Technical Appendix F

Summary of Differences between RSEI Data and the TRI National Analysis

Table of Contents

1 Introduction

The RSEI model currently uses a data set that EPA uses in publishing the annual Toxics Release Inventory (TRI) National Analysis. This data set is known as the National Analysis data freeze, as it represents a static reference point for the TRI data, which is continually updated and revised. Currently, RSEI uses the National Analysis data freeze. However, both the National Analysis and RSEI revise the data to increase its accuracy, and some of the revisions are performed for one dataset and not the other. The following sections discuss each of the types of the revisions made, and how they differ between RSEI and the National Analysis.

2 Off-site Facility Consolidation

In TRI, 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 hard to determine unique facilities. Determining unique facilities is important for allowing comparisons of waste volume, 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 POTW to "City Water Department," that transfer may get manually assigned to a POTW in that city. For details on the method used, please see Technical Appendix D.

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

Facilities reporting chemical releases to TRI no longer report their latitude and longitudes. RSEI uses 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 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 ZIP code. See Technical Appendix D for details on the deriving the locational data for both reporting facilities and off-site facilities.

4 Adjustments for Double-Counting

TRI facilities must report any chemicals that are transferred off-site to other facilities. These recipient facilities can dispose of the wastes in various ways, some of which are modeled by RSEI. Some of the off-site facilities that receive chemicals from TRI facilities are treatment,

storage and disposal (TSD) facilities regulated under the federal Resources Conservation and Recovery Act (RCRA). These facilities were required to report their own releases to TRI for the first time in 1998. The new reporting requirement means that there is the potential for doublecounting releases that are transferred by a TRI facility to an off-site facility (and so reported to TRI), then released by a RCRA-regulated TSD (and reported again). The National Analysis uses facilities' RCRA identification numbers to match releases reported by TSD facilities to off-site 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 transfers to incineration after 1998 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 and 4-digit NAICS codes are used to estimate chromium speciation.

When submitting Form R reports by chemical, facilities are asked to always list the primary NAICS code for the entire facility first, and then list up to five additional 6-digit NAICS codes for other "establishments" (defined as "distinct and separate economic activities [that] are performed at a single physical location") which are associated with reportable releases and other waste management and source reduction activities. At least one Form R report is required for each chemical, but some facilities report multiple Form R reports for a single chemical to reflect the activities involving a TRI chemical at each establishment or group of establishments. This could lead to more than six 6-digit NAICS codes being reported for a single facility. When aggregating TRI releases and transfers by industry, the TRI Program uses the form-level NAICS codes, so in some cases a facility's form Rs could be included in different industries.

In previous years, RSEI performed a frequency analysis and picked the most frequently reported primary NAICS code and the five additional most frequently reported NAICS codes, for a total of up to 6 codes for each facility. Unlike for other TRI data products, each facility was classified solely by its primary NAICS, and a facility's releases and transfers could not be assigned to more than one industry. To avoid confusion between RSEI and other TRI data products when looking at results by industry, beginning with Version 2.3.6 (RY2016), RSEI retains the same process to assign NAICS codes to a facility, and uses the primary NAICS code assignment to derive the necessary modeling parameters, but for aggregations in user-facing applications like EasyRSEI, the TRI Program approach will be followed. The RSEI-assigned NAICS code will be called the "Modeling NAICS" to avoid confusion.

[revised 3/1/2022]