

Technical Appendix F

Summary of Differences between RSEI Data and the TRI National Analysis

Table of Contents

- 1 Introduction..... 3**
- 2 Off-site Facility Consolidation 3**
- 3 Off-site and On-site Facility Latitude/Longitude Revisions 3**
- 4 Adjustments for Double-Counting 4**
- 5 Selection of Industry Classification Codes 4**

1 Introduction

The RSEI model currently uses a data set that EPA uses in publishing EPA’s annual [Toxics Release Inventory \(TRI\) National Analysis](#). This data set is known as the National Analysis data freeze, and represents data contained in the TRI database at a given point in time each year in the autumn season following the July TRI reporting deadline for toxic chemical release reporting forms. This specific extraction point is when EPA “freezes” the TRI data contained in the TRI database and is a cutoff point that the Agency uses so that it can prepare official EPA analyses of the TRI data and other Agency deliverables. Both the TRI National Analysis and RSEI further process and refine this data set for analytical purposes, where some of these processes or refinements may be performed for one dataset and not for the other. The following sections discuss each of the specific processing or refinements made and how they differ between RSEI and the National Analysis.

2 Off-site Facility Consolidation

In TRI, chemical waste transfers to off-site facilities are reported by the facility transferring the chemical, rather than the receiving facility, which results in two problems: 1) names and/or addresses of the receiving facility may be reported incorrectly or incompletely; and 2) the same facility name and address may be reported in slightly different forms by different facilities, making it difficult to determine unique facilities. Determining unique facilities is important for allowing comparisons of waste quantities, hazard, and score across facilities, and can improve overall locational quality by matching records with correct and complete addresses to records that reference the same facility but with incomplete or inaccurate addresses. In some cases, when incomplete information is submitted, assignments are made based on existing information. For example, if a facility reports a transfer to a publicly owned treatment works (POTW) facility to “City Water Department,” that transfer may get manually assigned to a POTW in that city. For specific details on the methods used, please see Technical Appendix D - Locational Data for TRI Reporting Facilities and Off-site Facilities.

3 Off-site and On-site Facility Latitude/Longitude Revisions

Facilities reporting chemical waste management activities to TRI no longer report their latitude and longitudes. RSEI and the TRI National Analysis both use EPA’s Facility Registry Service (FRS) as the primary source of locational data. Additionally, RSEI performs manual verification of high-scoring facilities, and in some cases adjusts the locational coordinates based on satellite data. TRI does not contain coordinates for off-site facilities. It should be noted, however, that in geocoding facilities the quality of the coordinates varies, from an exact match based on street address with verification based on satellite data, to a match based only on a 5-digit Zone Improvement Plan (ZIP) code. See Technical Appendix D for details on the derivation of the locational data for both reporting facilities and off-site facilities for the RSEI model.

4 Adjustments for Double-Counting

TRI facilities must report chemical quantities in waste that are transferred off site to other facilities for further waste management activities. These recipient facilities manage the chemical waste in various ways, some activities of which are then modeled by RSEI. In addition, some of the off-site facilities that receive chemical waste from TRI reporting facilities are treatment, storage, and disposal (TSD) facilities that are regulated under the federal Resource Conservation and Recovery Act (RCRA). These type of facilities were made subject to TRI reporting for the first time beginning in reporting year 1998.¹ The new reporting requirement meant that there was a potential for double counting chemical waste management quantities that are first transferred by a TRI facility to an off-site facility (and so reported to TRI), then managed by a RCRA-regulated TSD (and reported again to TRI). The TRI National Analysis uses facilities' RCRA identification numbers (RCRA IDs) to match waste management activities reported by TSD facilities to off-site chemical waste transfers from other TRI reporters, and omits the matching off-site transfer from the summary. RSEI does not use this same matching routine, but does drop chemical waste transfers that are sent off site for the purpose of incineration where the receiving facility is also a TRI reporter in the North American Industry Classification System (NAICS) code 562211 (Hazardous Waste Treatment and Disposal).

5 Selection of Industry Classification Codes

The consideration of industrial sectors is an essential component of RSEI. The foremost reasons are that the 6- and 4-digit primary NAICS codes for a facility are used to estimate the stack air modeling parameters for those facilities for which facility-specific information is not available. NAICS codes are also important in the estimate of chromium speciation assignments.

When submitting TRI reporting forms, facilities are required to provide the primary NAICS code for the entire facility first, and then list up to five additional 6-digit NAICS codes for other "establishments" (i.e., economic units, generally at a single physical location, where business is conducted or where services or industrial operations are performed) which are associated with reportable environmental releases and other waste management and source reduction activities. At least one reporting form is required for each chemical or chemical category if reporting thresholds are met, but some facilities report multiple forms for a single chemical or chemical category to reflect the activities involved at each establishment or group of establishments. This can lead to more than six 6-digit NAICS codes being reported for a single facility to the TRI Program. When aggregating environmental releases and chemical waste transfers by industry, the TRI Program uses the form-level NAICS codes, so in some cases a facility's reporting forms could be included in different industries depending on the reported data.

In previous years, EPA had performed a frequency analysis for the RSEI model and picked the most frequently reported primary NAICS code and the five additional most frequently reported

¹ These TSD facilities subject to TRI reporting are limited to commercial hazardous waste treatment facilities regulated under the RCRA Subtitle C, 42 U.S.C. section 6921 *et seq.*

NAICS codes, for a total of up to 6 NAICS codes for each facility. Unlike for other TRI data products, each facility was classified solely by its primary NAICS code, and a facility's environmental releases and chemical waste transfers could not be assigned to more than one industry for RSEI. To avoid confusion between RSEI and other TRI data products when looking at data by industry, beginning with RSEI Version 2.3.6, the RSEI model employs the NAICS code assignment process used by the TRI Program. Primary NAICS code assignments are still used to derive the necessary RSEI modeling parameters, but for user-facing applications like EasyRSEI that aggregate available data and results, the TRI Program process is followed that uses the reporting-form level NAICS code data. The RSEI-assigned NAICS code is called the "Modeled NAICS" to avoid confusion.

[revised 4/4/2023]