Singularity | Brattle

Where matters: Integrating deliverability into voluntary clean energy market boundaries

AUTHORED BY

Singularity Energy

Gregory Miller, Ph.D.

Gailin Pease

Wenbo Shi, Ph.D.

WITH CONTRIBUTIONS FROM

The Brattle Group

Long Lam, Ph.D.

Kathleen Spees, Ph.D.

Jadon Grove

Ivy Yang

August 2023

Prepared for and supported by
Google LLC

Authorship notes: Singularity authored the report, including developing the framework, evaluation criteria, case study, and modeled evidence. The Brattle Group contributed subject-matter expertise on electricity systems and markets, and survey of procurement boundaries in standards and policies. The authors wish to acknowledge members of the Google climate team for thoughtful comments and inputs on earlier drafts of this report.

Table of Contents

1. Introduction

In the past several decades, many programs have emerged to incentivize voluntary action on grid decarbonization, including the market-based Scope 2 emissions accounting guidance in the Greenhouse Gas (GHG) Protocol, the U.S. EPA's Green Power Partnership program, and the RE100 program. These provide guidance on how actors, such as corporations, cities, and universities, should structure their clean energy procurement to meet the program goals and hopefully accelerate the decarbonization of the electric grid.

One of the most influential of these frameworks has been the GHG Protocol's Guidance on Scope 2 accounting. In particular, the Protocol's market-based accounting framework is designed to allow GHG reporters to accurately reflect the GHG footprint of the electricity or electricity attributes they contractually purchase, rather than based on the electricity they physically consume, as is reflected in the complementary location-based accounting method.[1]

In 2022, the GHG Protocol began a multi-year revision process for the Scope 2 guidance. A major question under consideration for the update is how the procurement should be accounted for based on when (temporal matching) and where (spatial matching) the associated generation occurs relative to the reporter's electricity consumption.

The topic of more granular temporal matching of clean energy procurement with load has received much attention in recent years,[2] but the criteria for appropriate spatial boundaries for deliverability of procured clean energy have remained relatively unexplored, especially in the context of accounting standards that purport to reflect emissions of clean energy purchasing decisions. This paper begins to address this gap.

In the context of the GHG Protocol's market-based Scope 2 accounting framework, the primary question is: how can standards define energy procurement boundaries to accurately reflect the effect of clean energy purchasing decisions on the emissions intensity of a reporter's consumed electricity? At one end of the spectrum, there would be no limitations on where clean energy could be sourced from, such that a company based in New York could procure clean energy from a solar farm in China. At the other end of the spectrum, it would be required that all procured energy be directly delivered to the buyer, such that a company could only procure from an onsite or direct line-connected generator. However, as with the design of reliable and efficient electricity market boundaries, an effective market-based accounting boundary should strike a balance between the needs of the grid and the needs of participants.

Paper Overview

- **Section 2** provides an overview of how voluntary market boundaries are defined in various contexts today;
- **Section 3** explores the context of existing physical, market, and regulatory boundaries in three key geograpahies: the United States, the European Union, and China;
- **Section 4** introduces the boundary evaluation framework;
- **Sections 5–7** explain the structural relevance, physical deliverability, and practicality considerations and introduce the evaluation criteria for each;
- **Section 8** illustrates how this framework could be applied in practice to the GHG Protocol Scope 2 market-based accounting framework;
- **Section 9** concludes with a discussion of how the framework can be broadly applied when designing voluntary market boundaries in any context.

1 See Section 2A for complete definitions of these GHG accounting terms.

2 For example, see the UN Sustainable Energy for All 24/7 Carbon Free Energy Compact at https://gocarbonfree247.com/

Except in the case of emissions offsets, the concept of physical deliverability is central to credible market-based procurement and claims. Many companies claim to be powered by 100% renewable energy, but most claims are not substantiated by physical deliverability. Physical deliverability refers to the ability of generated electricity to be delivered through the grid to the buyer.[3] Although physical deliverability is not the only salient criterion, some level of physical deliverability in voluntary clean energy market design is important. In voluntary markets, there is currently no mechanism for ensuring that clean power supply resources are individually or simultaneously deliverable to the subset of customers that procure them to reduce their carbon footprint.

This paper introduces a framework that designers of voluntary clean energy markets can consider when defining the spatial boundaries that determine from where buyers can source clean energy. This framework enables designers to balance considerations of structural relevance, physical deliverability, and practicality depending on their goals. While this framework is broadly applicable to various types of voluntary markets and incentive programs, we include a case study that examines the current voluntary market boundaries defined in the GHG Protocol Scope 2 market-based accounting guidance. We find that the market boundaries, as interpreted today, perform poorly against most of the metrics in our framework, suggesting they should be updated.

3 In some cases, identifying whether generated electricity is deliverable to a buyer is straightforward. For example, it is physically impossible for electricity generated in California to be delivered to a buyer in Hawaii because these two locations belong to separate electricity grids that are separated by thousands of miles of ocean with no transmission lines between them. In other cases, answering this question is not as clear-cut: Is electricity produced in Northern California deliverable to a customer in Southern California? What about electricity generated from neighboring Arizona, where electricity is often imported from? Or from Wyoming, which is part of the same synchronous grid but is 1,000 miles away? We explore several approaches for answering these questions in the Physical Deliverability section of this paper.

2. Procurement Boundaries in Voluntary Standards and Government Policies

Key Takeaways

- Energy procurement boundaries are defined for a range of voluntary programs and government mandates.
- The GHG Protocol Scope 2 Guidance is one of the most widely referenced standards for voluntary energy procurement and GHG accounting, but focuses more on energy attribute market boundaries than electricity markets and lacks clear guidance on implementation.

Many regulatory programs, such as renewable portfolio standards and resource adequacy programs, may offer examples of boundaries that balance deliverability and practicality. Geographic boundary definitions are used in several ways in the context of policies, standards, and regulations related to tracking, reporting, and reducing GHG emissions from the power sector, such as:

- In Scope 2 market-based GHG emissions reporting, electricity consumers use geographic boundaries for two purposes: (1) for estimating the GHG emissions associated with any electricity purchases for which the reporter does not have a clearly identifiable supply resource; and (2) for determining the market region from which clean energy can be sourced and considered to allow for GHG-free electricity consumption. The second purpose is not only relevant to Scope 2 GHG emissions reporting, but also to other voluntary energy procurement programs, including RE100 and 24/7 carbon-free energy.
- In implementing renewable and clean energy policy requirements (such as renewable portfolio standards), governments adopt a geographic boundary for determining the eligibility of resources and for creating fungible market instruments that responsible entities can use to demonstrate compliance with the standard.
- In GHG cap or cap-and-trade policies, emissions are tracked and are tradeable within a particular geographic boundary.

The selection of geographic boundaries has substantial implications for the accuracy of resulting Scope 2 accounting allocations, the credibility of the consumers' clean energy supply claims, and the effectiveness of government policy. To the extent that revised boundary definitions can improve the accuracy of GHG emissions accounting, this can help drive higher impact.

Boundaries in the Greenhouse Gas Protocol

The GHG Protocol, a global standard for GHG emissions accounting, is referenced and used in a number of mandatory and voluntary disclosure standards and frameworks (see Table 1 above). For issues related to the concept of geographic boundaries for Scope 2 emissions reporting, these standards and frameworks broadly defer to guidance from the GHG Protocol. The Protocol provides two methods for calculating emissions associated with consumed electricity. In the location-based method, the reporting company calculates its emissions by relying on the emissions intensity of the electricity grid where operations and energy consumption occur. The **market-based method**, the focus of this paper, estimates emissions associated with energy procurement by accounting for contractual instruments such as clean energy attribute certificates and power purchase agreements (PPA), as well as an electricity supplier's specific emissions rates.

Under the market-based accounting method, boundaries are intended to credibly link the power supply (both fossil and clean) from specific supply resources to specific consumers. Reporters are advised that regional transmission organizations (RTOs), power pools, or balancing areas are typical electricity market boundaries, though boundaries for certificates may be broader, and to check and use boundaries established by regulatory authorities, certification bodies, and/or issuing bodies.[4] However, if those boundaries are unclear, a single market may be one country (*e.g.*, the U.S.) or multiple countries that have "a set of common market rules and a regional connection" (*e.g.*, the European Union).[5] The guidance encourages the use of boundaries within which there is a consistent method to track and retire certificates and to calculate residual mixes, and that certificates be "sourced from regions reasonably linked to the reporting entity's electricity consumption."[6]

Less clear in the Protocol, however, is how to implement its guidance on market boundaries. Although electricity markets play a fundamental role in determining which generation resources are dispatched on consumers' behalf, the Protocol does not offer clear nor specific guidance on which types of market structures (*e.g.*, electricity markets versus attribute markets; market regions, zones, or nodes) should be used. The Protocol does reference the use of North American energy attribute tracking systems, several of which are defined by boundaries of wholesale electricity markets (*e.g.*, PJM), while others are not (*e.g.*, M-RETS).

The Protocol provides a quality hierarchy for contractual instruments, with energy attribute certificates having the highest precision, followed by energy contracts, supplier/utility emissions rates, residual mix, and other national or sub-national grid-average emissions factors.[7] The Protocol stops short of presenting a preferred hierarchy of boundaries for contractual instruments, or of providing guidance on what boundary should apply to the instruments. Rather, in describing the quality hierarchy, the Protocol

4 GHG Protocol Scope 2 Guidance, WRI (2014): pages 64–65.

5 GHG Protocol Scope 2 Guidance, WRI (2014): page 65.

6 *Ibid.*

7 GHG Protocol Scope 2 Guidance, WRI (2014): page 48.

Table 1: Geographic Definitions Across Accounting Frameworks, Tools, and Policy

Framework/Source	Role	Geographic Boundaries Guidance & Rationale
Green-e (by Center for Resource Solutions)	Certifies EACs and publishes residual mix emission factors for the U.S.	Generators must be located in or connected to certificate region, which is not necessarily dependent on physical interconnection.[8] Allows more geographically constrained boundaries. Residual mix calculations are based on eGRID subregions.
Association of Issuing Bodies	EU required reporting of emissions related to power supply	Uses EU national boundaries
Clean and Renewable Energy Standards: U.S.: States' renewable and clean energy portfolio standards Europe: Clean Energy Package China: Green Energy Certificates	Establish and enforce policy requirements to increase the share of renewable and clean energy supply Market instruments created under these programs are frequently used by voluntary Scope 2 reporters (as above)	Regional boundaries used to establish qualified supply resources, sometimes limited to jurisdictional boundaries but often clean resources outside the home jurisdiction are considered as qualified (*e.g.*, in the same electricity market area).
GHG Protocol: Revised Version Scope 2 Guidance Document	Standard for corporate accounting and reporting of emissions	Companies should use existing certificate market boundaries if available, otherwise use political or regulatory boundaries, or source from regions "reasonably linked" to a reporter's electricity consumption.[9]
Voluntary Disclosure Initiatives: Carbon Disclosure Project (CDP) The Climate Registry Science-Based Targets Initiative	Voluntary emissions reporting programs used by companies to standardize disclosure	All three programs reference the GHG Protocol. The Climate Registry refers to eGRID and IEA emissions factors as the baseline source.
RE100	Corporate renewable energy purchasing initiative	Individual countries are distinct markets for renewable electricity, except the U.S. and Canada are considered a single market, and all of Europe is considered a single market except countries that are not AIB members or grid-connected to the European market.[10]
Government Standards: SEC Draft Climate Disclosure Rule EU Corporate Sustainability Reporting Directive China's Guidance for Enterprise ESG Disclosure	Government mandated reporting for publicly traded companies	U.S.: Proposed rule broadly defers to the GHGP, cites the EPA's methodology for calculating indirect emissions (which is derivative of the Protocol), and links to the same set of emissions factors from the EPA, which refers to eGRID.[11,12] EU: ESRS E1: Climate Change points to the GHG Protocol and GRI 305 (which itself references the GHG Protocol). China: Guidance for Enterprise ESG Disclosure points to national GHG reporting guideline GB/T 32150-2015, which references the GHG Protocol.
GHG Reduction Policies (*e.g.*, GHG tax, cap-and-trade, rate-base limits): U.S.: Western Climate Initiative, Regional Greenhouse Gas Initiative Europe: EU Emissions Trading System China: Emissions Trading System	Create government-enforceable limits on GHG emissions to achieve policy goals over time	Regional boundaries consistent with home jurisdiction only, as consistent with the scope of authorities to direct GHG-emitting fossil resources to limit their emissions. Emissions deemed to be embedded in electricity imports may be held accountable under applicable limits.

Sources and Notes: Association of Issuing Bodies, CDP, EU Corporate Sustainability Reporting Directive, GHG Protocol: Revised Version, GHG Protocol: Scope 2 Guidance Document, Green-e, Guidance for Enterprise ESG Disclosure, RE100, Science-Based Target Initiative, SEC Draft Climate Rules, The Climate Registry, The World Bank's State and Trends of Carbon Pricing 2022

8 Center for Resource Solutions, Green-e Framework for Renewable Energy Certification, Version 1.0.

9 GHG Protocol Scope 2 Guidance, WRI (2014): pages 64–65.

10 RE100 Technical Criteria, Appendix B.

11 The Enhancement and Standardization of Climate-Related Disclosures for Investors – Proposed Rule, Securities and Exchange Commission (2022): page 148.

12 *Id.*, page 193.

aims to inspire higher accuracy while providing flexibility, acknowledging that actual implementation depends on local market options and company-specific conditions. However, this flexibility may also introduce a high degree of subjectivity in determining market boundaries, which can potentially make comparison across emissions inventories challenging and increase the risk of double-counting.

The lack of explicit rationale, incentivization, or clear guidance on market boundaries means that many reporting companies in practice often rely on the most readily available data, simplest methodology, or even options highlighted as illustrative examples. For example, even though the GHG Protocol references eGRID emissions rate data as one of the lowest-precision options in its current market-based data hierarchy, eGRID data is widely used in reporting today because it is readily available across the U.S.[13] In addition, companies do not have to report what boundaries are used, making comparison and aggregation across companies impossible. Until there is more specificity or transparency on geographic boundary and definition, comparability of claims across companies will continue to be challenging.

Boundaries in the CDP

Some disclosure standards and frameworks also echo the Protocol's flexible approach to selecting market boundaries. For example, the Carbon Disclosure Project (CDP), an international nonprofit organization that supports companies, cities, states, and regions across the world to disclose their environmental information, refers to the GHG Protocol for calculating emissions.[14,15] With respect to claiming renewable clean energy use, CDP stipulates "companies must source renewable electricity from within the boundary of the market in which they are consuming the electricity," and defines "market boundary" as an area in which:

- The laws and regulatory framework governing the electricity sector are consistent between the areas of production and consumption.
- There is a physical interconnection between the point of generation and the point of consumption of renewable electricity. When interconnection happens across different grids, there must be a level of system-wide coordination between such grids.
- The countries' utilities/energy suppliers recognize each other's energy sourcing instruments and have a system in place to prevent double counting of claims.

However, the CDP considers the U.S. and Canada to be a single market, even though there are many different state and province-level regulatory frameworks that govern the electricity sector in these two countries. This overly broad and imprecise boundary definition means that a power plant located in British Columbia and a customer located in Virginia are considered to be in the same electricity market, which is not the case in reality.

---

13 GHG Protocol Scope 2 Guidance, WRI (2014): page 48.

14 CDP Disclosure Insight Action, What we do, You can't manage what you don't measure. (2023).

15 CDP Disclosure Insight Action, CDP Technical Note: Accounting of Scope 2 emissions: CDP Climate Change Questionnaire, Version 9 (Rev. March 11, 2022).

Other boundary options

Many non-voluntary market-based frameworks, such as state renewable portfolio standards (RPS) in the U.S., have also had to grapple with the question of the locational eligibility for procured resources, and may offer practical and credible examples for voluntary programs. To be eligible for California's RPS program, for instance, facilities must be located in California, or if not, must demonstrate a first point of interconnection to a California balancing authority along with several other requirements.[16]

In summary, standards and policies today offer examples of boundary guidance that often leave room for interpretation, but indicate the importance of structural relevance, physical deliverability, and practicality. In the next section, we provide a background on how the electricity system is structured and how it operates.

---

16 Green, Lynette, Christina Crume. Renewables Portfolio Standard Eligibility Guidebook, Ninth Edition. California Energy Commission, Publication Number: CEC-300-2016-006-ED9-CMFREV. (2017): page 33.

3. Electricity Systems Background

Key Takeaways

- Electricity systems have a physical dimension, a regulatory dimension, and a market dimension, all of which affect how electricity can be transacted.
- "Synchronous grids" or "interconnections" are the broadest boundaries of the physical system, representing a single, synchronized machine consisting of generators, transmission and distribution, and loads.
- Regulation affecting the planning, reliability, and operation of the grid is often split between national or supranational entities and smaller regions.
- Electricity markets vary in size and structure but share several common features including instantaneously balancing supply and demand, ensuring deliverability of supply, efficiently pricing energy, and ensuring the financial accountability of transactions.

A clean energy procurement framework should be anchored in the principles of electricity markets and operations. This section highlights the foundation of electricity market boundaries and related boundaries for carbon accounting policies in the United States, Europe, and China, which in 2021 were collectively responsible for more than 50% of total global CO_2 emissions from electricity and heat.[17]

The Physical Grid

The electricity grid consists of three main components: generation, transmission, and distribution. Electricity generated from a power plant is delivered to customers through the high-voltage transmission system and the low-voltage distribution system. A country may have one or multiple synchronous alternating current power grids, which can be tied together with direct current transmission lines. There are three major synchronous grids that make up the contiguous U.S. electric system: the Eastern Interconnection, the Western Interconnection, and the Texas Interconnection.[18] Similarly, an interconnected electricity system consisting of five major synchronous areas serves the majority of the European continent. In China, the State Grid serves about 80% of China's total electricity consumption, with most of the remainder served by the smaller China Southern Power Grid.

---

17 From Ritchie H. *et al.*, CO_2 and Greenhouse Gas Emissions, Our World in Data (August 2022).

18 North America consists of nine major synchronous grids: three in Mexico (the Baja California, Baja California Sur, and National Interconnections), the two Alaskan Interconnections, the Texas Interconnection, the Western and Eastern Interconnections that span the U.S. and Canada, and the Quebec Interconnection in Canada.

Within an interconnection region, all of the electric utilities are electrically tied together during normal system conditions and operate at a certain synchronized frequency (*e.g.*, 60 Hertz in the U.S.). Power within each synchronous area is distributed through alternative current (AC) lines, and loads are balanced within the subnetwork, together creating the physical structure that enables electricity to be delivered and traded across large areas.

Regulatory Authority

There is a wide spectrum of regulatory structures governing the electricity sector across the globe, and jurisdictional boundaries that are politically or administratively similar may not reflect similar grid structures.

In some regions, one single national entity is responsible for all aspects of the power system, including short-term and long-term system planning, operations, and reliability standards. The electrical systems of many island nations operate under this model.

In other regions, the regulatory authority is shared between national and sub-national regulators. For example, the U.S. federal government has oversight of wholesale electricity sales in most of the country through its constitutional authority over interstate commerce, and has delegated this power to the U.S. Federal Energy Regulatory Commission (FERC). FERC regulates independent system operators (ISOs) and regional transmission organizations (RTOs), which facilitate open access to transmission, independently operate the transmission system as a free and fair market, and promote competition for electricity generation within their jurisdiction. Some of FERC's authority is delegated to the North American Electric Reliability Corporation (NERC), which ensures and promotes the reliability of the bulk power system on a regional basis by issuing mandatory reliability standards. Regulatory matters that do not pertain to FERC's authority are generally under the purview of state regulatory agencies, whose specific mandate and authority are prescribed and approved by state legislatures and vary from one another.[19] For example, public utilities commissions (PUCs) typically oversee resource adequacy requirements for utilities serving their states (if states do not participate in an RTO), regulate retail electricity rates and utility programs, and implement environmental policy as appropriate.[20]

By contrast, in China, the central government through the National Development and Reform Commission (NRDC) maintains major administrative and policy-making responsibilities, including long-term electricity system planning, reliability standards, and interregional regulation, as well as ratemaking and environmental enforcement. Meanwhile, the provincial governments are responsible for short-term planning, small project approval, and system operations.

---

19 States that are not part of an RTO are responsible for their own system planning and operations.

20 For instance, while the Pennsylvania PUC is responsible for implementing the state's Alternative Energy Portfolio Standard, the PA Department of Environmental Protection is responsible for issuing rules related to the state's participation in the regional GHG cap-and-trade system.

The European Union (EU) reflects yet another model, where the regulation of electricity markets emphasizes collaboration among European bodies and national governments, regulators, and transmission system operators (TSOs). The executive branch of the EU, the European Commission, also has the power to write laws governing electricity markets, which are then passed by the European Council and European Parliament (together making up the legislative branch). At the national level, each national energy regulator can establish and amend rules related to the country's electricity market, and the market operator then develops more detailed market rules.[21, 22]

Electricity Markets

Within an electricity market, at any given moment, supply must match demand, a task carried out by grid operators in each balancing authority area. The balancing authority maintains load-resource balance within its area. Further, electricity delivery depends on transmission capacity constraints. In competitive markets, market operators determine the electricity price based on an economically optimized model, where energy resources are secured and dispatched based on their short-run marginal cost, starting with the least-cost resources. This sets prices at the individual node (generally a substation) of the transmission system (nodal pricing), reflecting transmission constraints. This is the case in the U.S. electricity markets, and electricity markets in China are moving toward a nodal pricing system as well.

Prices can also be set at the zonal level, as is the case in European electricity markets. In this approach, electricity is traded across "bidding zones" based on available transmission capacities. The European bidding zone borders are shaped by country borders.[23] Because transmission lines within a particular zone are assumed to have unlimited capacity, wholesale electricity prices are uniform within that zone.

It is important to note that individual customers do not have any means of expressing their own willingness to pay for GHG abatement or lower-emissions-rate delivered power in the wholesale market. Additionally, while all energy across the regional system is simultaneously deliverable in a broad sense (and transmission planning is designed to build system upgrades in response to regionally identified needs of simultaneous deliverability), this system does not guarantee deliverability of any specific clean energy supplier to any specific customer. For these reasons, market instruments (*e.g.*, RECs and

---

21 Local TSOs act as the parties taking final responsibility for managing grid imbalances; the role of the electricity market operator is separate and distinct from the role of the TSO. This is unlike the U.S., where the role of balancing authority or transmission system operator was transferred to the ISO/RTO entity that also runs the independent markets. The balancing authority integrates resource plans ahead of time and maintains load/import/export/generation balance within a balancing authority area whose boundaries are defined by the transmission networks.

22 While individual member countries are responsible for establishing their reliability standards and assessing their national resource adequacy, the EU is moving toward a more comprehensive planning framework for the entire block through initiatives such as the European Resource Adequacy Assessment which monitors and evaluates the resource adequacy of the EU-wide power system. See ENTSO-E, European Resource Adequacy Assessment 2022 Edition.

23 A bidding zone refers to "the largest geographical area within which market participants are able to exchange energy without capacity allocation," and it may span multiple countries, as is the case for Germany and Luxembourg. See Commission Regulation (EU) No 543/2013.

power purchase agreements) exist such that customers can express their willingness to pay to reduce emissions and set their own goals to support clean energy development in these markets.

Implications for boundary design

In summary, electricity markets exist to determine what resources are dispatched to which customers and at what price to ensure that supply and demand are reliably matched at any given moment. Though the details of how electricity markets are regulated and operate vary by jurisdiction, they share several key features, as described in Table 2 below.

Table 2: Shared characteristics of wholesale electricity markets

Feature	Description
Instantaneous Balance of Supply and Demand	Supply and demand must be balanced at all points in time down to the timescale of fractions of a second (though market metering and pricing are typically only distinguished down to 1-minute, 5-minute, or 15-minute granularity)
Deliverability	Supply must be simultaneously deliverable to demand across the transmission lines and subject to prevailing transmission limits, considering that these binding limits may differ in every dispatch interval (though there tend to be strong patterns)
Prices	Prices reflect the incremental cost of dispatching power at all times and places, including a geographically granular expression of these costs in consideration of transmission constraints
Financial Accountability	Customers are financially accountable to pay energy supply resources dispatched on their behalf, again as calculated on a temporally and geographically granular basis

Despite these commonalities, the boundaries of what constitutes an electricity market are not consistently defined:

- The broadest and most ambiguous definition of a market boundary can be considered as a region across which there is a material level of energy trade among players. At a minimum this means that there must be at least some transmission interconnection between the regions and an agreed practice for scheduling power flows and making payments. However, these areas may not have a centralized market operator or structure, nor a common jurisdictional authority. Under this definition, market boundaries may overlap and not be mutually exclusive.
- The most commonly used definition of an electricity market boundary is a region that utilizes a single centralized market operator (such as a balancing authority or independent system operator) and a common set of rules for market pricing and resource dispatch, subject to transmission constraints.
- Within each market, smaller zones or subregions may exist that align with individual transmission system owners or with areas that experience chronic transmission congestion. These zones are utilized for transmission planning and routing market settlement costs through the responsible utilities.

- Some markets (*e.g.*, in the U.S. and China, but not Europe) additionally include “nodes” that represent individual power injection and withdrawal points on the electricity grid (generally corresponding with a substation) that are separately metered and settled against market prices.

This variability contributes to the challenge of defining clear market boundaries for voluntary energy procurement standards, and emphasizes the importance of clearer guidance by standards-setting bodies.

4. Introducing the Energy Procurement Boundary Framework

Given the important role that standards play in incentivizing voluntary clean energy procurement and influencing specific procurement decisions, it is critical that the procurement boundaries be intentionally and thoughtfully designed to achieve the intended outcomes. If grounded in the physical, economic, and regulatory realities of how electricity markets and electric power systems function, carefully designed procurement rules can encourage voluntary purchasers to make impactful and sustainable contributions to the decarbonization of the electric grid. If poorly designed, these rules may hinder voluntary participation, misalign incentives with the needs of the grid, or marginalize the role of the voluntary market in grid decarbonization.

We have identified four potential approaches that voluntary market program designers can use to define the locational eligibility of procured energy. The **boundary-based approach** involves defining a static geographic boundary within which both the buyer and generator must be located in order for the resource to be eligible. Many existing programs, including the GHG Protocol Scope 2 guidance, undertake this approach, and it is the primary focus of this paper. In an **instrument-based approach**, buyers can procure only certain types of market instruments that intrinsically reflect a deliverability aspect (such as physical PPAs). The **modeling-based approach** establishes eligibility by conducting power flow analysis to determine whether power generated by a resource can physically flow through the grid to the buyer. Finally, a **rule-based approach** defines a set of rules that can be used to establish the eligibility of a resource, and may simply refer to the existing eligibility rules used by the electricity market or regulatory context in which the buyer is located. The rule-based approach may be a combination of several other approaches, such as defining a static geographic boundary within which resources are automatically eligible coupled with a requirement to procure transmission capacity rights to ensure deliverability if the resource is located outside of that boundary. Each of these approaches has different advantages and disadvantages that can be explored using the framework introduced in this paper.

To help standards setters intentionally and thoughtfully design these voluntary market boundaries, the framework introduced in this paper is structured around three primary considerations:

1. **Structural Relevance:** The boundaries should align voluntary actions with the needs of the grid and with existing market and regulatory structures, which are increasingly organized around decarbonizing the grid in addition to their traditional roles of ensuring the safe, reliable, and affordable delivery of electricity.

2. **Physical Deliverability:** The boundaries should seek to ensure that clean energy procured anywhere within the boundary is deliverable to a buyer anywhere within the same boundary in order to improve credibility and avoid greenwashing.

3. **Practicality:** The boundaries should consider the abilities of all voluntary actors to ensure an appropriate level of access and participation.

Each of these considerations can be evaluated using a set of criteria introduced in Table 3 below and explained in detail in the following sections.

Table 3: Overview of the voluntary energy procurement boundary evaluation framework

Framework considerations	Motivations	Evaluation Criteria	Question
Structural Relevance	Grid reliability, efficient market operation, alignment with non-voluntary market demand signals	Electricity market alignment	Are voluntary buyers participating in their local electricity markets? Does the boundary allow voluntary buyers to procure from generally the same pool of resources that other market participants can procure from?
		Regulatory alignment	Does the boundary support reliable grid operations? Does the boundary support the effective functioning of emissions regulations (*e.g.,* cap-and-trade)?
Physical Deliverability	Credible claims, alignment with demand signal sent by physical consumption, alignment of clean energy capacity with load growth	Synchronous grid boundaries	Does the voluntary market boundary exist within one synchronous grid, or rather allow for procurement across multiple synchronous grids?
		Transmission congestion	Is the uniform deliverability of new generation within a boundary limited by transmission congestion?
		Delivery distance	How large is the market boundary relative to the average distance power flows travel from generators within the boundary?
		Expected deliverability	What fraction of generators within a boundary deliver some amount of electricity to load in the same boundary?
		Regional connectivity	Do boundaries reflect regions with relatively high delivery between generators and loads?
Practicality	Broad participation	Achievability	Can participants achieve the program goals by only procuring resources from within the boundary? If not, do buyers have any pathways to expanding access?
		Understandability	Is the boundary definition conceptually simple enough and explained sufficiently clearly for voluntary buyers of all abilities to understand and act upon?
		Universality	Is the boundary definition consistent across all geographies where the standard will be used?
		Stability	Are the boundaries stable over time to reduce the risk of "stranded" voluntary purchases?

The importance of considering program goals

Because each program, market, or standard has different goals and priorities, it is important to recognize that there may be no universally or objectively "best" market boundary. For any one program, the boundary approach that best optimizes for a set of intended outcomes is not necessarily well suited for other contexts.

Prior to applying the energy procurement boundary framework proposed in this paper, program designers should identify the goals and outcomes that they are seeking to achieve to determine the relative importance of each of the criteria. For example, if a standard seeks to maximize the credibility of claims,[24] then an approach that prioritizes physical deliverability may be most suitable for that program. Or, if the program is focused on the emissions offset value of clean energy investments, physical deliverability may be less important. However, effective program design will likely involve balancing all three considerations to some extent.

Program designers should also weigh the tradeoffs among the criteria and their broader impact on the program. For example, a more practical approach may lead to greater impact if it encourages greater participation, but too relaxed of an approach may risk each participant's actions having little impact. Likewise, an approach that encourages physical deliverability may help increase the impact of any one action, but too strict of an approach may risk limiting participation and thus aggregate impact.

24 For example, the Federal Trade Commission's green marketing guides: https://www.ftc.gov/legal-library/browse/rules/green-guides.

5. Consideration One: Structural Relevance

Key Takeaways

- As voluntary energy buyers deploy an increasing volume of clean energy, their impact on the dynamics of electricity markets and regulatory outcomes becomes more pronounced. If this voluntary clean energy deployment continues to be divorced from grid dynamics, it could have major unintended impacts on system reliability and costs.
- The institutions which govern the delivery of power are increasingly supporting grid decarbonization through market and regulatory design. The efficacy of these efforts may be undermined if a large portion of the market engages in transactions outside of the boundaries of these policies.
- Electricity market alignment assesses the alignment with market participation rules and cost allocation.
- Regulatory alignment assesses the alignment with regulatory programs to advance decarbonization and ensure reliability.

Background and Justification

Considering structural relevance involves identifying how well a voluntary market boundary aligns with the structure and needs of the market and regulatory contexts within which a clean energy buyer is located. Aligning the voluntary procurement boundaries with market and regulatory boundaries is important because it matches the physical demand signal of a consumer (which helps determine market prices and supply response) with their contractual demand signal. This is important because it can:

- Help ensure that system reliability needs are met,
- Help promote the efficient operation of markets,
- Limit the risk of cost shifting to other grid users, and
- Align voluntary clean energy demand with other demand from other market participants (*e.g.*, from utilities) to strengthen the aggregate demand signal for clean energy.

As an example, procuring generation in Spain to match load in Germany can increase system costs and decrease value compared to procurement between areas that are more highly interconnected, as illustrated in a 2022 report by the Energy Track and Trace Initiative.[25]

The purpose of common electricity market features, outlined in Table 2 above, is to provide a signal for the efficiency and effectiveness of grid operations. This is relevant to a voluntary procurement boundary for the following reasons:

- Existing electricity market boundaries often define which generators can supply demand in a given region;
- Existing regulatory constructs, such as resource adequacy, are designed to ensure that there is enough generation capacity available to meet peak demand within a region, and therefore often define which generators qualify based on physical delivery to the reliability area;
- Traditional electricity suppliers are subject to rules or physical realities that limit where they can procure generation from; and
- Voluntary clean energy procurement can interact with existing regulatory structures, like cap-and-trade, and so the impact of the procurement may be distorted or implementation may inherit complexity if boundaries are not aligned.

Evaluation Criteria

To help standards setters evaluate the structural relevance of a proposed voluntary market boundary, we have identified two primary evaluation criteria: electricity market alignment and regulatory alignment. These criteria are summarized in Table 4 and described in more detail below.

Table 4: Summary of structural relevance evaluation criteria

Evaluation Criteria	Characteristics of more structurally relevant boundaries	Characteristics of less structurally relevant boundaries
Electricity market alignment	The boundary encourages voluntary buyer participation in their local electricity markets, and allows voluntary buyers to procure from generally the same pool of resources that other market participants can procure from.	Voluntary buyers are subject to much stricter or much looser procurement boundaries than other participants in their market, and thus may encourage procurement that may weaken demand signals, shift costs, or undermine wholesale market price signals.
Regulatory alignment	The boundary supports the effective functioning of emission regulations (*e.g.,* cap-and-trade) and supports the reliable operation of the grid.	The boundary allows voluntary procurement that weakens the effectiveness of regulatory programs in decarbonizing the grid or maintaining grid reliability.

25 Energy Track and Trace, System Benefits of Granular Certification (July 2022). Available at https://energytrackandtrace.com/wp-content/uploads/2022/08/ETT_System-benefits-paper_1_0.pdf.

Electricity market alignment: Comparing the voluntary procurement boundaries to the boundaries of existing electricity markets is a useful first step to understand whether the voluntary boundaries are smaller, larger, or overlapping with these boundaries. While it may not be necessary to perfectly align voluntary procurement boundaries with existing market boundaries, it is nonetheless important to consider how the incentives created by the voluntary boundary will interact with the functioning of electricity markets. To consider the potential impacts of boundary (mis-)alignment, one can assess whether it undermines the price signals that the grid operator sends in order to operate the grid efficiently and reliably, how it might shift costs to other consumers, whether it aligns voluntary demand with other demand, and how it might impact wholesale market prices.

When evaluating electricity market alignment, it can be helpful to first ask whether the boundary subjects voluntary buyers to stricter or looser requirements than other market participants. Requiring voluntary buyers to follow much stricter rules than other market participants may be unnecessary and limit participation from voluntary buyers in electricity markets. However, setting much looser rules for voluntary buyers has the potential to distort or erode markets and regulations.

Regulatory alignment: It is also important to consider how a voluntary boundary definition may interact with the goals of regulations such as resource adequacy, renewable portfolio standards (RPS), and cap-and-trade programs. Such programs often have their own deliverability and resource eligibility rules that can serve as a useful starting point for defining reasonable voluntary boundaries. Resource adequacy eligibility definitions in particular can serve as useful examples, since their criteria are often designed to be practical, work within the context of existing electricity markets, and serve the physical needs of the grid. For example, some programs use market boundaries to establish eligibility, but also allow participants to procure capacity from outside these boundaries if they also procure transmission capacity rights to ensure the capacity is deliverable. Additionally, if voluntary boundaries are not aligned with those of programs like RPS and cap-and-trade, it may erode the effectiveness or increase the cost of complying with such programs.

6. Consideration Two: Physical Deliverability

Key Takeaways

- Physical deliverability is directly linked to the credibility of clean energy matching claims by emphasizing the underlying grid dynamics. Improving alignment with physical deliverability can more accurately reflect when an energy buyer is consuming clean energy and align load growth with additional clean energy capacity and grid development.
- Deliverability can be determined using less precise fungible deliverability metrics, or more robust, flow-based deliverability metrics, such as delivery distance, expected deliverability, and regional connectivity.
- Recently published research has shown that generated electricity is physically delivered within a relatively small radius (50 miles). However, smaller amounts of this generated electricity may be *deliverable* to loads much further away.
- Because of the range of region sizes within a defined set of boundaries (*e.g.*, some balancing areas are city-sized while others span thousands of miles), certain regions may reflect physical deliverability better than others within a single boundary type (*e.g.*, balancing areas).

Background and Justification

While structural relevance reflects institutional features of the underlying grid, physical deliverability reflects the physical dynamics of the grid. **Physical deliverability** refers to the ability of the electricity generated at a specific power plant to reach, or be "delivered" to, a specific consumer or purchaser based on the flow of power through the transmission and distribution network that connects the two. Physical deliverability is important because it:

- Enhances the accuracy of claims and helps avoid any appearance of greenwashing,
- Strengthens the aggregate demand signal for clean energy by aligning purchasing-driven demand with consumption-driven demand, and
- Ensures that clean energy capacity growth matches load growth.[26]

---

26 In addition to these factors, recent research suggests that deliverability may be important to help minimize the consequential emissions impact of new load growth. Ricks *et al.,* "Minimizing emissions from grid-based hydrogen production in the United States." *Environmental Research Letters* (2023).

The deliverability of generated electricity can be defined according to two different but complementary conceptual models of how physical power systems operate: **fungible deliverability** and **flow-based deliverability**. Although the fungible and flow-based frameworks use different conceptual definitions of deliverability, they are both grounded in power systems principles. However, fungible metrics are best thought of as simpler, qualitative approximations of deliverability, while flow-based metrics are more precise quantitative measures of deliverability.

Fungible and Flow-based Deliverability

Flow-based deliverability metrics trace the flows of power[27] through the transmission and distribution network from a specific generator to a specific load. These metrics are more robust and precise, but require using more sophisticated analytical techniques. When used to evaluate deliverability, power flow tracing can answer questions such as how far, and in what direction, electricity travels from a specific generator, and which generators deliver electricity to a consumer.

Fungible deliverability metrics are less precise generalizations that evaluate the ability of the consumer's load to be balanced by generation in another part of the grid, and treats all generators within a region as equally deliverable to all loads in a region. This framework recognizes that, while power from a specific generator may not flow to a specific load, all of the generators and loads together support the overall balance and functioning of the system, subject to congestion or bottlenecks. These metrics are more straightforward to evaluate but their results are subject to more exceptions and caveats and are less definitive for determining deliverability.

Physical deliverability analysis can be used to answer several related, but distinct questions. Determining the **deliverability** of generation involves answering whether *it is possible for any* electricity generated by a generator to be delivered to a load. In contrast, determining the **delivery** of generation involves answering *how much* of the electricity generated by a generator is actually delivered to a load. For example, electricity from a specific generator may be deliverable to a buyer even if only a small portion of the generated electricity is actually delivered, or if the electricity is only delivered at certain times of day.

It is also important to note that evaluating delivery and deliverability may require slightly different analyses depending on the specific context. Illustrative examples of different evaluation use cases include:

27 Power flow modeling is fundamental to power systems engineering and flow tracing was independently introduced by two groups of researchers in the mid-1990s as a solution for allocating transmission costs to individual generators and loads in restructured markets. Since then, hundreds of papers have been written on the topic, refining the theory and expanding the application of power flow tracing. A 2013 paper by Khan and Agnihotri titled "A Comprehensive Review of Embedded Transmission Pricing Methods Based on Power Flow Tracing Techniques" provides one early review of this research.

- A user buying a PPA might want a guarantee that a significant amount of that power is delivered to their load. They could evaluate this using project-specific *delivery* modeling that estimates the power delivered from their proposed PPA to their load over the course of the year.
- A voluntary market program designer might be interested in identifying market boundaries that support credible procurement claims. They could define a credible claim as procurement from a generator that delivers at least some of its power to the buyer's load over the course of the average year, since this indicates that the power is *deliverable* (*i.e.*, it is not only theoretically possible but practically possible for power to be delivered from the generator to the load). Once this boundary is established, they could evaluate the *expected deliverability*, or what percentage of load/generator pairs in a given boundary satisfies the criteria.
- A different designer might prioritize boundaries with high *regional connectivity*, which generally align with the most common delivery patterns on the grid.

As shown in these examples, physical deliverability is a tool that can support a variety of analysis goals. Across each of these approaches, there is a shared commitment to accurately reflect underlying grid dynamics, while highlighting the features most important to a given user.

Evaluation Criteria

To help standards setters evaluate the physical deliverability of a proposed voluntary market boundary, we have identified two fungible deliverability metrics and three flow-based deliverability metrics. At a high level, these five criteria seek to answer the following questions:

- **Synchronous grid boundaries:** Does the voluntary market boundary exist within one synchronous grid, or rather allow for procurement across multiple synchronous grids?
- **Transmission congestion:** Is the uniform deliverability of generation within a boundary limited by transmission congestion?
- **Delivery distance:** How large is the market boundary relative to the average distance power flows travel from generators within the boundary?
- **Expected deliverability:** What fraction of generators within a boundary deliver some amount of electricity to load in the same boundary?
- **Regional connectivity:** Do boundaries reflect regions with relatively high delivery between generators and loads?

It is important to note that none of these metrics by itself defines deliverability; rather they each provide a different perspective of deliverability and should be considered together. These criteria are summarized in Table 5 and described in more detail below.

Table 5: Examples of how to evaluate boundaries using the physical deliverability criteria

Type of metric	Evaluation criteria	Characteristics of more physically deliverable boundaries	Characteristics of less physically deliverable boundaries
Fungible (less precise qualitative generalization)	Synchronous grid boundaries	All energy procured by a buyer comes from within the same synchronous grid	A buyer is able to procure energy generated in a different synchronous grid from the one they are located in
	Transmission congestion	When procuring energy from a newly built generator, there are no transmission constraints between the generator and the buyer's load.	There are significant transmission constraints between a buyer and generation from a newly built generator.
Flow-based (more precise quantitative metric)	Delivery distance	Region size is similar to the median distance that electricity travels from generators to loads in a region	Region is much smaller or larger than delivery distances
	Expected deliverability	A high percentage of loads receive some amount of electricity from every generator in the region	Most loads in a region receive no generation from most generators in a region
	Regional connectivity	High modularity compared to optimal	Low modularity compared to optimal

In the following sections on fungible and flow-based deliverability, we describe how this range of metrics can serve the evaluation goals above. The fungible deliverability criteria can be a good starting point for evaluating boundaries, but for reasons explained below are not definitive and should not be used by themselves. The flow-based deliverability metrics can more definitively establish deliverability where it is left unclear by the fungible metrics. However, although power flow tracing is a well-established analytical approach, its application in the context of voluntary clean energy markets is still relatively nascent, and thus there is no single "best" flow-based deliverability metric that has- yet emerged. Thus, we propose the use of multiple recently developed flow-based deliverability metrics that illustrate the utility of deliverability to policy makers and electricity customers.[28] The evaluation of physical delivery therefore parallels the overarching three criteria presented in this report, where different users may use their specific priorities to provide relative weights to each criteria. As users and researchers continue to explore delivery and use delivery-derived metrics to evaluate electricity grids and boundaries, some metrics may prove to be more useful than others.

---

28 Pease, Gailin *et al.* "Tracking physical delivery of electricity from generators to loads using power flow tracing." Preprint at doi.org/10.5281/zenodo.8117207 (2023).

Fungible deliverability criteria

Synchronous grid boundaries: The largest boundary within which electricity could be uniformly deliverable is a synchronous grid, also known as an interconnection. A synchronous grid is essentially a single large machine: within its boundaries, all generators connected to the alternating current (AC) grid spin in sync with each other at the same frequency (hence "synchronous"), together supporting the stability of the grid. Synchronous grids may be connected to each other via DC interties that allow for a limited exchange of power, but they otherwise operate as physically independent power systems.

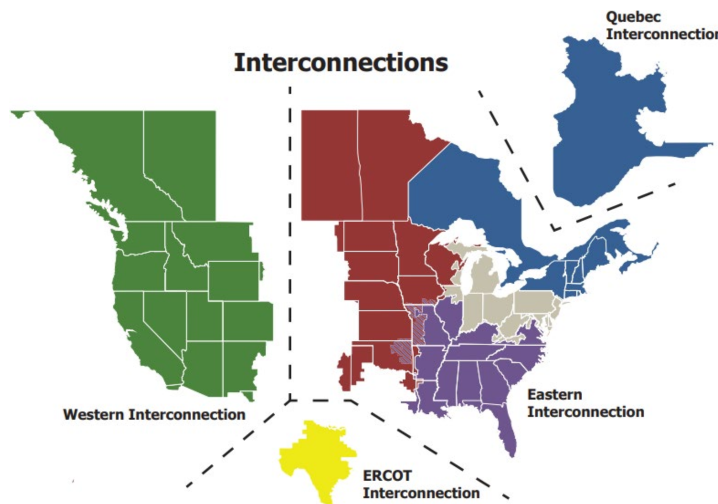

Figure 1. *This map shows four of the nine major synchronous grids (interconnections) in North America (excluding Alaska and Mexico). Evaluating boundaries using the physical grid structure metric involves examining whether the voluntary market boundaries span across multiple synchronous grids or not.*[29]

This straightforward evaluation criterion thus involves comparing the voluntary market boundaries to synchronous grid boundaries to ensure that a voluntary market does not span multiple synchronous grids. Generally, electricity from one synchronous grid is not physically deliverable to another. In some cases, this is clear cut: California is separated from Hawaii by thousands of miles of ocean, so a boundary that allows a buyer in Hawaii to procure energy from California is physically impossible. In other cases, the answer may not be as straightforward, for example in the case of two synchronous grids that exchange a large volume of electricity with each other via DC interties. This is an example of how this metric may not always provide a definitive answer about deliverability, and would need to be supplemented with flow-based deliverability analysis.

29 Image from the North American Reliability Council. https://www.nerc.com/AboutNERC/keyplayers/PublishingImages/NERC%20Interconnections.pdf

Transmission congestion: Even within a single synchronous grid, electricity may not be uniformly deliverable if transmission congestion exists. Transmission congestion occurs when a transmission line is operating at or near its maximum capacity, and is unable to support all requests for transmission services across the line.[30] Unlike traffic congestion, transmission congestion does not mean that power is not flowing across the line (in fact a large amount is flowing), but rather that new or additional "traffic" cannot travel across the line. This distinction has important implications for the relevance of congestion as a metric for evaluating deliverability: even with congestion, a large amount of existing generation may be deliverable across congested zones, but additional generation may not be. Thus, congestion is more helpful for evaluating the physical deliverability of *new* generation, or the deliverability of existing generation to new loads.

To evaluate transmission congestion, some grid operators define congestion zones, in which case the voluntary market boundaries can be easily compared. In wholesale electricity markets where locational marginal prices (LMPs) are used, such as in the United States, these prices usually include some congestion component since congestion imposes costs on the efficient operation of the electricity market. Whenever LMPs are uniform across a region, this can indicate a lack of congestion, and thus uniformly deliverable electricity across the region. Whenever prices are substantially different in a region, this suggests that each cluster of prices may be a separate congested zone. However, it is important to understand what is driving price differences, as LMPs can also contain components besides congestion (such as costs of energy and losses).

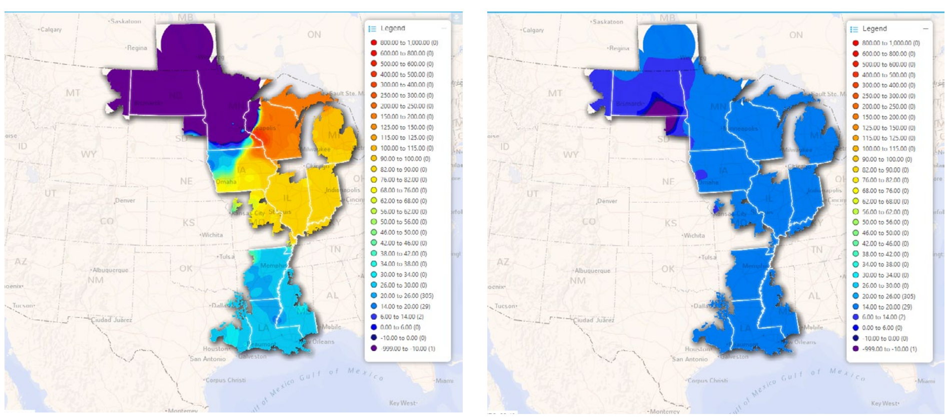

Figure 2. *These maps show LMP prices in MISO at two different times.[31] On the left, substantially different LMP prices in the North, Central, and South parts of the region indicate congestion between the regions. However, at another time, on the right, prices are relatively uniform across all of MISO, indicating little congestion. Using LMPs to approximate deliverability requires evaluating these prices over time to understand general trends.*

30 Lawrence Berkeley National Laboratory, "Electricity Transmission Congestion Costs: A Review of Recent Reports" (2003). https://www.energy.gov/oe/articles/electricity-transmission-congestion-costs-review-recent-reports

31 https://api.misoenergy.org/MISORTWD/lmpcontourmap.html

One challenging aspect of congestion-based evaluation is that congestion is not always static across time, as illustrated in Figure 2. A line that is congested at one time of day may not be congested at another time of day, and a new transmission line being built may structurally change the dynamics of congestion. It is only through evaluating congestion trends over time that one can fully understand how congestion might affect the deliverability of generation and the appropriateness of a voluntary procurement boundary.[32] This fungible deliverability metric can be useful to approximate deliverability between two locations, but flow-based deliverability metrics may be necessary to definitively determine deliverability.

Flow-based deliverability criteria

To determine flow-based *deliverability*, power flow tracing is used to trace generation from each generator on a power grid to each load on that grid. We call this **bilateral delivery**, since it describes the flow-based delivery between a specific generator-load pair on the grid at a specific time. Bilateral delivery can be quantified from either the consumer's upstream perspective, or the generator's downstream perspective. **Upstream delivery** describes what percentage of the total load delivered to each node originates from the generator, and can be used to answer "Where is the electricity I consume coming from?" Alternatively, **downstream delivery** describes what percentage of the total power generated by the generator is consumed at each node, and can be used to answer the question "Where is the electricity I generate going?"

32 One resource in the U.S. is the triennial National Electric Transmission Congestion Study (now called the National Transmission Needs Study) published by the U.S. Department of Energy: https://www.energy.gov/oe/national-electric-transmission-congestion-study

Highlight: Tracing Bilateral Delivery

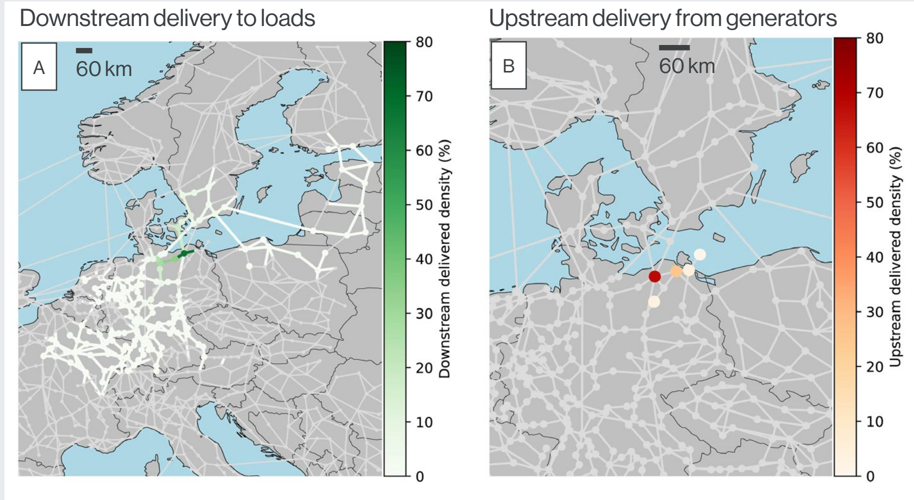

Figure 3. *This figure illustrates bilateral delivery at a specific time, and shows how it can be traced downstream from a single generator to many loads (A), or upstream from a single load to many generators (B). Due to the direction of power flows, electricity may not be delivered symmetrically: in panel A, we see that power flows primarily to the North and Southwest of the generator. In panel B, we see what percent of the electricity delivered to the load originated at each generator.*[33]

Flow-based deliverability can be evaluated on both transmission and distribution grids, wherever grid topology and time-varying generation and load are known. Because of constraints on publicly available data, this may require either using sensitive data owned by grid operators or using synthetic data that approximates the real grid behavior. Importantly, open energy system models that can enable more widespread access to flow-based deliverability analysis are becoming increasingly available for more global geographies through initiatives like PyPSA Meets Earth.[34] To illustrate these metrics in practice, we highlight key results from a new academic paper titled "Tracking physical delivery of electricity from generators to loads using power flow tracing," which uses open-source, synthetic grid models to explore flow-based deliverability metrics in more depth.[35]

Evaluating deliverability for an entire region across a longer timespan requires aggregating these bilateral deliverability metrics. This introduces some design choices about how the features should be aggregated. In this section, we present three options for geographic boundary-level metrics based on aggregating bilateral delivery analysis in different ways.

---

33 Data and figures adapted from Pease *et al.* (2023).

34 https://pypsa-meets-earth.github.io/

35 Pease, Gailin *et al.*, "Tracking physical delivery of electricity from generators to loads using power flow tracing." Preprint at doi.org/10.5281/zenodo.8117207 (2023). Specifically, the paper uses Python for Power Systems Analysis (PyPSA) with a clustered, publicly available model of the European grid (PyPSA-Eur) as well as a synthetic version of the U.S. grid developed by Breakthrough Energy.

Because each aggregated metric evaluates a different facet of the underlying delivery relationships, practitioners interested in evaluating how boundaries reflect physical delivery may wish to consider multiple metrics.[36]

Flow-based Deliverability Concepts

Deliverability evaluation is context dependent, depending on whether one i s evaluating the deliverability between a specific generator-load pair or the deliverability of generation within an entire region.

Bilateral deliverability analysis looks at the flow of power between a generator-load pair on the grid at or over a specific time period. For the generator-load pair, power flow analysis can discover:

- Is any generation from the generator delivered to the load, *i.e.*, is it **deliverable?**
- How much electricity generated by the generator is delivered to the load, *i.e.*, what is the total **delivery?**

Regional deliverability analysis looks at the flow of power between all generator-load pairs in a region and can help answer:

- **Delivery distance:** What is the median distance power traveled from the generator or to the load (across all generator-load pairs that are deliverable)?
- **Expected deliverability:** What is the proportion of generator-load pairs within a region between which some amount of electricity can travel?
- **Regional connectivity:** Do boundaries reflect regions with relatively high delivery between generators and loads?

Delivery distance. Delivery distance is a metric that measures the median distance (in kilometers) traveled by power from a given generator, or to a given load. Delivery distance can be used to draw a radius around a generator or load within which the generation can be considered deliverable. However, one limitation of delivery distance is that it does not reflect the direction that power flows, meaning electricity from a generator may not be uniformly deliverable to all loads within the delivery distance radius if the power flow is primarily in a single direction. Delivery distances for all generators or loads in a region can be averaged and compared to the size of the region itself to help evaluate the deliverability of an entire region. Ideally, the average delivery distance should be about the same as the average geographic distance between all generator-load pairs in the region.

36 See "Regional expected deliverability" in Pease (2023) for detail on why it can be important to consider multiple metrics.

Highlight: Delivery Distances in Europe

When mapped over space, delivery distances can show trends in delivery across the network. The maps below show the average distance that power travels to loads (red) and from generators (blue) in each location across Europe over a year. Spatial trends reflect regional differences in electricity delivery. For example, the median distance traveled by power serving loads in Great Britain is 67 km, while loads in the Baltics and Balkans are served on average by power from generators much farther away.

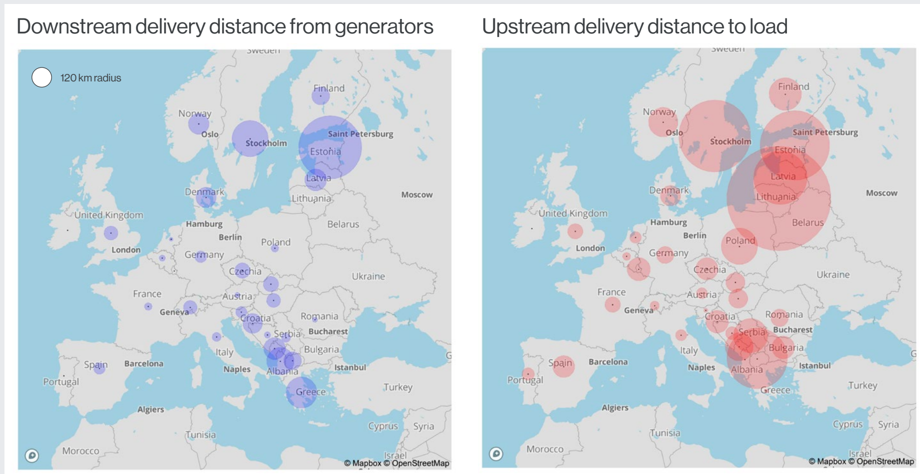

Figure 4. Most electricity is delivered relatively close to where it is produced: in both Europe and the U.S. Eastern Interconnect, on average 50% of a generator's electricity is delivered to loads within 82 km (50 mi). *This suggests that customers desiring strict physical* delivery *as a criterion for procurement likely need to look close to their load centers.*[37]

Expected Deliverability. Because delivery distance represents the *median* distance that power can travel, it does not represent how far smaller volumes of power can travel. Generation is *deliverable* to a location even if only a small amount of power is actually delivered there, since it is physically possible for power to travel between them. Thus, expected deliverability is a regional deliverability metric that describes the percentage of all generator-load pairs within a region between which some amount of electricity can travel. The minimum threshold according to which a pair of nodes is considered deliverable is defined by a **deliverability cutoff** value, which is the number of hours in a year that a minimum volume of electricity is delivered (for example, at least 1 kWh delivered for at least 168 of the hours of a year).[38]

---

37 Data and figures adapted from Pease *et al.* (2023).

38 The choice of cutoff value is subjective and depends on how strictly one wants to define deliverability. One kWh for 168 different hours is highlighted here as an example of a reasonable and intuitive metric. The 2023 Pease *et al.* paper suggests that the relative performance of different regions will generally not change depending on the cutoff value used, although the absolute scores will change (stricter cutoffs, *i.e.* more kW over a longer time period, lead to lower scores).

Intuitively, expected deliverability is related to delivery distance: generally, smaller regions and regions with longer delivery distances have higher expected deliverability. However, because expected deliverability considers the relationships between all nodes in a region, it reflects the direction of power flow in addition to the distance. This nuance highlights the importance of considering multiple metrics to fully capture physical delivery trends across regions.

Highlight: Expected Deliverability in Europe

Although a generator will deliver most of its power within a relatively short radius, there is often a much larger range of loads that receive at least some power from a generator that is farther away. For example, in Germany, electricity is deliverable between 55.6% of generator-load pairs,[39] even though the median delivery distance from generators in Germany is only 78 km (for reference, the distance between Munich and Berlin is approximately 500 km).

Averaging the expected deliverability scores within each boundary type can be used to evaluate how well the boundary as a whole performs. As shown in panel A, bidding zones and countries perform similarly, with 64.2% of generation deliverable to load,[40] while within synchronous grids, generation is less deliverable to load on average. However, as panel B shows, the expected deliverability scores within a single boundary type can vary widely, from 37% in France to 100% in Ireland.

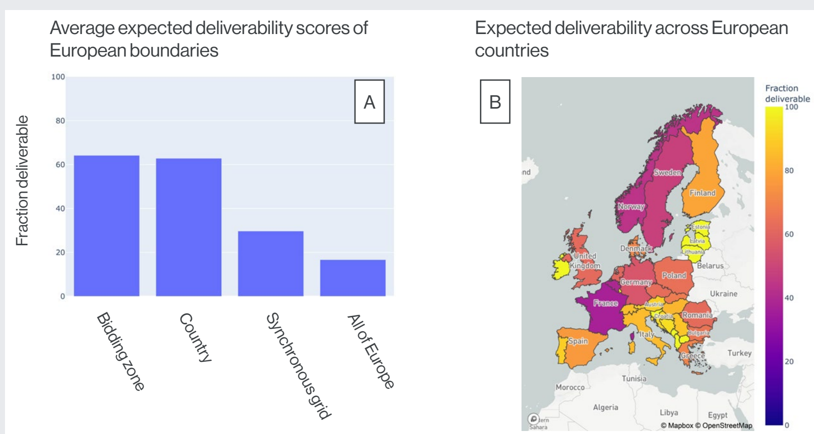

Figure 5. *Data and figures adapted from Pease et al (2023).*

---

39 Using a deliverability cutoff value of at least 1 kWh for at least 168 hours of a year.

40 The small difference between bidding zones and countries is likely the result of aggregating the Baltic states, which each have very high deliverability, into larger bidding zones: Slovenia, Croatia, and Bosnia and Herzegovina are combined (zone SHB); and Serbia, Montenegro, and North Macedonia are combined (zone SMM). This effect on the aggregate deliverability score is offset, and the overall score slightly improved, by increased deliverability from the division of Italy, Norway, and Sweden into multiple bidding zones.

Regional connectivity

Regional connectivity describes how well a set of boundaries matches the natural divisions between strongly connected and weakly connected parts of the larger grid. Regional connectivity is quantified using a metric borrowed from the field of network theory called *modularity*, which compares the strength of the bilateral delivery relationships of each node within each boundary to those of the entire network. In contrast to the previous two criteria, regional connectivity is a single score for the boundary definition as a whole (*e.g.*, national borders), and not for an individual region (*e.g.*, France). Although regional connectivity is somewhat less intuitive than expected deliverability, it is more robust as it does not depend on the choice of a cutoff value, which may affect the expected deliverability score.

Highlight: Regional Connectivity

Regional connectivity is determined by graphing the strength of the connection between every set of pairs in the entire network (A). Each line represents average power delivered over one model day between a generator and a load, with line weight corresponding to average power delivery. Note that connections on this plot do not reflect actual transmission topology; instead, they represent bilateral delivery relationships between generators and loads. Additionally, note that the regional connectivity score for all of Europe is 0 because by definition the modularity of an entire graph is 0.

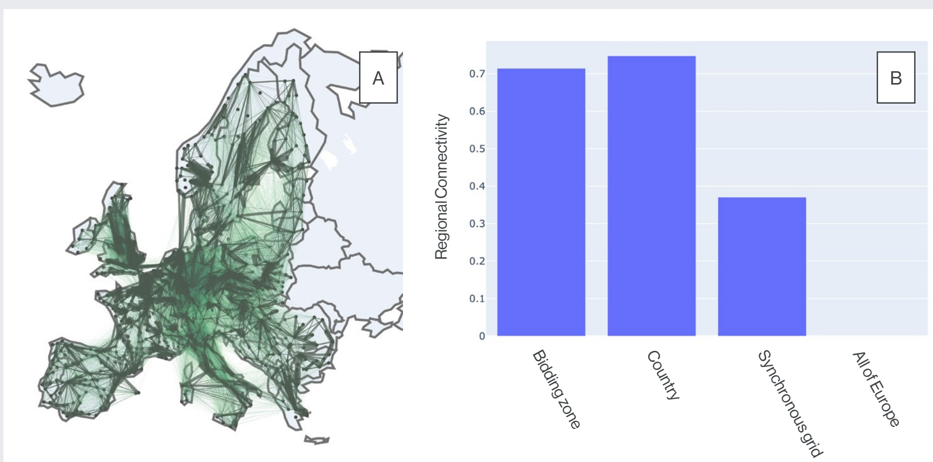

Figure 6. *Evaluating different types of boundary definitions in Europe (panel B) reveals that on average, bidding zones and countries perform similarly and better than synchronous grids.*

Data and figures adapted from Pease et al (2023).

7. Consideration Three: Practicality

> **Key Takeaways**
>
> - Practicality is a measure of reporters' ability to act. Actionability can be negatively impacted by lack of consumer choice, increased complexity, and emerging risks.
> - Standards setters must consider the tradeoff between practicality and impact. For example, if there is no barrier to entry, there may be a lot of actors and actions, but each action may have little impact or unintended consequences. On the other hand, if the barrier to entry is too high, each action may have a large, intended impact, but there may be very few actors, leading to low impact overall.

Background and Justification

In order for a voluntary market boundary to encourage widespread participation and achieve the intended decarbonization impact, it must be practical. The practicality of a voluntary market boundary can be defined according to its:

- **Achievability:** Can participants achieve the program goals by only procuring resources from within the boundary? If not, do buyers have any pathways to expanding access?
- **Understandability:** Is the boundary definition conceptually simple enough for voluntary buyers of all abilities to understand and act upon?
- **Universality:** Is the boundary definition consistent across all geographies where the standard will be used?
- **Stability:** Are the boundaries stable over time to reduce the risk of "stranded" voluntary purchases?

When evaluating these criteria, it is important to consider the rapid pace of innovation in this sector and that what is not achievable or accessible today may be so in the near future. It is also important to consider that a widely used voluntary standard (such as the GHG Protocol) may actually act as a forcing function to accelerate innovations or market developments that do not exist today. Thus, when designing voluntary market boundaries, it may be prudent that standards setters design the boundaries based on what will be, or what they hope will be, practical in the future.

Additionally, standards setters should consider what pathways may exist for voluntary buyers to enhance the practicality of a boundary over time and what incentives the boundary definition may provide for them to do so. For example, if the voluntary procurement boundary was based on the boundaries of existing electricity markets, while this may not be practical for all buyers today, it may incentivize the voluntary market in these areas to advocate for the creation of a wholesale electricity market, or to expand the boundaries of an existing market through regionalization. Or, if the definition allowed voluntary buyers to procure energy from outside the voluntary boundary if they could prove deliverability, it may encourage more voluntary market investment in transmission expansion and upgrades.

Evaluation Criteria

To help standards setters evaluate the practicality of a proposed voluntary market boundary, we have identified four primary evaluation criteria: achievability, understandability, universality, and stability. These criteria are summarized in Table 6 and described in more detail below.

Table 6: Examples of how to evaluate boundaries using the practicality criteria

Evaluation Criteria	Characteristics of more practical boundaries	Characteristics of less practical boundaries
Achievability	Sufficient capacity of diverse and cost-effective resources are (or can be) available within the boundary such that all buyers can achieve program goals.	There is not enough potential supply of carbon-free energy resources with a boundary to meet the needs of all loads/buyers wit=hin that boundary.
Understandability	A first-time buyer could easily understand where they can and cannot source electricity from on a map.	Boundary definitions are complex, vague, full of exceptions, and/or require independent analysis to understand.
Universality	The boundary exists in and represents functionally the same structure in all geographies where the program is used.	The boundary does not exist in, has different meanings in, or is functionally different in different geographies.
Stability	The boundary does not change or get more restrictive over time.	The boundary changes frequently and/or gets smaller over time.

Achievability: In order for a boundary definition to support achievable goals, it must allow all participants to procure the resources needed to meet the program goals. If one of the program goals is to encourage participants to procure enough renewable energy to match 100% of their load, then the boundary must be wide enough to make that feasible. For example, a definition that requires all participants to procure electricity generated within 25 miles of their load would likely not be feasible for most participants, as it may not be possible to build enough capacity that close to load centers (especially in urban areas). The feasibility of a boundary could be defined not only by the total capacity, but also the diversity of the resource types and the cost effectiveness of the resources that could be built within that boundary. For example, if a goal of the program is to encourage around-the-clock supply of clean energy, then a boundary that is too small may not allow for the geographic and resource diversity needed to meet such a goal.

Conversely, the cost-effectiveness of the current resources that could be built within a boundary may be less important if one of the goals of the program is to encourage voluntary buyers to help buy down the cost of emerging technologies or in emerging markets. Achievability does not necessarily mean that it is *easy* to achieve the goal, just that it is *possible.* For example, if a goal is not immediately achievable due to a boundary, it would presumably encourage voluntary buyers to engage with regulators and/or utilities to expand access to resources, expand market boundaries, or create incentives for new supply in the region.

Understandability: The way that the voluntary market boundary is defined should be conceptually simple enough for participants of all levels of sophistication to easily understand and act on. The understandability of a boundary may be limited if its definition contains many rules and exceptions that are difficult to understand or require each participant to conduct independent research to identify where they are able to procure energy from. Boundaries may also not be easily understandable if they require complex tools that are difficult to use or data that may not be widely available. To evaluate understandability, boundary designers could ask: Would a first-time energy buyer without a specialized staff be able to understand where they could procure electricity from just by looking at a map?

Universality: Universality refers to whether a boundary definition is relevant in each of the geographies where the program will be used, and its importance may depend on the scale of the program being designed. For example, if designing an international standard, the boundary definition used should be functionally consistent and relevant across the world. For example, requiring that all generation be procured from within the same NERC region as the buyer would not be universal because NERC regions are unique to the United States. Likewise, using national boundaries would not be universal because even though national boundaries exist around the world, they are not functionally consistent. In the EU, for example, national boundaries generally represent a single grid balancing area, but in the U.S., the national boundary includes multiple distinct synchronous grids and dozens of grid balancing areas. However, if an incentive program is being designed for a single country or electricity market, the universality of the approach may be less important. Similarly, if a standard chooses to tailor boundary definitions to each regional context, universality may be irrelevant.

Stability: The stability of a boundary refers to whether the boundary changes over time. Boundary stability is important because it reduces the risk of “stranded assets,” *i.e.* resources that may be procured because they qualify today, but may not qualify in the future if the boundary changes. Boundary stability makes the program more accessible: if boundaries change frequently or unpredictably, it may be more difficult for less sophisticated participants to track and comply with them. Thus, intentionally dynamic boundaries may not be well suited for use in all programs. However, boundaries need not be permanent. Electricity market boundaries sometimes change, especially given ongoing trends of market regionalization. But in these cases, changes are fairly infrequent and generally expand the boundaries of a market, rather than constrict them, limiting the risk of stranded assets.

8. Case study: Evaluating the GHG Protocol Market-based Boundaries

Key Takeaways

- The current GHGP market boundary guidance lacks clarity and does not reflect physical deliverability or structural relevance well, which provides no incentive for the voluntary market to advance electricity market expansion and regionalization, transmission buildout, and greater interconnection of different synchronous grids.
- A single boundary definition may not work well in all global contexts. For global or cross-regional standards, there may be benefits to selecting boundaries on a more regional basis or creating new boundaries that better balance the most important considerations.
- Region size within an existing boundary type can range widely, complicating the evaluation of deliverability and practicality. Larger voluntary market boundaries tend to be more practical but less structurally relevant and representative of physical deliverability.
- Consistency with other program-specific requirements should be considered.

In this section, we illustrate how the voluntary market boundary evaluation framework could be applied in the case of the GHG Protocol Scope 2 market-based accounting boundaries. The goal of this case study is to illustrate the use of the framework in action, rather than to provide any specific recommendation about which boundary is "best" or should be used for market-based accounting in the Scope 2 guidance. Likewise, although this framework could be applied globally, this case study does not attempt to comprehensively evaluate all possible boundaries across the entire globe.

In this case study, we use the framework to evaluate the market boundaries as currently defined in the GHG Protocol, as well as three additional potential boundary definitions that could be considered. Although the GHG Protocol is a global standard, we will focus on comparing these boundaries across the U.S. and EU for the sake of simplicity and data availability.[41] Based on the grid models we have available at the time of writing, flow-based

41 The geographic scope of our case study is limited to these two geographies based on the grid models we had available at the time of writing to evaluate some of the flow-based physical deliverability metrics.

physical deliverability will only be analyzed for Europe and the Eastern Interconnect of the U.S.

How are the current GHG Protocol market boundaries defined?

As described in Section II, the current guidance from the GHG Protocol recommends that reporters use the following market boundaries:[42]

- Balancing areas and RTOs are typical electricity market boundaries, but energy attribute certificate (EAC) boundaries may be broader;
- Companies should check with regulatory authorities, certification bodies, and/or issuing bodies to see what boundaries they have established, and use those;
- However, if those boundaries are not clear, treat the entire U.S. or EU as a single market;
- However, in order to use the boundary, there also must be a consistent way to calculate a residual mix and retire EACs within that boundary to avoid double-counting;
- Additionally, one should only source contractual instruments from regions that are "reasonably linked" to the region where you consume electricity.

Although this current definition is not necessarily practical given the logic is vague and somewhat challenging to follow, it does emphasize both physical deliverability ("reasonably linked") and structural relevance (checking with regulatory authorities, alignment with electricity markets, consistent accounting frameworks).

In this case study, we will interpret this guidance to allow for a U.S.-wide market boundary in the U.S. and a Europe-wide market boundary in the EU, which is how much of the voluntary market interprets this guidance.

What other boundaries are considered in this case study?

Besides the current GHG Protocol boundaries, we will evaluate three additional types of boundaries that are possible options for market-based accounting: synchronous grid boundaries, balancing areas, and state boundaries.

We chose synchronous grids as an example of a boundary that requires procurement to be "reasonably linked" but is also quite geographically expansive in many cases. We chose balancing areas as an example of a boundary with high structural relevance that is linked to electricity markets, rather than attribute markets.

We chose states as an example of a boundary that is somewhat vague on a global scale and also is a political, rather than necessarily structurally relevant boundary, although in the case of the EU and U.S., states often have some structural relevance. In our example, a U.S.-based buyer may interpret "state" to mean one of the U.S.'s sub-national political

42 GHG Protocol Scope 2 Guidance, WRI (2014): pages 64–65.

boundaries (*e.g.*, Florida, New York), and that an EU-based buyer may interpret "state" to mean nation-state (*e.g.*, Germany, France) rather than a sub-national administrative region. This is done intentionally to illustrate how vague and non-universal terminology can lead to confusion, and also because in Europe, nation-state boundaries mostly coincide with balancing area boundaries. This allows us to illustrate how two functionally similar but differently defined boundaries can lead to varying evaluation scores.

Step 1: Weighting the evaluation criteria

Before beginning the process of evaluating each criterion, it can be helpful to first establish the relative importance of each of the considerations and criteria based on the goal of the program. If standards designers wish to quantitatively score each of the options under consideration to identify the "best" option for their program, assigning a weighting factor for each of the factors is a helpful way to approach this.

In the context of the GHG Protocol revisions, the weighting and evaluation scheme would be determined by the Protocol secretariat through a stakeholder process, and the specific weights here are intended to be illustrative, rather than prescriptive. Hypothetically, if the GHG Protocol's goals were to prioritize a rigorous and credible accounting framework, they might weight physical deliverability and structural relevance higher than practicality. In this example, we assign a combined score for the performance of each boundary across both regions we consider (the U.S. and EU), but this scoring could also be applied on a region-by-region basis.

Sample weighting scheme for evaluating boundaries

There is no single best way to perform this weighting, but we introduce an example system here for illustrative purposes. In this example, we first weight the three criteria, and then each of the evaluation criteria within each criteria. However, an alternative approach could be to simply weight all 11 criteria across the three categories directly.

First, we distribute 10 points to each of the three considerations based on their relative importance to your program goals. The more points each consideration is assigned, the higher it will be weighted.

Structural Relevance: 4
Physical Deliverability: 4
Practicality: 2

Next, within each consideration, we assign a percentage (out of 100%) to each evaluation criterion based on its importance. A higher percentage will be weighted higher, so it should reflect criteria that are most important to the program.

Structural Relevance	Electricity Market Alignment	50%
	Regulatory Alignment	50%
Physical Deliverability	Physical grid boundaries	20%
	Transmission congestion	20%
	Delivery distance	20%
	Expected deliverability	20%
	Boundary connectivity	20%
Practicality	Achievability	30%
	Understandability	30%
	Universality	15%
	Stability	25%

Finally, we define a point scale that can be used to evaluate each criterion. In this example, we'll choose a 3-point scale, with 0 being bad, 1 being neutral/mixed, and 2 being good. A larger scale (such as a 5-point or 10-point scale) could allow for more nuanced and flexible evaluation.

0 (bad)	**1 (mixed/neutral)**	**2 (good)**

To calculate the final score, we multiply each criterion score by the criteria weight and the consideration weight. Then, all of these weighted points would be summed, up to a possible total of 20 points in this example.

Step 2: Evaluate structural relevance

For this illustrative case study, we evaluate how structurally relevant each of these boundaries are across the U.S. and EU.

Criteria	Entire U.S./EU (Current GHGP)	Synchronous Grid	State	Balancing Area
Electricity Market Alignment	There is no single U.S.-wide or Europe-wide electricity market that allows for unrestricted trade of electricity.	Synchronous grids can sometimes overlap with regionalized market boundaries, but often they are quite large and may not encourage investment in the same electricity market where the buyer is located.	Although electricity and state markets sometimes overlap (in Europe), market boundaries in the U.S. often span many states (*e.g.*, MISO), or there can be multiple markets in a single state (*e.g.*, Florida).	Market boundaries are often aligned with balancing areas. However, with market regionalization, some market boundaries can be much larger than a single balancing area.
Regulatory Alignment	In the U.S., although certain aspects of wholesale markets are regulated nationally by FERC, most electricity markets are regulated at the state/regional level. In Europe, there is regulatory coordination among EU members, but regulations are still set at the national level.	Often within a single synchronous grid there is an overlapping patchwork of regulatory standards including portfolio standards, cap-and-trade systems, and reliability rules.	In the U.S., state utility commissions often set reliability rules, and portfolio standards are often set at the state level. Even so, even these programs can allow out-of-state eligibility of resources.	Procurement within a balancing area would generally support grid reliability needs, although reliability is not always coordinated at the BA level, and other regulatory policies such as cap-and-trade can span multiple BAs.

Step 3: Evaluate physical deliverability

For this illustrative case study, we evaluate how well each of these boundaries reflect physical deliverability across the U.S. and EU.

Criteria	Entire U.S./EU (Current GHGP)	Synchronous Grid	State	Balancing Area
Physical Grid Boundaries	The current boundaries span multiple synchronous grids in both the U.S. and Europe.	Synchronous grids by definition align with the physical boundaries of the grid and do not allow procurement from a different synchronous grid.	Political state boundaries sometimes span multiple synchronous grids, such as Texas (parts of the state are in the Texas, Western, and Eastern Interconnects), or nation-states that include islands (such as the island of Sardinia in Italy).	Balancing authorities exist within the boundaries of a single synchronous grid.
Congestion	The current geographically expansive boundaries would include many different congested zones.	Although some synchronous grids are small and do not have any transmission congestion, others span continents where there are substantial spatial differences in wholesale electricity prices driven by congestion.	In addition to sometimes spanning multiple interconnections, states can be quite large and often have congested zones.	Some balancing areas are quite large and include distinct congestion zones within their borders.
Delivery Distance	U.S. and Europe-wide boundaries are much larger than typical delivery distances.	Synchronous grids are generally larger than delivery distances.	State boundaries can be smaller or larger than average delivery distances. However, because of the wide variation in state boundary size, some state boundaries may be appropriate.	BA boundaries can be smaller or larger than average delivery distances. However, because of the wide variation in boundary size, some BAs may correspond with delivery distances.

Step 3: Evaluate physical deliverability (continued)

Criteria	Entire U.S./EU (Current GHGP)	Synchronous Grid	State	Balancing Area
Expected Deliverability	The current boundaries are quite large and will have lower expected delivery and deliverability.	Synchronous grids are generally large boundaries and will have lower expected delivery and deliverability.	Smaller states and states with longer delivery distances have higher expected deliverability, while larger states have lower scores. States that are a similar size may have different expected deliverability scores if they have different delivery distances.	Smaller BAs and BAs with longer delivery distances have higher expected deliverability, while larger states have lower scores. BAs that are a similar size may have different expected deliverability scores if they have different delivery distances.
Regional Connectivity	In general, there are many smaller regional grid clusters in the U.S. or Europe.	In general, there are many smaller regional grid clusters within a single synchronous grid.	In the U.S., state boundaries are not related to underlying power flows, except where they happen to align with physical grid boundaries, so there is no natural clustering of generators within state boundaries. In Europe, where nation-state boundaries are also balancing areas, there is a greater amount of clustering in these boundaries.	Balancing authorities generally align with underlying power flows, although depending on the size of the BA and amount of trade with neighboring BAs, the natural regional clusters may not always match BAs.

Step 4: Evaluate practicality

For this illustrative case study, we evaluate how practical each of these boundaries is across the U.S. and EU.

Criteria	Entire U.S./EU (Current GHGP)	Synchronous Grid	State	Balancing Area
Achievability	U.S. and Europe-wide boundaries give buyers a large amount of flexibility and choice when procuring energy.	Synchronous grids are often broad enough to provide much choice to purchasers. Although this may not be the case in smaller island grids, voluntary procurement outside of this boundary will not help decarbonize the islanded grid.	Some states can be quite small, and thus it may not be practical or cost effective to source all generation from within that boundary.	Some BAs can be quite small (*e.g.*, a single city), and thus it may not be practical or cost effective to source all generation from within that boundary.
Understandability	The current definition is vague and requires each purchaser to do independent research to determine their local boundaries.	Synchronous grids are mutually exclusive and their boundaries are clear, although maps of synchronous grids are not always currently widely or easily available.	It is easy to look at a map and understand state boundaries.	Although maps of BAs are widely available, BA boundaries sometimes overlap (especially in the U.S.) or can have both physical and commercial definitions.
Universality	The current definition can have different and inconsistent interpretations in different regions.	Synchronous grids are universal across the globe.	"States" are somewhat vague and not functionally universal across the globe.	Functionally, BAs exist in both geographies.
Stability	The boundary definition rule is based on potentially shifting market rules, tracking systems, and data availability, but as interpreted, the boundaries of the U.S. and Europe are stable.	Synchronous grid boundaries are generally quite stable, and any changes generally involve decadal-scale planning processes and expand rather than contract boundaries.	Except in extreme geopolitical circumstances, state boundaries do not change.	BA boundaries are generally stable, but do change as markets evolve (often growing larger rather than smaller).

Step 5: Score the boundaries

Based on the scores and weights assigned in this illustrative example, the weighted score for each of the considerations would be (out of 2 points):

Category Weight	Category	Weighted Criteria Score (0–2)			
		Current GHGP Boundaries	Synchronous Grid	State	Balancing Area
4x	**Structural Relevance**	0.5	1	1	1
4x	**Physical Deliverability**	0	0.4	1	1.2
2x	**Practicality**	0.85	1.7	1.4	1.4
Total Weighted Score (0-20) (multiply category weight by weighted criteria score and sum)		3.7	9.0	10.8	11.6

Key takeaways from implementing the framework in practice

The purpose of this limited, illustrative example is not to identify which of the four boundaries would be "best" to use for scope 2 market-based accounting. In order to do so, a larger set of boundaries would need to be evaluated more robustly across more geographical contexts than only the U.S. and EU, and the scoring and weighting would need to be informed through a transparent, multi-stakeholder process. However, it is clear that the current market-based scope 2 boundaries are not sufficient. We find that although the GHGP emphasizes physical deliverability ("reasonably linked") and structural relevance (checking with regulatory authorities, alignment with electricity markets, consistent accounting frameworks) in its written guidance, today's boundaries as currently interpreted perform poorly across both of these metrics, suggesting that they should be updated to align with the stated intent of the standard.

This exercise highlights a number of challenges and subsequent recommendations for framework implementation.

Defining global boundaries vs. tailoring boundaries for each context: Assessing boundaries for a global standard, like the GHG Protocol, can be a particularly challenging exercise because a boundary definition that performs well in one context may fail in another. While we did not evaluate an exhaustive list of potential boundary options in this case study, it is possible that a more comprehensive, real-world exercise may also fail to identify one universal best option. In a global context where it's hard to consistently evaluate the same boundary across geographies, another valid approach is to craft different, yet *specific* boundary definitions for each geography. As a hypothetical example, perhaps the guidance for North America would be to use balancing area boundaries, while the guidance for Europe would be to use synchronous grid boundaries, and the guidance for China would be to use national boundaries. Or, perhaps the answer is to not use a boundary-based definition at all, but instead use a rule-based or instrument-based definition. In any case, this framework can be used flexibly to evaluate all of these options.

Inconsistent region sizes make evaluating deliverability and practicality difficult: We also found that actual regions within a single boundary type may be very different sizes, which can affect their practicality and physical deliverability. For example, the PJM balancing area in the U.S. covers over 386,000 square miles and includes 13 states within its boundaries, while the Tampa Electric Company (TEC) balancing area covers a single city in a single U.S. state.[43] The former area has an expected deliverability score of 18%, while the latter's score is 68%.[44] In the first case, there is generation in some parts of PJM that are not deliverable to other parts of PJM, and in the latter case, there are resources outside of Tampa that would be deliverable to the city. Likewise, while it would be achievable for a buyer in PJM to source all of their clean electricity from within PJM, it is unlikely that a company operating in Tampa could source all of their clean energy from within the metropolitan area.

Re-using existing boundaries vs. drawing new, custom boundaries: If the variability in quality of a pre-defined boundary makes it too difficult to evaluate, it is also possible to use this framework to draw new, custom boundaries for a specific program. Existing boundaries are convenient because they are already defined and understood, but these boundaries can be defined by historical happenstance or structural needs that have little relevance to the needs of the voluntary program being designed. Inventing new boundaries to meet a specific voluntary program need is not without precedent. For example, in the U.S., the Environmental Protection Agency developed "eGRID subregions" to better represent consumption boundaries for location-based GHG accounting, and designed them to find a balance between the limitations of balancing areas and interconnects. This example shows that the GHG Protocol could design a new set of market boundaries to optimize for the goals of market-based accounting.

Consistency with other program-specific requirements: In certain contexts, it is also important to consider criteria specific to the standard. For example, in the GHG Protocol, if a reporter does not have contractual emissions factors or utility-specific emissions factors for all of their load, they are instructed to use a "residual mix" emissions factor, which represents the average emissions intensity of all generators to which no other grid user has a contractual claim to the electricity. For the sake of consistency and to avoid double counting, it is necessary to use the same boundary for the residual mix as is used for voluntary procurement. Thus, the GHG Protocol revision process may need to additionally consider the consistency of the voluntary market boundary with the boundaries used to calculate residual mix emission factors.

---

43 "PJM At a Glance." https://www.pjm.com/~/media/about-pjm/newsroom/fact-sheets/pjm-at-a-glance.ashx

44 Pease, Gailin *et al.*, "Tracking physical delivery of electricity from generators to loads using power flow tracing." Preprint at doi.org/10.5281/zenodo.8117207 (2023).

9. Conclusion

The voluntary procurement boundary framework introduced in this paper is a flexible and powerful tool that can be used by the designers of programs and standards to identify or design credible, impactful, and practical boundaries for voluntary energy procurement. Each of the three pillars of the framework is intended to help designers evaluate and think critically about the broader impact of their design choices on electricity markets, grid needs, the users of their standards, and global decarbonization. The first pillar, structural relevance, asks designers to consider how their program will either support or undermine electricity markets and regulations, and whether it amplifies the voluntary demand signal by aligning it with that of other actors in the market. The second pillar, physical deliverability, asks designers to examine the physical grid structure and consider whether the claims and intended impacts of the program are supported by the actual dynamics of power flows on the grid. The third pillar, practicality, asks designers to consider how their design may affect the adoption of the standard and thus the impact it may have.

These three considerations were chosen to help voluntary market designers strike a critical balance to achieve the intended impact of the program. Large boundaries may encourage widespread participation in the market, but may cause questionable impact if the action taken by each participant is less meaningful or relevant to driving structural change. On the other hand, while stricter boundaries may encourage more individually impactful action, they also may limit broad participation, and thus the aggregate impact may be smaller.

Although the case study in this paper focuses on the GHG Protocol Scope 2 market-based accounting standard, the framework is equally applicable to other standards seeking to encourage or regulate voluntary procurement of clean energy, including RE100, the 24/7 Carbon-free Energy Compact, the Emissions First Partnership, the U.S. EPA Green Power Partnership, the U.S. FTC Green Guides, the Science-Based Targets Initiative (SBTI), corporate climate disclosure regulations, building energy performance standards, and green hydrogen regulations. In each of these examples, the use of the framework could be tailored to the specific geographic scope of the program (from state, to national, to global), or the specific aims of the program (energy accounting, emissions offsets, credible and accurate claims). For example, the U.S. Federal Trade Commission might put a higher emphasis on physical deliverability to maximize credibility of claims.[45] The EPA's Green Power Partnership, with its focus on markets, might prioritize structural

45 The Federal Trade Commission's Green Guides are designed to help marketers avoid making environmental claims that mislead consumers. https://www.ftc.gov/news-events/topics/truth-advertising/green-guides

relevance.[46] Conversely, the Emissions First Partnership might prioritize practicality over physical deliverability or structural relevance.[47]

This work is intended as a starting point for a broader, more robust conversation on the topic of deliverability in voluntary clean energy markets, and there is still much work to be done. Future directions for this work could include:

- Refining the flow-based physical deliverability metrics introduced by the research cited in this paper and expanding access to the data and models needed to evaluate flow-based deliverability.
- Examining the applications and limits of the congestion-based fungible deliverability metric that uses LMP data, particularly as it applies to the context of new loads, new generation, and new transmission.
- Ensuring the universal applicability of this framework and introducing refinements based on different global contexts.
- Exploring how the methods described in this paper can be applied to better understand the appropriate spatial granularity for location-based carbon accounting. Power flow tracing can also be used to create nodal estimates of consumed emissions, which could help identify how much the carbon intensity of consumed electricity varies across space. This could be useful to inform whether the ongoing use of regional-average emission factors is accurate.

The path toward global decarbonization will involve voluntary efforts by companies, and deliberate procurement and accounting frameworks can better harness the power of corporate purchases to spur the growth of clean electricity. The voluntary market boundary framework introduced in this paper is designed to help standards setters critically evaluate the broader impact of their design choices on electricity markets, power system needs, the users of their standards, and – ultimately – global decarbonization. However, the framework presented in this paper is intended as a starting point for a broader, more robust conversation on the topic of deliverability in voluntary clean energy markets, and we invite others to apply the framework and put it to the test to help refine it over time.

---

46 Some of the goals of the program include advancing green power markets and supporting basic market principles. https://www.epa.gov/greenpower/about-green-power-partnership

47 The partnership seeks to value clean energy procurement targeted to locations with maximum decarbonization impact at the global level, irrespective of grid or market boundaries. https://www.emissionsfirst.com/principles