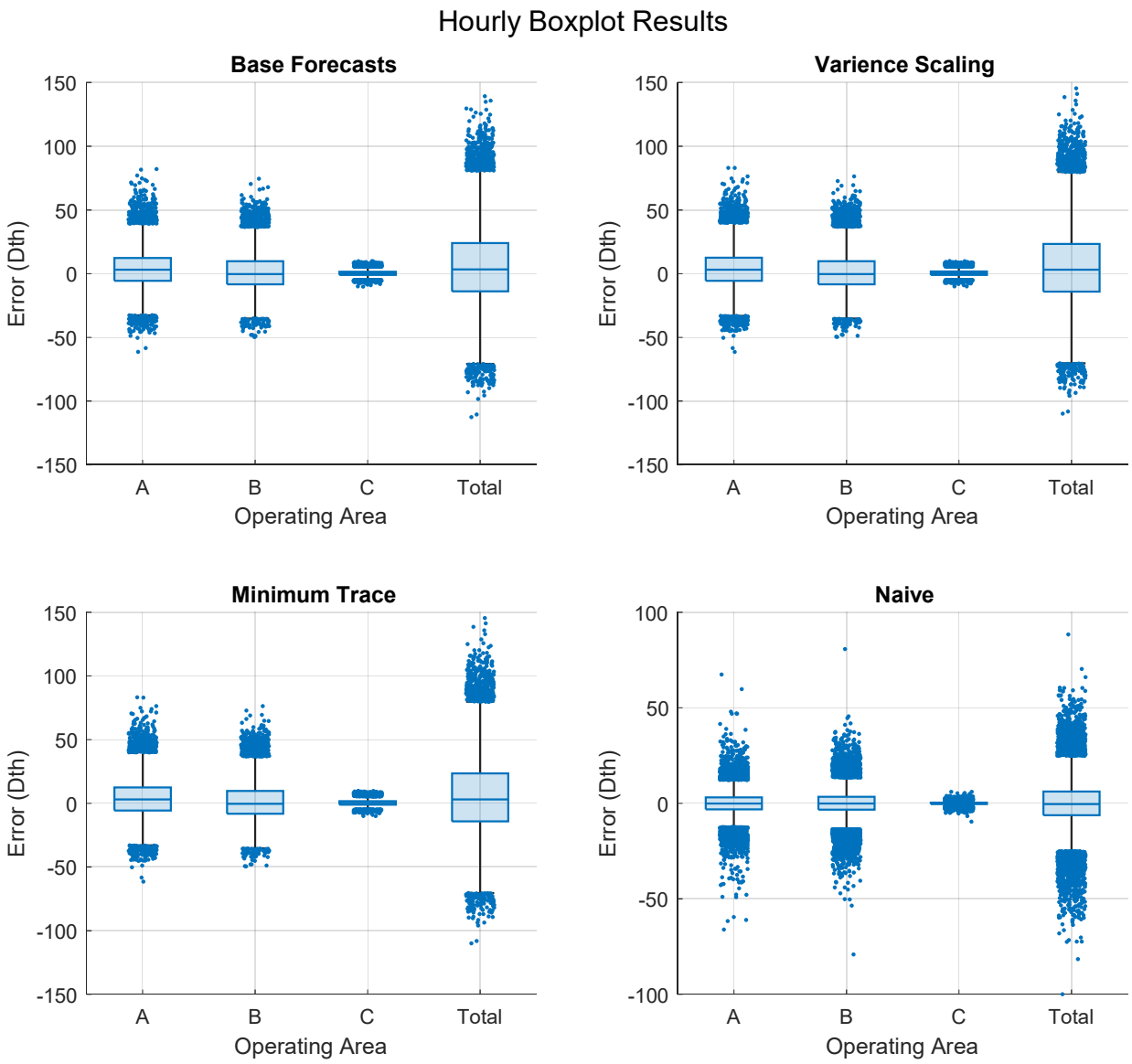

Figure 6.4 – Hourly box plot results for the error distribution across different operating areas, comparing each reconciliation method (variance scaling, minimum trace, and naïve).

Figure 6.5 shows the daily results for the same areas.

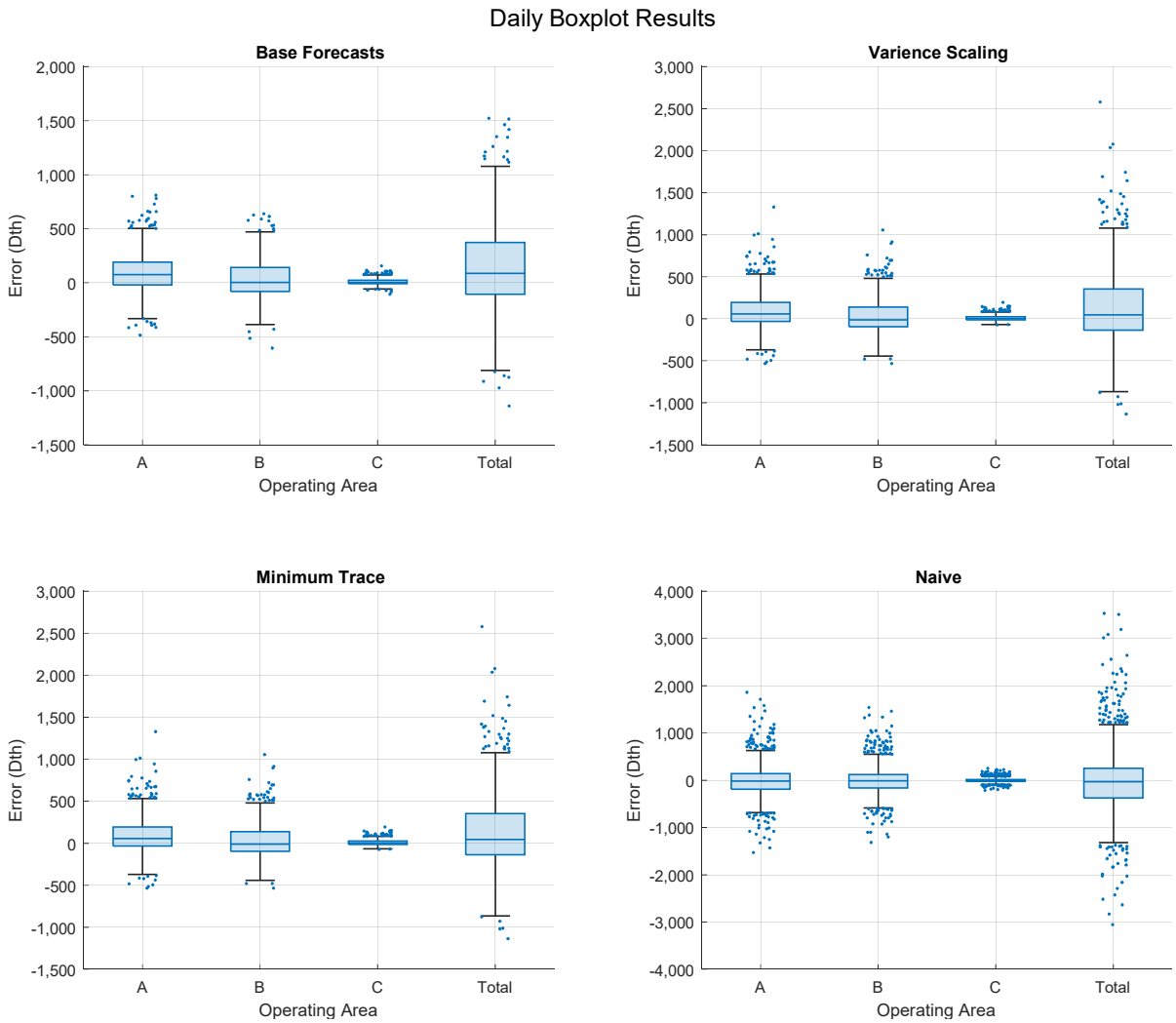

Figure 6.5 – Daily box plot results for the error distribution across different operating areas, comparing each reconciliation method (variance scaling, minimum trace, and naïve).

Natural gas demand forecasting performance is typically assessed using scaled-dependent or percent-error metrics evaluated at a single level. Both Figure 6.4 and Figure 6.5 show errors measured in dekatherms (Dth), which are calculated by taking the difference between actuals y and forecasts $\hat{y}$ for the naïve, base forecast, VS, and MinT methods. Unlike the relative out-of-sample error metrics in Table 7, these results are scale-dependent. Across both hourly and daily plots, errors are distributed around zero except for fourth quantile outliers. The VS and MinT residuals in Figure 6.4 and Figure 6.5 appear

nearly identical. This occurs because the generalized least squares solution for MinT converges closely to the ordinary least squares solution for VS. This suggests that the ordinary least squares solution effectively summarizes much, but not all, intra-hierarchy interaction. A trade-off occurs between computational efficiency and information gain while working with large hierarchies. While the VS method is outperformed in terms of accuracy, it is computationally quicker to calculate than MinT. This trade-off must be considered, especially if reconciliation is carried out over large temporal hierarchies.

Figure 6.6 and Figure 6.7 give another view of the unscaled residuals over the 2023 heating season.

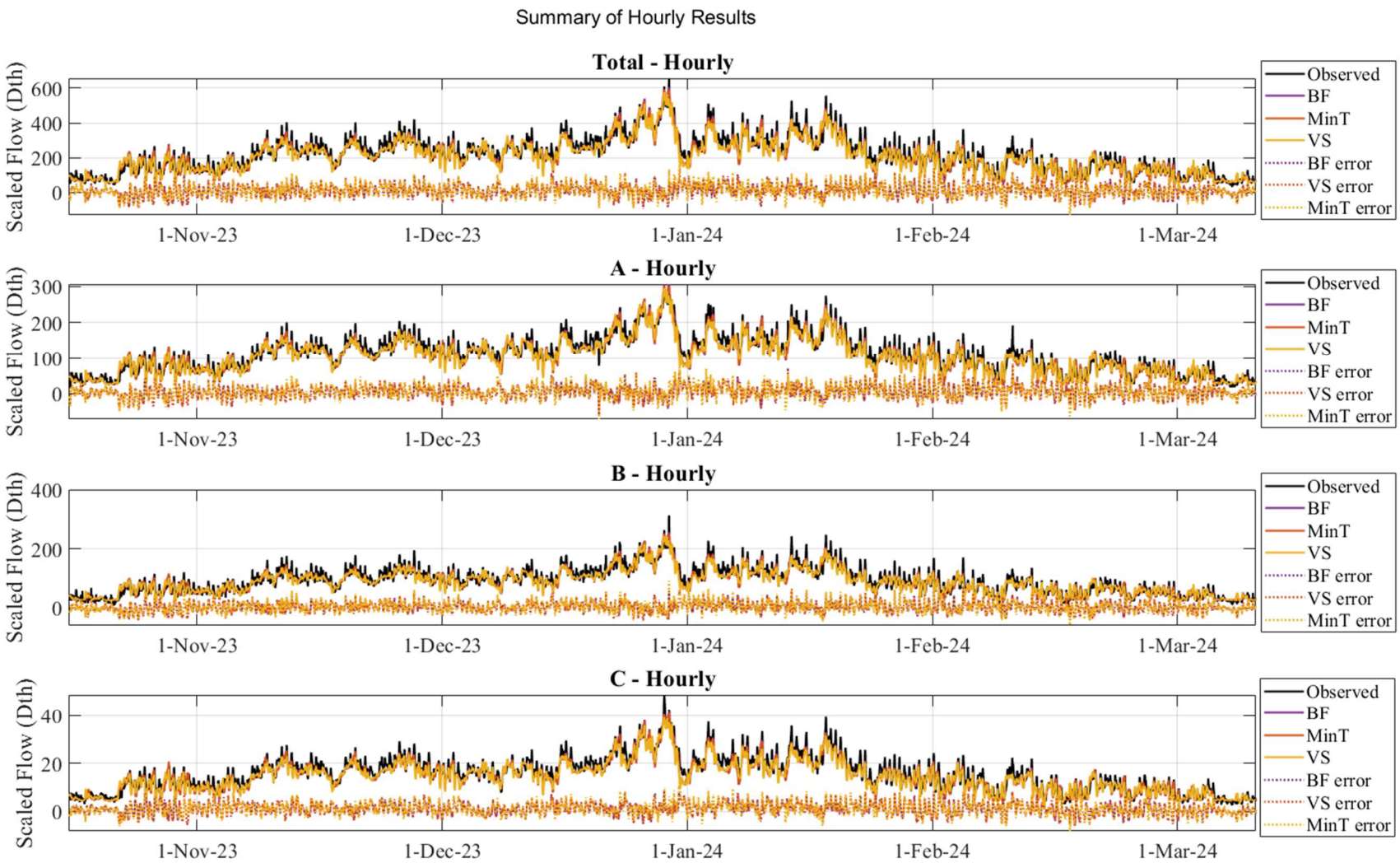

Figure 6.6 – Hourly reconciled forecasts compared to observed data, with the differences plotted below in corresponding dashed colors.

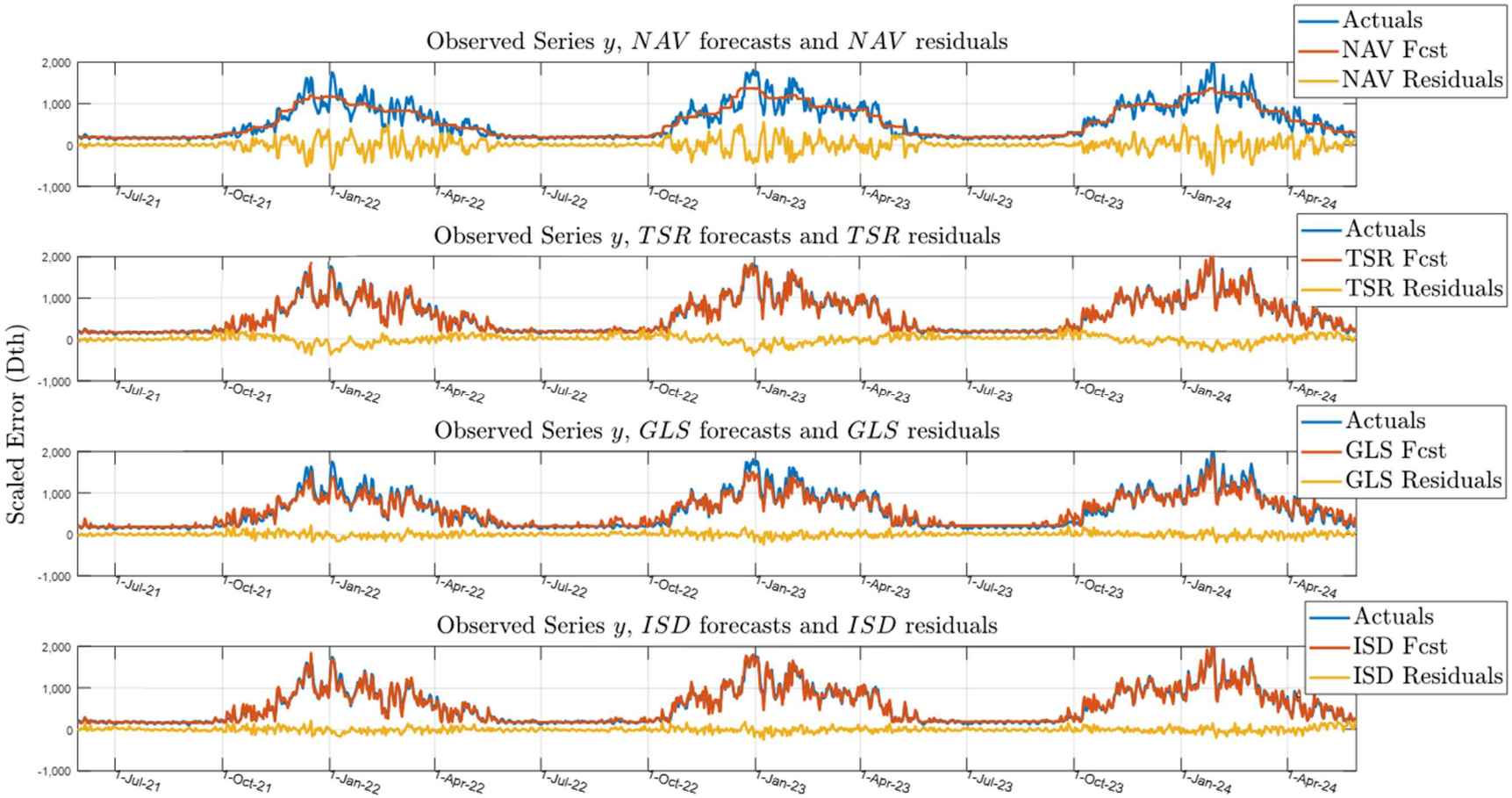

Figure 6.7 – Daily reconciled forecasts compared to observed data, with the differences plotted below in corresponding dashed colors.

Observing Figure 6.6 and Figure 6.7, the actuals are shown in black, BF in purple, MinT in orange, and VS in yellow. The corresponding dashed lines in the same colors indicate the differences between each method and the observed gas consumption. Gas practitioners are particularly interested in accurate forecasts during peak winter periods [6], [66], [80]. In Figure 6.6, the orange peaks stand out, indicating that MinT reconciliation effectively leveraged the correlation between exogenous weather variables and gas consumption in areas A, B, C, and $Total$. This is in part due to the MinT methods ability to model inter-hierarchy series interaction (via inclusion of off-diagonals of G). The VS method performed similarly but does not consider these interactions and does not perform as well on these peaks. Figure 6.7 shows daily estimates with a similar pattern of orange peaks, though not as consistently. This suggests that the smoother, lower-frequency daily series benefited from MinT reconciliation, but not as significantly as the hourly results in Figure 6.6. This also motivates further research in gas demand reconciliation and suggests that training specialized forecasting models for each hierarchical level could better leverage different exogenous correlations.

Given the inherent incoherence in the natural gas historical data, the results depicted in Figure 6.6 and Figure 6.7 show satisfactory performance when contrasting the coherent VS and MinT forecasts with the incoherent naïve and independently generated base forecasts. These findings underscore the applicability and effectiveness of MinT in improving forecasting accuracy in temporal and hierarchical contexts.

6.4 Cross-temporal Hierarchical Reconciliation Conclusion

We use cross-temporal forecast reconciliation methods tailored to the needs of local gas distribution. Most time series reconciliation efforts focus on a particular temporal or cross-sectional hierarchy, raising doubts about the effectiveness of such methods in the context of gas delivery [57]. We hypothesized that insights from single-dimensional approaches could apply to cross-temporal implementations with suitable aggregation structures. Cross-temporal reconciliation is made possible by organizing gas demand series into a hierarchical time series structure, which is then used to constrain forecasts across both measurement resolution (time) and operating area (space), ensuring coherence. We propose two reconciliation schemes to transform incoherent natural gas base forecasts into coherent sets of gas demand forecasts. A case study is carried out focusing on the natural gas demand across three operational regions, forecasting at various geographical levels, and analyzing both hourly and daily frequencies. Cross-temporal reconciliation using the MinT method results in a 10% improvement in hourly forecasts and a 3% improvement in daily forecasts compared to incoherent base forecasts. Additionally, MinT yields a 7% improvement in hourly forecasts and a 9% improvement in daily forecasts compared to coherent VS forecasts.

Future research on cross-temporal reconciliation of natural gas demands will focus on integrating demand forecasting with other energy systems such as electricity and heat generation to optimize energy management. Investigating the regulatory and policy implications within the hierarchical constraints set by government regulators is also a realistic application. Additionally, analyzing different customer segments and their consumption behaviors can lead to tailored forecasting approaches for each segment. These

ideas contribute to a more efficient gas distribution system while also providing significant financial incentives to local distribution companies.

The superior performance of MinT and VS methods over base forecasts is attributed to combining information from all levels of the natural gas demand hierarchy. Future directions include exploring error propagation based on individual gas demand levels rather than the entire hierarchy, as well as adjusting for the natural incoherence of gas data before reconciliation. In summary, MinT's capability to leverage hierarchical structures and ensure coherence across aggregation levels renders it superior for hierarchical time series forecasting and reconciliation in gas demand forecasting

CHAPTER 7

DISCUSSION, PRACTICAL IMPLICATIONS, AND FUTURE WORK

7.1 Summary of Contributions

This dissertation advances the field of natural gas demand modeling and hierarchical reconciliation by addressing the challenges associated with data incoherence in both historical and forecasted gas consumption datasets. Through a structured multi-method approach, this research introduces methodological innovations in disaggregation, hierarchical reconciliation, and forecast reconciliation, all aimed at improving the accuracy and reliability of natural gas modeling at multiple levels of aggregation.

One of the key contributions of this dissertation is the development of the Multi-Source Iterative Shifting Disaggregation (ISD) Algorithm, introduced in Chapter 4, which refines traditional approaches for disaggregating historical gas consumption data. The ISD algorithm was designed to allocate coarse gas consumption data to finer spatial and temporal resolutions more effectively than traditional interpolation methods. By incorporating multiple data sources and iterative shifting techniques, ISD captures localized consumption patterns more accurately, providing a more detailed representation of natural gas usage. Through empirical evaluation on real-world LDC data, ISD demonstrated superior performance in preserving the structure and variability of observed consumption patterns.

Chapter 5 introduces a structured framework for single-dimension hierarchical reconciliation, ensuring that recorded historical gas consumption aligns with known oper-

ational constraints across either spatial or temporal hierarchies. This reconciliation process directly enhances the quality of training datasets used for demand forecasting, leading to improved out-of-sample forecast accuracy. By eliminating structural inconsistencies in historical gas consumption data, single-dimension hierarchical reconciliation strengthens the predictive capacity of machine learning and statistical forecasting models, as demonstrated in the case studies.

Building upon the principles of single-dimension reconciliation, Chapter 6 introduces cross-temporal hierarchical forecast reconciliation, a novel framework that extends hierarchical reconciliation techniques to natural gas demand forecasts. Unlike Chapter 5, which focuses on ensuring consistency in recorded historical gas consumption data, Chapter 6 addresses the challenge of maintaining coherence across both spatial and temporal levels in future demand projections. A cross-temporal summation matrix is developed to integrate spatial and temporal constraints into a unified reconciliation framework. The empirical evaluation demonstrates that applying cross-temporal hierarchical forecast reconciliation to natural gas demand forecasts reduces overall prediction error, resulting in more reliable short-term and mid-term demand projections that can improve operational decision-making for LDCs.

Each of these contributions builds upon the prior methodologies, forming a comprehensive framework for improving natural gas demand estimation. The methods presented in this dissertation equip LDCs with enhanced analytical tools for resource allocation, system planning, and regulatory compliance. By addressing inconsistencies in both historical gas consumption and forecasted natural gas demand, this research advances the

accuracy and operational relevance of hierarchical reconciliation techniques in real-world gas distribution systems.

7.2 Method Summaries and Performance Evaluation

The following section introduces and evaluates three key methods developed in this dissertation: (1) Multi-Source Iterative Shifting Disaggregation (ISD), (2) Single-Dimension Hierarchical Reconciliation, and (3) Cross-Temporal Hierarchical Forecast Reconciliation. Each method contributes to enhancing gas demand forecasting by addressing challenges in temporal disaggregation, hierarchical coherence, and reconciliation across multiple dimensions.

7.2.1 Multi-Source Iterative Shifting Disaggregation (ISD) Algorithm

The ISD algorithm disaggregates multiple, nonuniformly sampled time series into a single high-frequency signal while maintaining coherence across aggregation levels. The algorithm operates iteratively, alternating between a prediction phase that applies multiple linear regression to estimate high-frequency values and an update phase that redistributes low-frequency observations to refine these estimates. Unlike existing disaggregation methods, ISD accommodates multiple overlapping input series, significantly improving accuracy. When applied to billing cycle data, ISD improves Weighted Mean Absolute Percentage Error (WMAPE) by 1.4–4.3%, while for residential consumption disaggregation, it achieves a 4.6–10.4% improvement.

7.2.2 Single-Dimension Hierarchical Reconciliation

Single-Dimension Hierarchical Reconciliation ensures internal consistency within either temporal or spatial hierarchies. Temporal reconciliation aligns hourly, daily, and

monthly forecasts to ensure coherence across different time levels. Spatial reconciliation enforces consistency across geographic regions by ensuring that aggregated consumption at smaller spatial scales sums correctly to the total system-level demand. The approach relies on a summation matrix to enforce these constraints, reducing inconsistencies and improving forecast reliability compared to traditional methods.

7.2.3 Cross-Temporal Hierarchical Forecast Reconciliation

Cross-Temporal Hierarchical Forecast Reconciliation extends single-dimension reconciliation by integrating both temporal and spatial constraints, ensuring consistency across all hierarchical levels simultaneously. This method maintains coherence between forecasts at different time frequencies while also aligning estimates across spatial operating regions. It improves hourly forecast accuracy by 10% and daily forecast accuracy by 3% compared to base forecasts. Compared to Variance Scaling (VS) reconciliation, Cross-Temporal Hierarchical Forecast Reconciliation using the Minimum Trace (MinT) approach improves accuracy by 7% for hourly forecasts and 9% for daily forecasts. These improvements highlight the effectiveness of hierarchical reconciliation in producing coherent and more accurate gas demand forecasts.

7.3 Future Work and Research Directions

Future research will focus on refining the methodologies introduced in this dissertation to further enhance the accuracy and coherence of gas demand forecasts. One key area of improvement involves enhancing the disaggregation process within the ISD algorithm by incorporating reconciliation techniques earlier in the workflow. Adjusting Step 1 of ISD to integrate hierarchical reconciliation could improve forecast accuracy while

maintaining coherence across all aggregation levels. Further investigation into the trade-offs between iterative-update disaggregation methods like ISD and optimal forecast reconciliation techniques will provide insight into improving both efficiency and accuracy. Since the ISD algorithm already introduces inter-series interaction through load shifting, future work may explore incorporating reconciliation-based load-shifting techniques in Step 3 to accelerate convergence and optimize forecast performance.

For hierarchical reconciliation, future studies will examine how enforcing coherence in training data affects model learning. The findings suggest that hierarchical inconsistencies, if left uncorrected, distort model coefficients and weaken true demand relationships. Single-dimension reconciliation effectively ensures coherence within either spatial or temporal hierarchies, but it does not address inconsistencies across time horizons. Future work should focus on extending reconciliation techniques to capture both cross-temporal and cross-spatial interactions, particularly for datasets with dynamically changing demand patterns.

Further research on cross-temporal hierarchical reconciliation will integrate gas demand forecasting with broader energy systems, such as electricity and heat generation, to optimize multi-energy management. Exploring the regulatory and policy implications of hierarchical constraints set by government agencies will also be a valuable direction, as aligning forecasts with regulatory requirements could enhance decision-making for local distribution companies. Investigating different customer segments and their consumption behaviors will allow for the development of tailored forecasting models that account for distinct usage patterns. The superior performance of MinT and VS reconciliation methods suggests that future studies should explore error propagation at individual gas demand

levels rather than applying a uniform reconciliation process across the entire hierarchy. Additional research into natural incoherence in gas data before reconciliation could further improve model performance. These efforts will contribute to a more efficient, accurate, and financially beneficial forecasting framework for local distribution companies.

7.4 Dissertation Conclusions

This dissertation introduced novel methods for improving natural gas demand forecasting through enhanced disaggregation, hierarchical reconciliation, and cross-temporal forecast integration. The challenges associated with nonuniformly sampled gas consumption data, hierarchical inconsistencies, and forecast incoherence were addressed using three key methodological advancements: the Multi-Source Iterative Shifting Disaggregation (ISD) algorithm, Single-Dimension Hierarchical Reconciliation, and Cross-Temporal Hierarchical Forecast Reconciliation. These approaches collectively improve forecast accuracy, maintain coherence across hierarchical levels, and provide a structured framework for integrating disparate gas consumption data sources.

The ISD algorithm demonstrated significant improvements in disaggregating multiple overlapping and nonuniformly sampled time series into a single high-frequency representation. By iteratively refining estimates through load shifting and regression modeling, ISD produced more accurate forecasts compared to traditional disaggregation techniques. The hierarchical reconciliation methods addressed inconsistencies within spatial and temporal hierarchies, ensuring that gas demand forecasts remained structurally sound across different levels of aggregation. Finally, cross-temporal reconciliation extended

these techniques by integrating spatial and temporal constraints simultaneously, leading to further improvements in forecast coherence and overall predictive performance.

The findings presented in this dissertation underscore the importance of maintaining data consistency in hierarchical forecasting systems. Forecast incoherence, if left unaddressed, can introduce systematic biases, distort demand relationships, and reduce the reliability of predictive models. By implementing reconciliation techniques at both the data and forecast levels, it is possible to mitigate these effects and produce forecasts that align more closely with actual gas consumption patterns. The results demonstrate that the proposed methods outperform traditional forecasting approaches, offering substantial gains in accuracy and reliability.

Beyond methodological advancements, this work contributes to practical forecasting applications for local distribution companies by enabling more informed decision-making, improving resource allocation, and reducing operational risks. The integration of hierarchical reconciliation techniques provides a scalable and adaptable solution for handling gas demand variability across different time horizons and geographic regions. As energy distribution systems continue to evolve, the approaches developed in this dissertation offer a robust foundation for future forecasting research, with potential applications extending beyond the natural gas industry to other hierarchical forecasting problems in energy management, finance, and infrastructure planning.

In summary, this dissertation advances the state of natural gas forecasting by providing a structured approach to data disaggregation, hierarchical reconciliation, and cross-temporal forecast integration. The findings highlight the necessity of coherence in

demand forecasting and demonstrate the effectiveness of hierarchical methods in producing more accurate and reliable predictions. These contributions lay the groundwork for future research aimed at further refining forecasting techniques and extending hierarchical reconciliation frameworks to broader energy and economic systems.

BIBLIOGRAPHY

[1] T. Hong, "Energy Forecasting: Past, Present, and Future," *Foresight Int. J. Appl. Forecast.*, vol. 32, pp. 43–48, 2013.

[2] L. Suganthi and A. A. Samuel, "Energy models for demand forecasting - A review," *Renew. Sustain. Energy Rev.*, vol. 16, no. 2, pp. 1223–1240, 2012, doi: 10.1016/j.rser.2011.08.014.

[3] P. E. Kaefer, "Transforming Analogous Time Series Data to Improve Natural Gas Demand Forecast Accuracy," Marquette University, 2015.

[4] F. Moauro and G. Savio, "Temporal disaggregation using multivariate structural time series models," *Econom. J.*, vol. 8, no. 2, pp. 214–234, 2005, doi: 10.1111/j.1368-423x.2005.00161.x.

[5] W. Wei and D. Stram, "Disaggregation of Time Series Models," *Wiley R. Stat. Soc.*, vol. 52, no. 3, pp. 453–467, 1990, [Online]. Available: https://www.jstor.org/stable/2345669.

[6] S. R. Vitullo, G. F. Corliss, M. Adya, F. Nourzad, and R. H. Brown, "An algorithm for disaggregating temporal natural gas consumption," *Can. Appl. Math. Q.*, vol. 21, no. 3, pp. 391–410, 2013.

[7] S. Askari, N. Montazerin, and M. H. F. Zarandi, "High-Frequency Modeling of Natural Gas Networks from Low-Frequency Nodal Meter Readings Using Time-Series Disaggregation," *IEEE Trans. Ind. Informatics*, vol. 12, no. 1, pp. 136–147, 2016, doi: 10.1109/TII.2015.2500883.

[8] S. L. Wickramasuriya, G. Athanasopoulos, and R. J. Hyndman, "Optimal Forecast Reconciliation for Hierarchical and Grouped Time Series Through Trace Minimization," *J. Am. Stat. Assoc.*, vol. 114, no. 526, pp. 804–819, Apr. 2019, doi: 10.1080/01621459.2018.1448825.

[9] J. G. Tamba *et al.*, "Forecasting natural gas: A literature survey," *Int. J. Energy Econ. Policy*, vol. 8, no. 3, pp. 216–249, 2018.

[10] American Petroleum Institute, "Impacts of the Oil and Natural Gas Industry on the US Economy in 2015," Washington, DC, 2017. [Online]. Available: https://www.api.org/~/media/Files/Policy/Jobs/Oil-and-Gas-2015-Economic-Impacts-Final-Cover-07-17-2017.pdf.

[11] F. J. Strum, *Trading Natural Gas: Cash, Futures, Options, and Swaps*, 1st ed. Tulsa, Oklahoma: PennWell Corporation, 1997.

[12] S. Mokhatab, W. Poe, and J. Mak, *Handbook of Natural Gas Transmission and Processing: Principles and Practices*, 4th ed. Houston, Texas: Gulf Professional Publishing, 2018.

[13] K. Costello, "Lost and Unaccounted-for Gas: Practices of State Utility Commissions," Silver Spring, 2013. [Online]. Available: https://pubs.naruc.org/pub/FA86BB52-AE3F-D8AC-B295-801BD6DC6435.

[14] "Natural Gas Glossary," *Pacific Gas & Electric*. https://www.pge.com/pipeline/library/doing_business/glossary/index.page (accessed Apr. 22, 2021).

[15] M. Kopalek, "October 2019 Monthly Energy Review," 2019. [Online]. Available: https://www.eia.gov/todayinenergy/detail.php?id=110.

[16] T. F. Golob, J. Torous, M. Bradley, D. Brownstone, S. S. Crane, and D. S. Bunch, "Commercial fleet demand for alternative-fuel vehicles in California," *Transp. Res. Part A Policy Pract.*, vol. 31, no. 3, pp. 219–233, 1997, doi: 10.1016/S0965-8564(96)00017-1.

[17] B. Soldo, "Forecasting natural gas consumption," *Appl. Energy*, vol. 92, pp. 26–37, Apr. 2012, doi: 10.1016/j.apenergy.2011.11.003.

[18] M. Fakoor, "Disaggregation: Inferring Daily Gas Flow from Billing Cycle Data," Marquette University, 2019.

[19] G. W. Emery and Q. Liu, "An Analysis of the Relationship between Electricity and Natural-Gas Futures Prices," *J. Futur. Mark.*, vol. 22, no. 2, pp. 95–122, 2002, doi: 10.1002/fut.2209.

[20] D. J. Sailor, J. N. Rosen, and J. R. Muñoz, "Natural gas consumption and climate:

A comprehensive set of predictive state-level models for the United States," *Energy*, vol. 23, no. 2, pp. 91–103, 1998, doi: 10.1016/S0360-5442(97)00073-X.

[21] H. G. Huntington, "Industrial natural gas consumption in the United States: An empirical model for evaluating future trends," *Energy Econ.*, vol. 29, no. 4, pp. 743–759, 2007, doi: 10.1016/j.eneco.2006.12.005.

[22] R. P. Timmer and P. J. Lamb, "Relations between temperature and residential natural gas consumption in the Central and Eastern United States," *J. Appl. Meteorol. Climatol.*, vol. 46, no. 11, pp. 1993–2013, 2007, doi: 10.1175/2007JAMC1552.1.

[23] J. Jager, *Climate and Energy Systems: A review of their interactions*, 1st ed. John Wiley & Sons, 1983.

[24] G. D. Merkel, R. J. Povinelli, and R. H. Brown, "Short-term Load Forecasting of Natural Gas with Deep Neural Network Regression," *Energies*, vol. 11, no. 8, 2018, doi: 10.3390/en11082008.

[25] Energy Information Administration, "Natural Gas Explained - Use of Natural Gas," *U.S. Energy Information Administration - Monthly Energy Reveiw*, 2020. https://www.eia.gov/energyexplained/natural-gas/use-of-natural-gas.php (accessed Jan. 06, 2021).

[26] J. G. Asbury, C. Maslowski, and R. O. Mueller, "Solar Availability for Winter Space Heating: An Analysis of SOLMET Data, 1953 to 1975.," *Science*, vol. 206, no. 4419, pp. 679–681, Nov. 1979, doi: 10.1126/science.206.4419.679.

[27] D. G. Baker, "Effect of Observation Time on Mean Temperature Estimation," *J. Appl. Meterology*, vol. 14, no. 4, pp. 471–476, 1975.

[28] M. Fakoor, G. F. Corliss, and R. H. Brown, "Prior Day Effect in Forecasting Daily Natural Gas Flow from Monthly Data," *IEEE Power Energy Soc. Gen. Meet.*, vol. 2018-Augus, 2018, doi: 10.1109/PESGM.2018.8586017.

[29] F. Taşpinar, N. Çelebi, and N. Tutkun, "Forecasting of daily natural gas consumption on regional basis in Turkey using various computational methods," *Energy Build.*, vol. 56, pp. 23–31, 2013, doi: 10.1016/j.enbuild.2012.10.023.

[30] R. Hribar, P. Potočnik, J. Šilc, and G. Papa, "A Comparison of Models for Forecasting the Residential Natural Gas Demand of an Urban Area," *Energy*, vol. 167, pp. 511–522, 2018, doi: 10.1016/j.energy.2018.10.175.

[31] D. J. Sailor and J. R. Muñoz, "Sensitivity of electricity and natural gas consumption to climate in the U.S.A. - Methodology and results for eight states," *Energy*, vol. 22, no. 10, pp. 987–998, 1997, doi: 10.1016/S0360-5442(97)00034-0.

[32] N. Wei *et al.*, "Short-term forecasting of natural gas consumption using factor selection algorithm and optimized support vector regression," *J. Energy Resour. Technol. Trans. ASME*, vol. 141, no. 3, pp. 1–10, 2019, doi: 10.1115/1.4041413.

[33] S. H. Mohr and G. M. Evans, "Long term forecasting of natural gas production," *Energy Policy*, vol. 39, no. 9, pp. 5550–5560, 2011, doi: 10.1016/j.enpol.2011.04.066.

[34] A. Valero and A. Valero, "Physical geonomics: Combining the exergy and Hubbert peak analysis for predicting mineral resources depletion," *Resour. Conserv. Recycl.*, vol. 54, no. 12, pp. 1074–1083, 2010, doi: 10.1016/j.resconrec.2010.02.010.

[35] S. Y. Kan, B. Chen, X. F. Wu, Z. M. Chen, and G. Q. Chen, "Natural gas overview for world economy: From primary supply to final demand via global supply chains," *Energy Policy*, vol. 124, no. September 2018, pp. 215–225, 2019, doi: 10.1016/j.enpol.2018.10.002.

[36] F. Furuoka, "Natural gas consumption and economic development in China and Japan: An empirical examination of the Asian context," *Renew. Sustain. Energy Rev.*, vol. 56, pp. 100–115, 2016, doi: 10.1016/j.rser.2015.11.038.

[37] A. M. Kiliç, "Turkey's natural gas necessity, consumption and future perspectives," *Energy Policy*, vol. 34, no. 14, pp. 1928–1934, 2006, doi: 10.1016/j.enpol.2005.02.004.

[38] R. Sari, B. T. Ewing, and U. Soytas, "The relationship between disaggregate energy consumption and industrial production in the United States: An ARDL approach," *Energy Econ.*, vol. 30, no. 5, pp. 2302–2313, 2008, doi: 10.1016/j.eneco.2007.10.002.

[39] H. Aras and N. Aras, "Forecasting Residential Natural Gas Demand," *Energy*

Sources, vol. 26, no. 5, pp. 463–472, Apr. 2004, doi: 10.1080/00908310490429740.

[40] E. F. Sánchez-Úbeda and A. Berzosa, "Modeling and forecasting industrial end-use natural gas consumption," *Energy Econ.*, vol. 29, no. 4, pp. 710–742, 2007, doi: 10.1016/j.eneco.2007.01.015.

[41] P. Musilek, E. Pelikán, T. Brabec, and M. Šimůnek, "Recurrent neural network based gating for natural gas load prediction system," *IEEE Int. Conf. Neural Networks - Conf. Proc.*, pp. 3736–3741, 2006, doi: 10.1109/ijcnn.2006.247390.

[42] A. S.D.I. and O. M.A., "Temporal Disaggregation of Time Series Data: A Non-Parametric Approach," *Math. Theory Model.*, vol. 2, no. 5, 2012.

[43] W.-S. Chan, "Disaggregation of annual time-series data to quarterly figures: A comparative study," *J. Forecast.*, vol. 12, no. 8, pp. 677–688, Dec. 1993, doi: 10.1002/for.3980120816.

[44] R. Shumway and D. Stoffer, *Time Series Analysis and Its Applications With R Examples*, 4th ed. Springer, 2017.

[45] G. C. Chow and A. Lin, "Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series," *Rev. Econ. Stat.*, vol. 53, no. 4, p. 372, Nov. 1971, doi: 10.2307/1928739.

[46] V. M. Guerrero and F. H. Nieto, "Temporal and contemporaneous disaggregation of multiple economic time series," *Off. J. Spanish Soc. Stat. Oper. Res.*, vol. 8, no. 2, pp. 459–489, Dec. 1999, doi: 10.1007/BF02595880.

[47] C. Sax and P. Steiner, "Temporal Disaggregation of Time Series," *R Joural*, vol. 5, no. 53389, pp. 80–87, 2013.

[48] O. Sprangers, W. Wadman, S. Schelter, and M. de Rijke, "Hierarchical forecasting at scale," *Int. J. Forecast.*, vol. 40, no. 4, pp. 1689–1700, 2024, doi: 10.1016/j.ijforecast.2024.02.006.

[49] R. J. Hyndman, A. J. Lee, and E. Wang, "Fast computation of reconciled forecasts for hierarchical and grouped time series," *Comput. Stat. Data Anal.*, vol. 97, pp. 16–32, 2016, doi: 10.1016/j.csda.2015.11.007.

[50] N. Kourentzes and G. Athanasopoulos, “Cross-temporal coherent forecasts for Australian tourism,” *Ann. Tour. Res.*, vol. 75, no. 2018, pp. 393–409, 2019, doi: 10.1016/j.annals.2019.02.001.

[51] E. Spiliotis, M. Abolghasemi, R. J. Hyndman, F. Petropoulos, and V. Assimakopoulos, “Hierarchical forecast reconciliation with machine learning,” *Appl. Soft Comput.*, vol. 112, pp. 1–25, 2021, doi: 10.1016/j.asoc.2021.107756.

[52] G. Athanasopoulos, R. A. Ahmed, and R. J. Hyndman, “Hierarchical forecasts for Australian domestic tourism,” *Int. J. Forecast.*, vol. 25, no. 1, pp. 146–166, 2009, doi: 10.1016/j.ijforecast.2008.07.004.

[53] G. Athanasopoulos, R. J. Hyndman, N. Kourentzes, and F. Petropoulos, “Forecasting with temporal hierarchies,” *Eur. J. Oper. Res.*, vol. 262, no. 1, pp. 60–74, 2017, doi: 10.1016/j.ejor.2017.02.046.

[54] C. W. Gross and J. E. Sohl, “Disaggregation methods to expedite product line forecasting,” *J. Forecast.*, vol. 9, no. 3, pp. 233–254, May 1990, doi: 10.1002/for.3980090304.

[55] S. Ben Taieb, J. W. Taylor, and R. J. Hyndman, “Coherent probabilistic forecasts for hierarchical time series,” *34th Int. Conf. Mach. Learn. ICML 2017*, vol. 7, pp. 5143–5155, 2017.

[56] L. Corrado and B. Fingleton, “Where is the economics in spatial econometrics?,” *J. Reg. Sci.*, vol. 52, no. 2, pp. 210–239, 2012, doi: 10.1111/j.1467-9787.2011.00726.x.

[57] G. Athanasopoulos, R. J. Hyndman, N. Kourentzes, and A. Panagiotelis, “Forecast reconciliation: A review,” *Int. J. Forecast.*, vol. 40, no. 2, pp. 430–456, Apr. 2024, doi: 10.1016/j.ijforecast.2023.10.010.

[58] R. J. Hyndman, R. A. Ahmed, G. Athanasopoulos, and H. L. Shang, “Optimal combination forecasts for hierarchical time series,” *Comput. Stat. Data Anal.*, vol. 55, no. 9, pp. 2579–2589, 2011, doi: 10.1016/j.csda.2011.03.006.

[59] M. Sobhani, A. Campbell, S. Sangamwar, C. Li, and T. Hong, “Combining weather stations for electric load forecasting,” *Energies*, vol. 12, no. 8, 2019, doi: 10.3390/en12081510.

[60] V. Majazi Dalfard, M. Nazari Asli, S. M. Asadzadeh, S. M. Sajjadi, and A. Nazari-Shirkouhi, "A Mathematical Modeling for Incorporating Energy Price Hikes into Total Natural Gas Consumption Forecasting," *Appl. Math. Model.*, vol. 37, no. 8, pp. 5664–5679, 2013, doi: 10.1016/j.apm.2012.11.012.

[61] S. R. Vitullo, R. H. Brown, G. F. Corliss, and B. M. Marx, "Mathematical Models for Natural Gas Forecasting," *Can. Appl. Math Q.*, vol. 17, no. 4, pp. 1–13, 2010, [Online]. Available: papers2://publication/uuid/5E8CDA33-908F-4B19-8F90-02A304882841.

[62] F. Theodosiou and N. Kourentzes, "Forecasting with Deep Temporal Hierarchies," *SSRN Electron. J.*, vol. 92, pp. 26–37, Apr. 2021, doi: 10.2139/ssrn.3918315.

[63] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 3rd ed. OTexts: Melbourne, Australia, 2021.

[64] T. van Erven and J. Cugliari, "Game-Theoretically Optimal reconciliation of contemporaneous hierarchical time series forecasts," *Lect. Notes Stat.*, vol. 217, pp. 297–317, 2015, doi: 10.1007/978-3-319-18732-7_15.

[65] L. Bai and P. Pinson, "Distributed reconciliation in day-ahead wind power forecasting," *Energies*, vol. 12, no. 6, pp. 1–19, 2019, doi: 10.3390/en12061112.

[66] B. Gaweł and A. Paliński, "Global and Local Approaches for Forecasting of Long-Term Natural Gas Consumption in Poland Based on Hierarchical Short Time Series," *Energies*, vol. 17, no. 2, p. 347, Jan. 2024, doi: 10.3390/en17020347.

[67] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: a review and evaluation," *IEEE Trans. Power Syst.*, vol. 16, no. 1, pp. 44–55, 2001, doi: 10.1109/59.910780.

[68] A. Panagiotelis, G. Athanasopoulos, P. Gamakumara, and R. J. Hyndman, "Forecast reconciliation: A geometric view with new insights on bias correction," *Int. J. Forecast.*, vol. 37, no. 1, pp. 343–359, 2021, doi: 10.1016/j.ijforecast.2020.06.004.

[69] T. Di Fonzo and D. Girolimetto, "Enhancements in cross-temporal forecast reconciliation, with an application to solar irradiance forecasts," *arXiv Prepr. arXiv2209.07146*, 2022, [Online]. Available: http://arxiv.org/abs/2209.07146.

[70] E. De Alba, “Disaggregation and forecasting: Bayesian analysis,” *J. Bus. Econ. Stat.*, vol. 6, no. 2, pp. 197–206, 1988, doi: 10.1080/07350015.1988.10509654.

[71] S. L. Wickramasuriya, G. Athanasopoulos, and R. J. Hyndman, “Forecasting hierarchical and grouped time series through trace minimization,” 15, 2015. [Online]. Available: https://econpapers.repec.org/RePEc:msh:ebswps:2015-15.

[72] H. Akaike, “Fitting Autoregressive Models for Prediction,” *Ann. Inst. Stat. Math.*, vol. 21, no. 1, pp. 243–247, 1969, doi: 10.1007/BF02532251.

[73] S. Vitullo, “Disaggregating Time Series Data for Energy Consumption by Aggregate and Individual Customer,” Marquette University, 2011.

[74] R. Weron and F. Ziel, “Electricity price forecasting,” in *Routledge Handbook of Energy Economics*, 2019, pp. 506–521, doi: 10.4324/9781315459653-36.

[75] H. Rob J and G. Athanasopoulos, *Forecasting: Principles & Practice*, 2nd ed. Melbourne, Australia: OTexts, 2018.

[76] R. H. Brown, S. R. Vitullo, G. F. Corliss, M. Adya, P. E. Kaefer, and R. J. Povinelli, “Detrending daily natural gas consumption series to improve short-term forecasts,” in *2015 IEEE Power & Energy Society General Meeting*, Jul. 2015, pp. 1–5, doi: 10.1109/PESGM.2015.7286138.

[77] R. J. Hyndman, G. Athanasopoulos, and H. L. Shang, “hts : An R Package for Forecasting Hierarchical or Grouped Time Series,” *Stat. Softw.*, 2014, [Online]. Available: https://cran.r-project.org/web/packages/hts/vignettes/hts.pdf.

[78] D. Girolimetto, G. Athanasopoulos, T. Di Fonzo, and R. J. Hyndman, “Cross-temporal probabilistic forecast reconciliation: Methodological and practical issues,” *Int. J. Forecast.*, vol. 40, no. 3, pp. 1134–1151, 2024, doi: 10.1016/j.ijforecast.2023.10.003.

[79] R. J. Hyndman, “Another Look at Forecast-Accuracy Metrics For Intermittent Demand,” *Int. J. Energy Stat.*, vol. 04, no. 02, pp. 43–47, Jun. 2006, [Online]. Available: http://www.worldscientific.com/doi/abs/10.1142/S2335680416500083.

[80] C. O. Quinn, R. J. Povinelli, and G. F. Corliss, “Alarm Forecasting in Natural Gas Pipelines,” Marquette University, 2020.