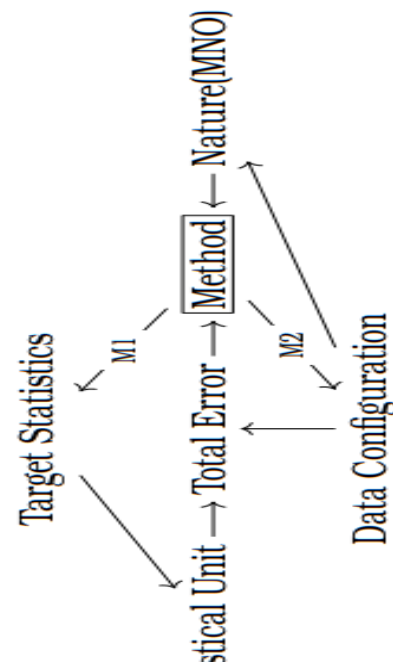


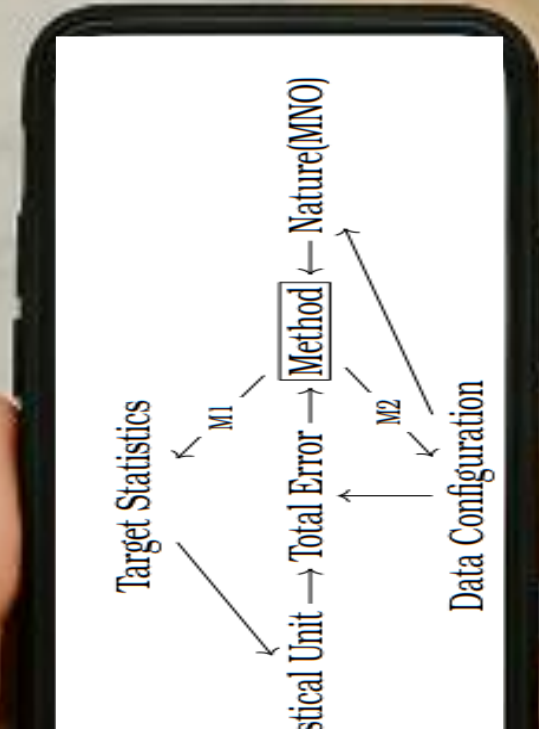
ROADMAP FOR TOTAL ERROR FRAMEWORK

Lost without
a Signal?



Sónia Quaresma
Pedro Cunha

ROADMAP FOR TOTAL ERROR FRAMEWORK



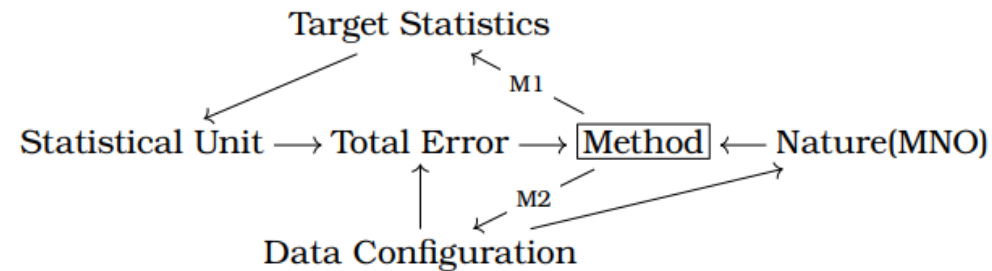
Lost without a signal?

Roadmap for Total Error Framework



Reference frame for methods to combine MNO and non-MNO data

D3.2



M1, M-executor methods; M2, M-enabler methods

To capture the range of problems in combining MNO and non-MNO data, a reference frame was proposed



Lost without a signal?

Roadmap for Total Error Framework

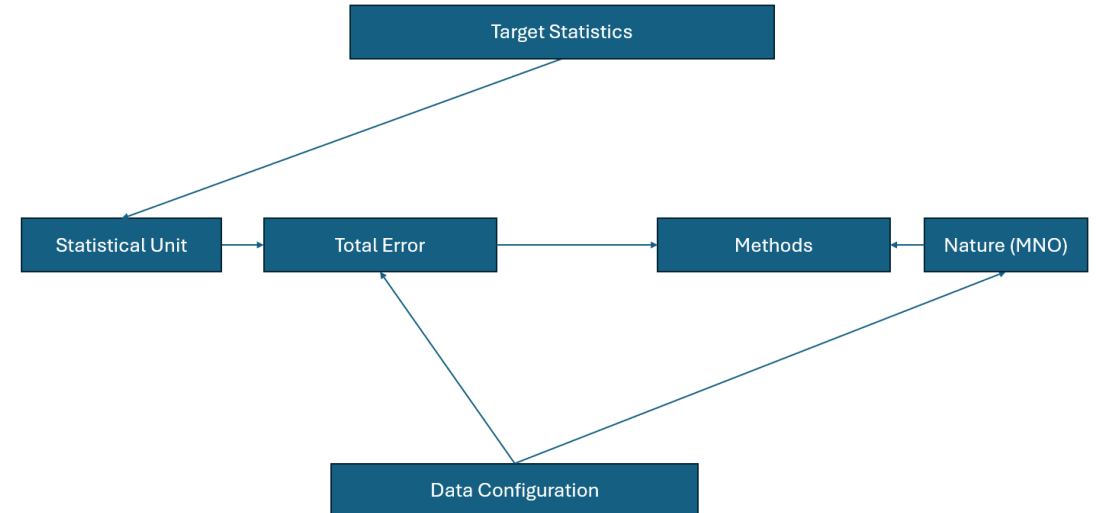
Reference frame for methods to combine MNO and non-MNO data

D3.2



Assumptions:

- Adopting a total error framework allows one to analyze and identify the most important error sources in each situation;
- Specifying a range of generic settings of the data to be combined can provide practical guidance to the relevant methods;



Lost without a signal?

Roadmap for Total Error Framework

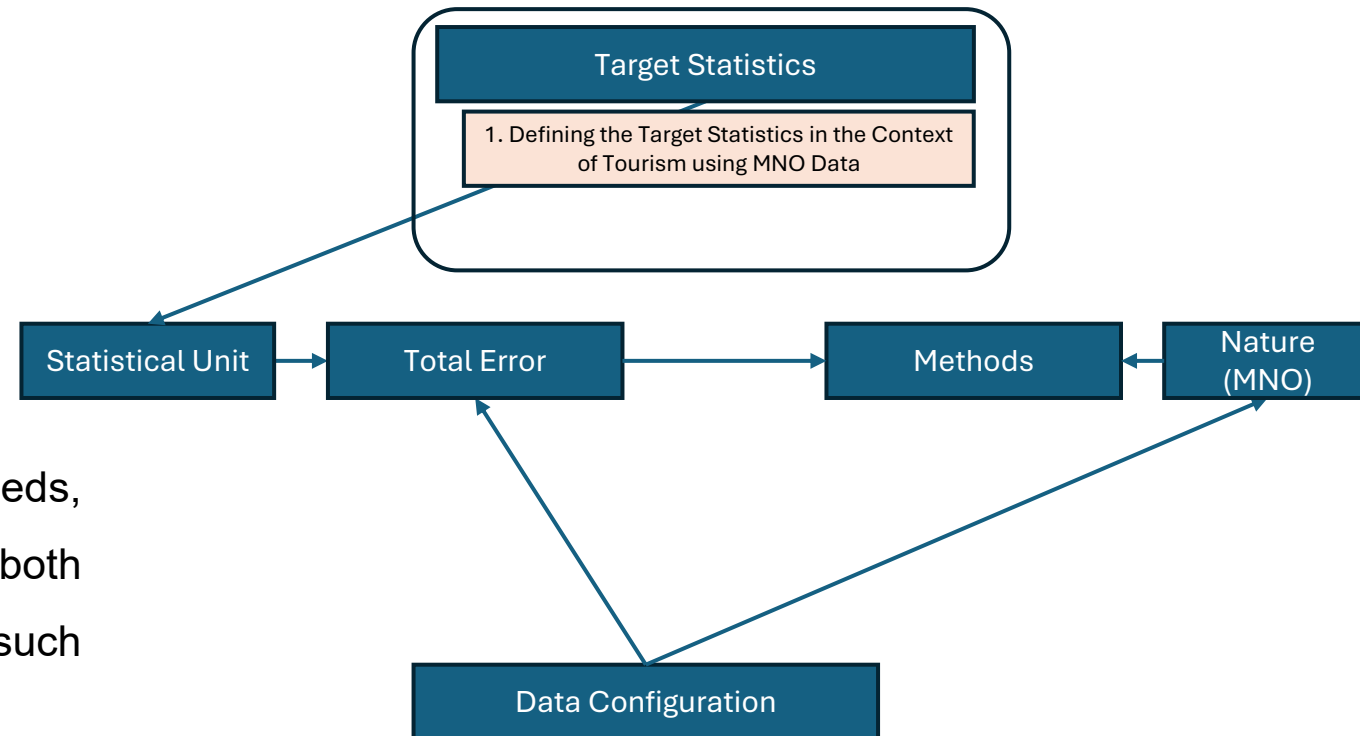
Roadmap

D3.4

Defining the Target Statistics:

- Clarifying the Measurement Objectives
- Defining the Statistical Units
- Specifying the Population of Interest
- Documenting the Definition Phase

This Step ensures alignment between user needs, statistical objectives, and the characteristics of both traditional and non-traditional data sources, such as MNO data.



Lost without a signal?

Roadmap for Total Error Framework

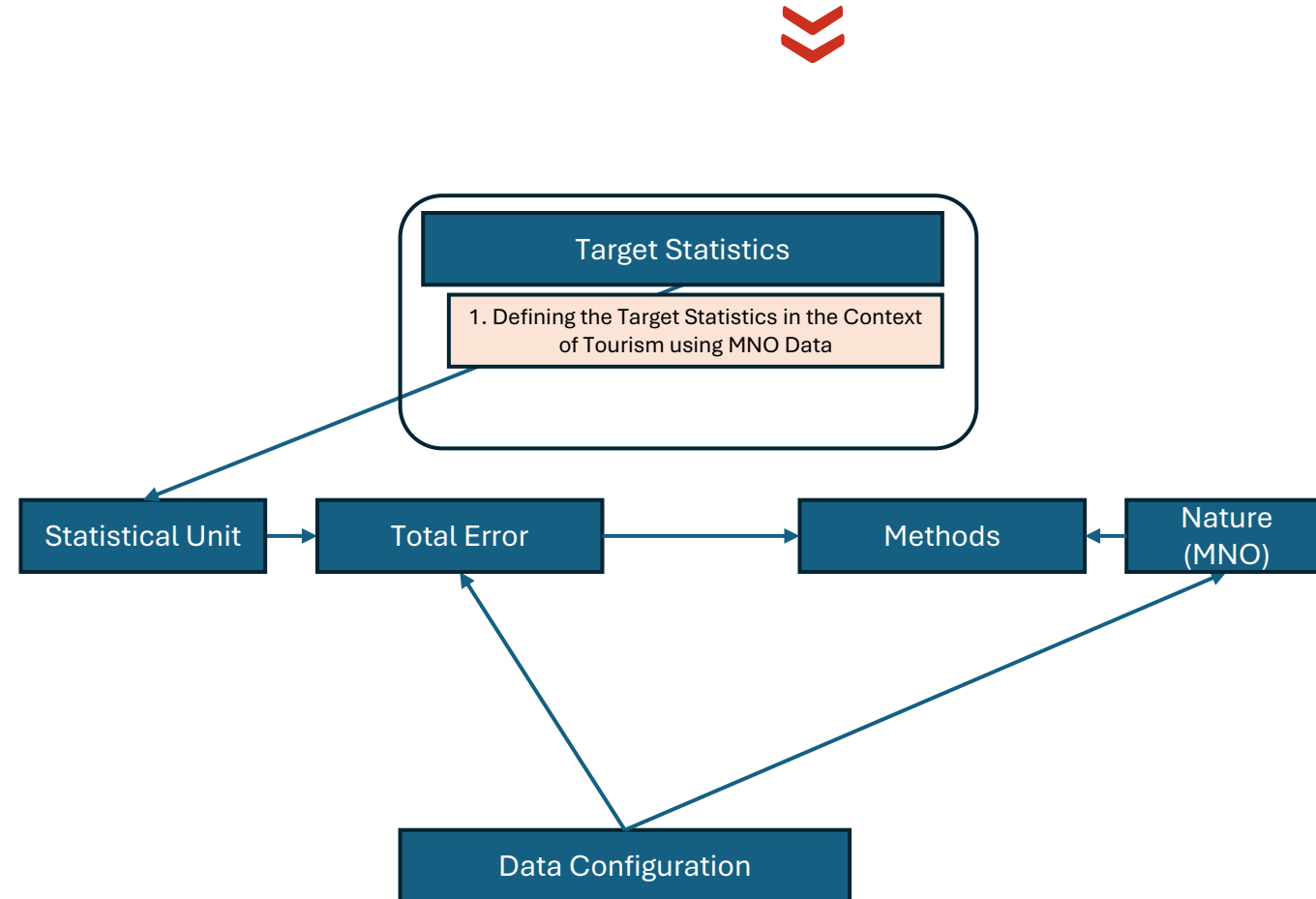
Roadmap

D3.4

Defining the Target Statistics:

- Clarifying the Measurement Objectives - Identify and formalize the main phenomena one intends to measure, in line with the statistical output requirements.

Example Objective Estimate the number of foreign visitors who stayed at least one night in Portugal during the month of Augusts.



Roadmap

D3.4

Lost without a signal?

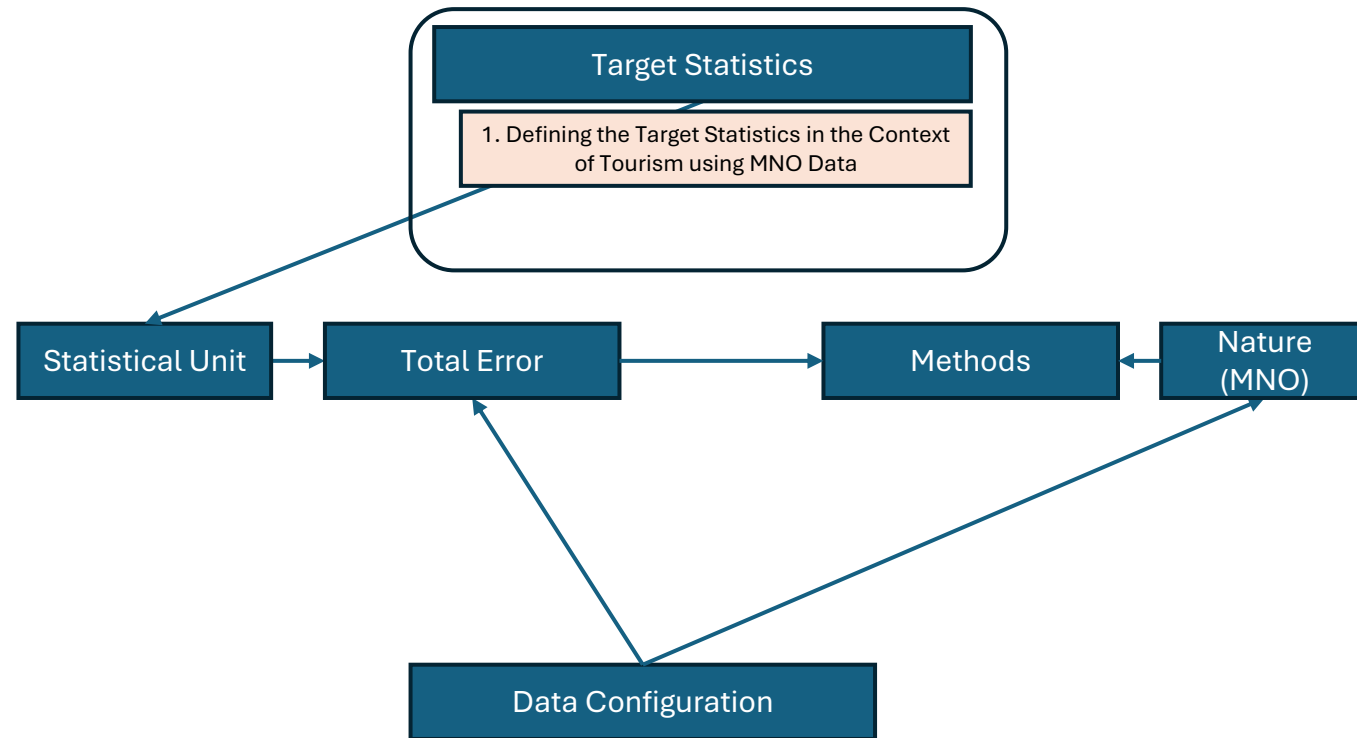
Roadmap for Total Error Framework



Defining the Target Statistics:

- Defining the Statistical Units - Statistical units are the basic entities being measured or counted. It is generally the case for official statistics that one is not interested in making statistics of mobile devices but of persons or spatial objects (such as a city centre).

Example Unit Foreign SIM card (not foreign visitor) recorded in a country..



Roadmap

D3.4

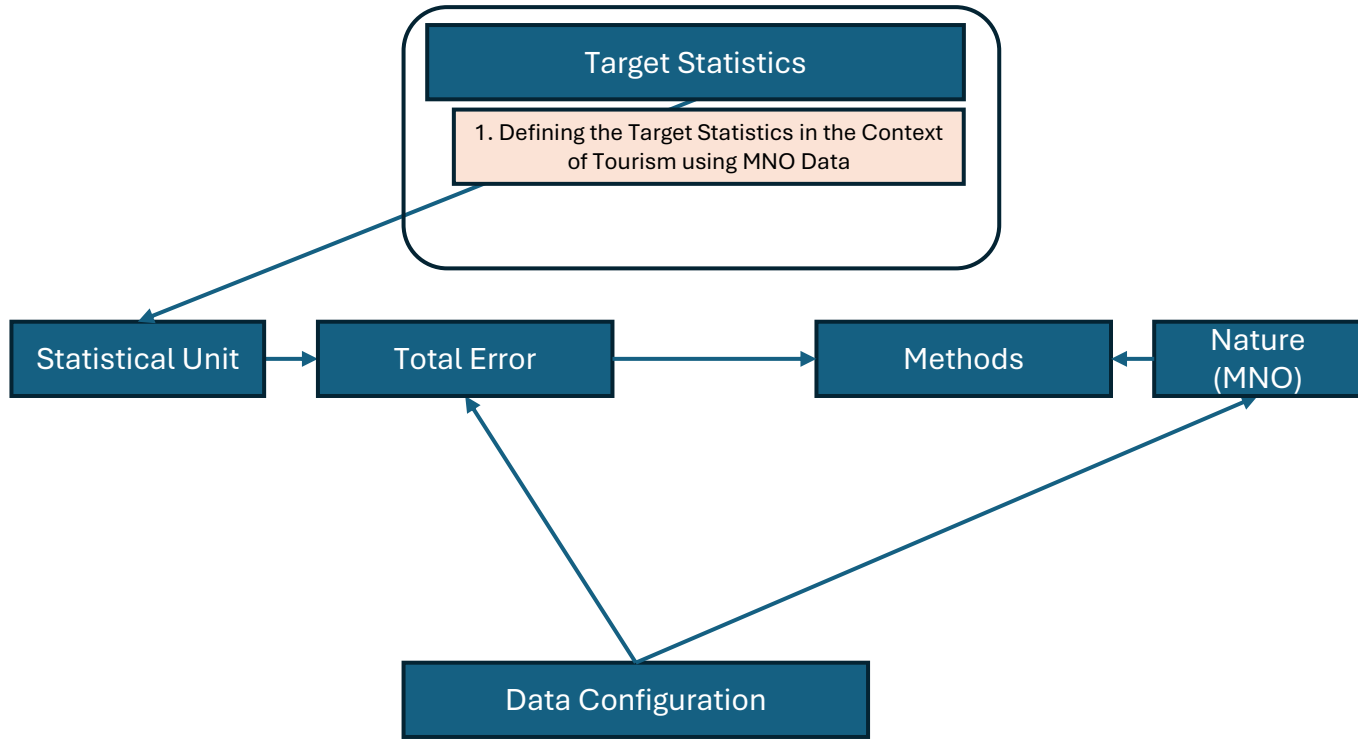
Lost without a signal?

Roadmap for Total Error Framework

Defining the Target Statistics:

- Specifying the Population of Interest The population of interest (or target population) refers to a specified collection of statistical units. It may be explicitly defined in terms of geography, time, and demographic or behavioural attributes. This is essential for assessing coverage errors and misalignments between the target and observed populations.

Example Population All non-resident individuals who entered Portugal with an active-roaming foreign SIM card and spent at least one night within national territory during August 2025. (This may or may not be the target population)



Lost without a signal?

Roadmap for Total Error Framework

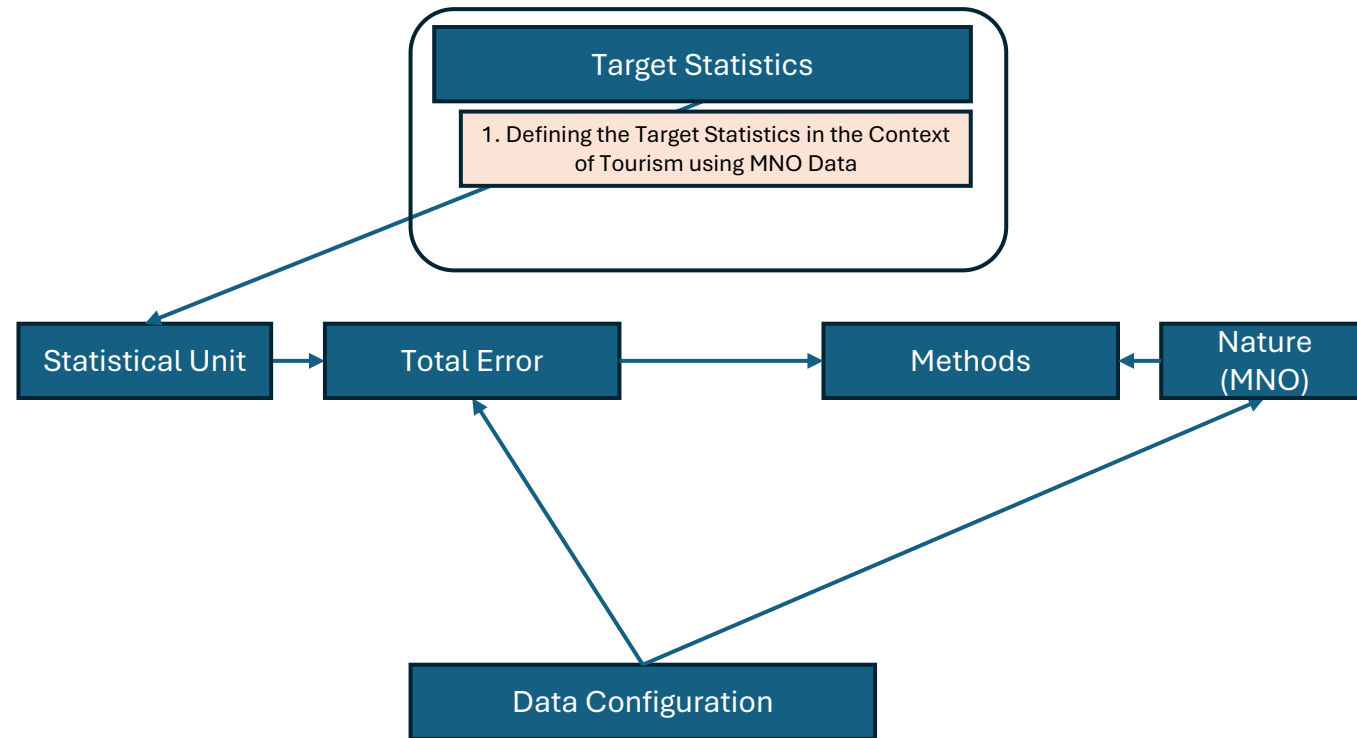


Roadmap

D3.4

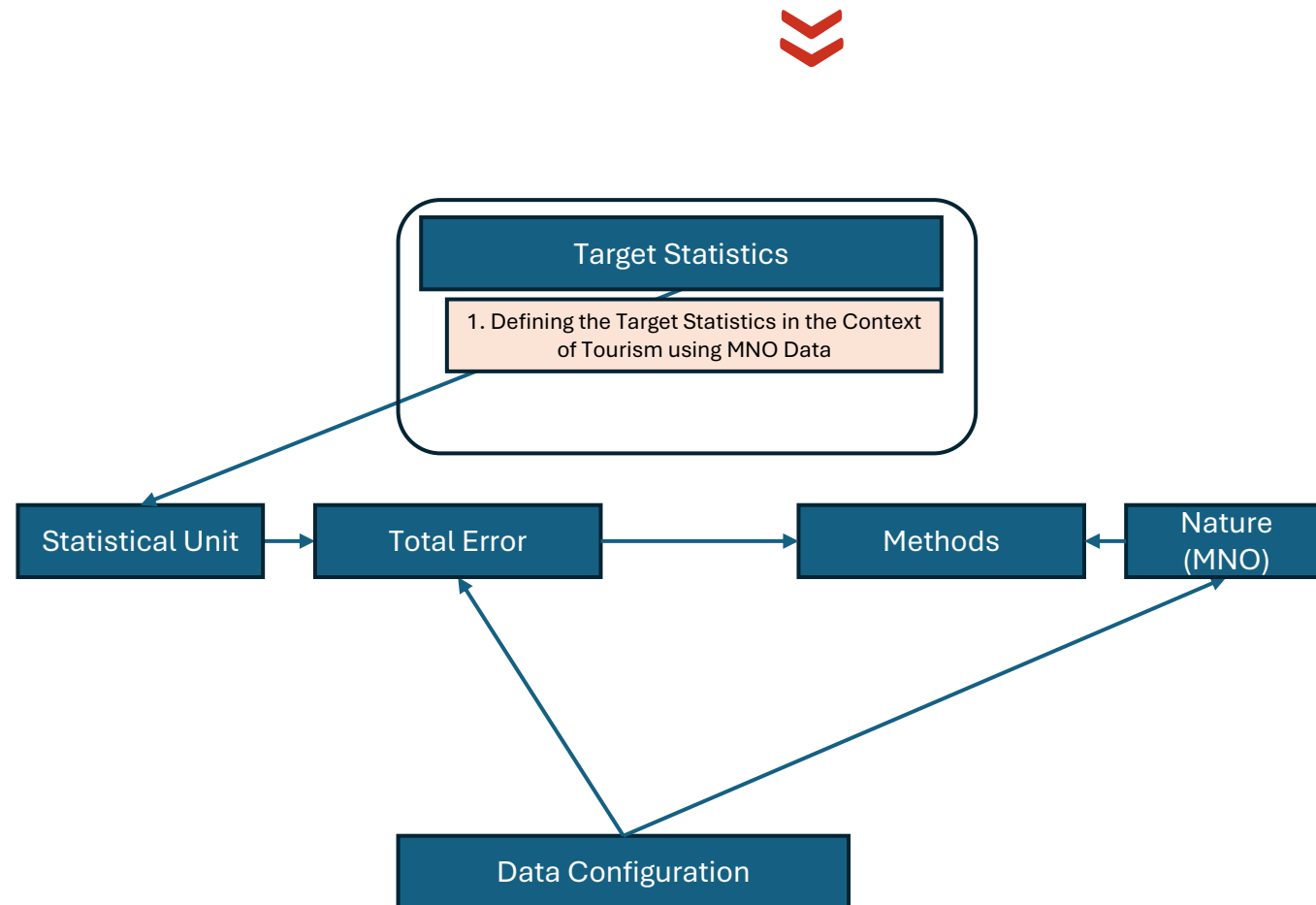
Defining the Target Statistics:

- Documenting the Definition Phase - This documentation plays both a methodological and institutional role: it provides a clear reference for those responsible for implementation, facilitates quality assurance and peer review, and supports future revisions or extensions of the statistical production process.



Documenting the Definition Phase:

- **Analytical purpose** **Example** To support regional tourism policy by providing timely estimates of foreign visitor flows at the NUTS-3 level during peak months.
- **Operational definitions** **Example** A foreign SIM card is considered to be associated with a visitor if it is observed in the national territory for at least four consecutive hours in a single day and is not flagged as a regular commuter.
- **Justification for using MNO data** **Example** MNO data allow for daily monitoring of tourism flows with a high spatial resolution, enabling subnational indicators not feasible with quarterly survey data or administrative data on accommodation.
- **Assumptions and limitations** **Example** One key assumption is that a foreign SIM corresponds to a foreign visitor, although this may not hold in cases such as long-term foreign residents or device sharing



Lost without a signal?

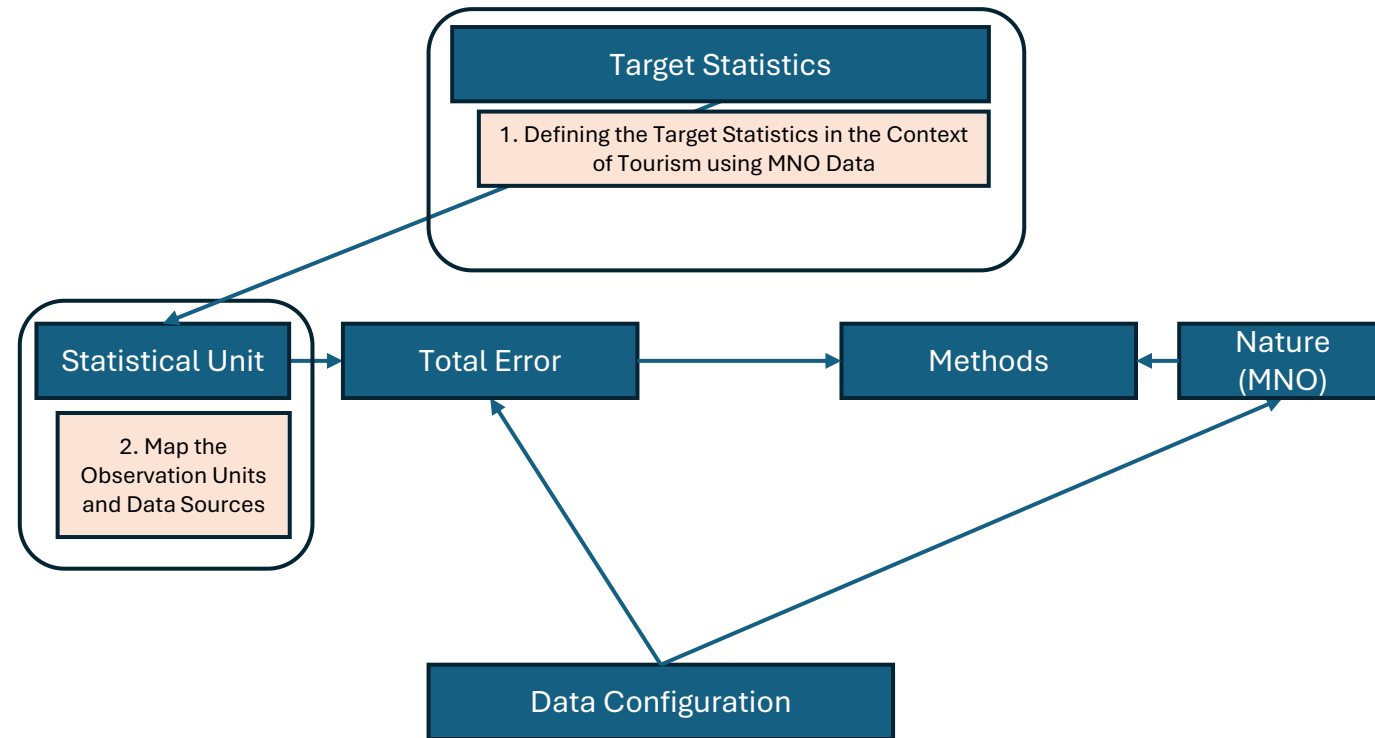
Roadmap for Total Error Framework

Roadmap D3.4

Map the Observation Units and Data Sources:

- Observation Units in MNO Data
- Observation Units in Non-MNO Data
- Mapping Relationships Between Observation Units

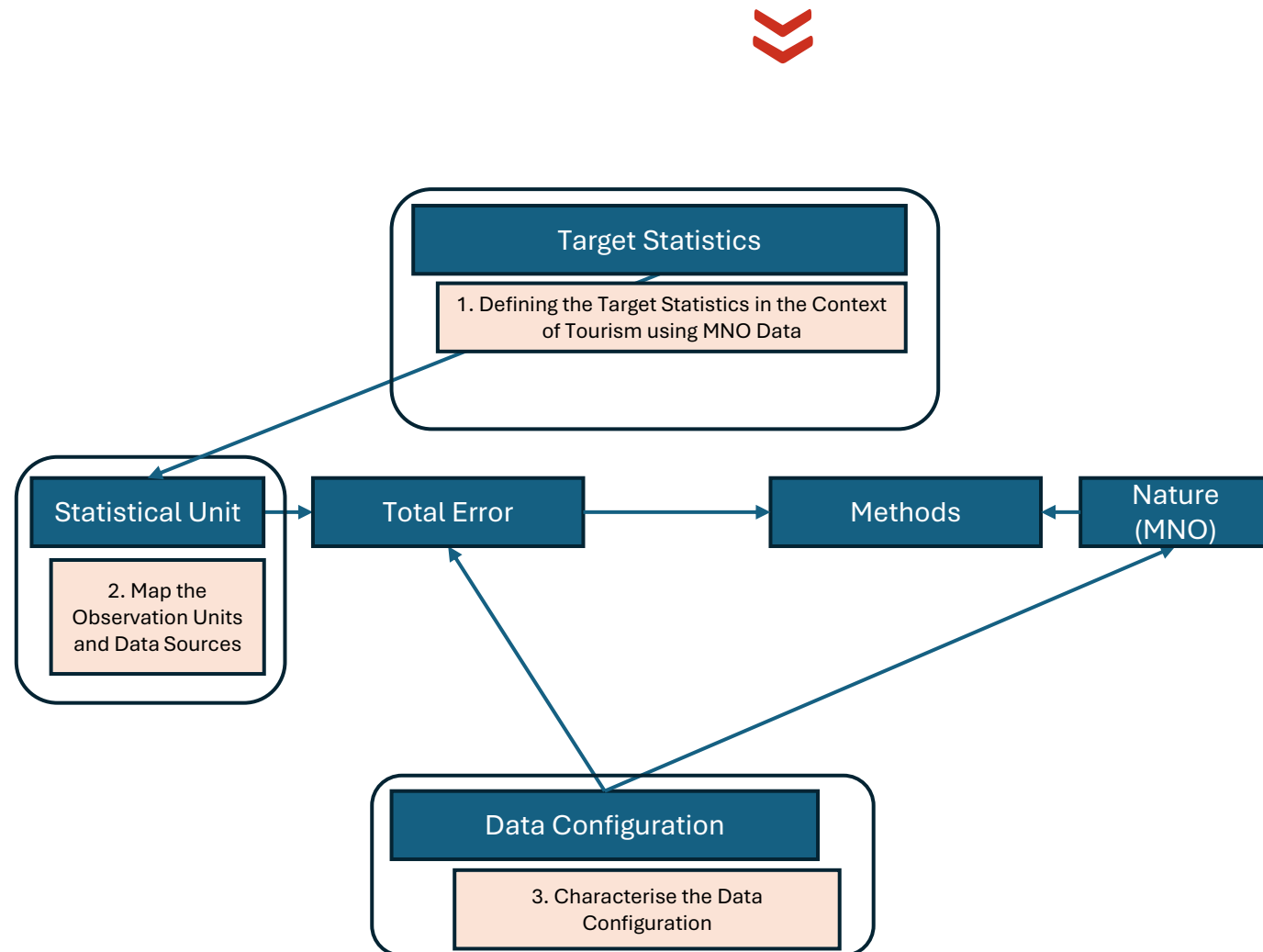
The purpose of this Step is to describe the observation units present in each source, assess how they relate to the statistical units defined, and identify any alignment discrepancy that may lead to misrepresentation, undercoverage, or bias.



Characterize the Data Configuration:

- Granularity (Technical and Analytical)
- Periodicity: Temporal Resolution (Raw Event Timestamps vs. Aggregated Intervals and Event-Based vs. Discrete Time-Series Data)
- Coverage and Scope (explicit or implicit inclusion criteria like only active devices, market share and client base influence the representativeness of the data provided by each MNO, relevant subpopulations of interest)
- Content and Use (Nearly All Statistical Variables from MNOs are Inferred or Derived so what is the Intended Role: Target vs. Auxiliary)
- Data Quality and Access Constraints (Data Reliability and Access, Privacy Transformations)

This Step helps to align the methodological requirements with the technical realities of data availability.



Lost without a signal?

Roadmap for Total Error Framework

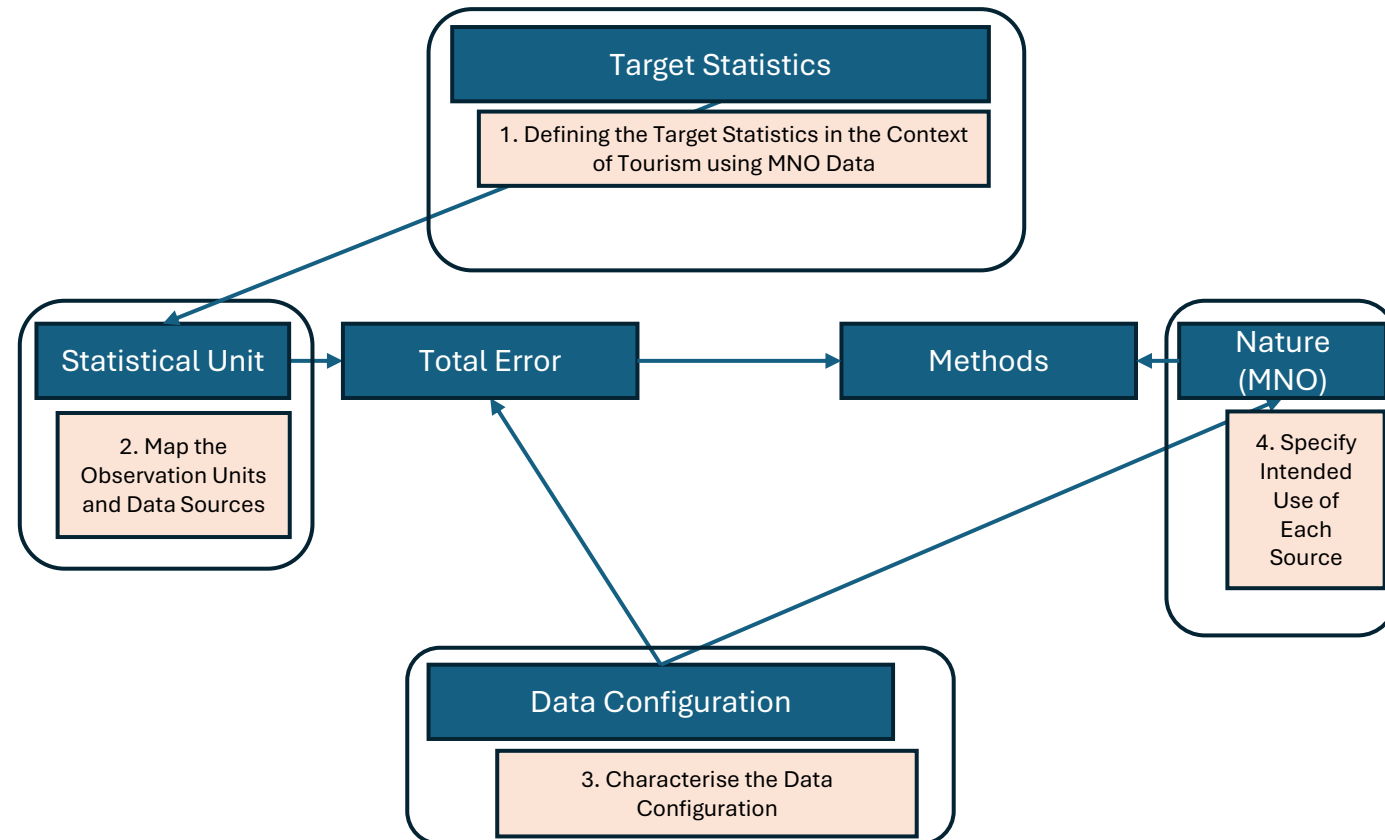
Roadmap

D3.4

Specify the Intended Use of Data:

- Adjusting MNO total by survey-estimated factor
- Sample survey weighting calibrated to MNO total
- Modelling non-MNO variables using MNO features
- Adjusting MNO totals for user selection effects

The purpose of this Step is to specify the methodological approach, by which the MNO and non-MNO data can be combined to produce the target statistic





Roadmap

D3.4

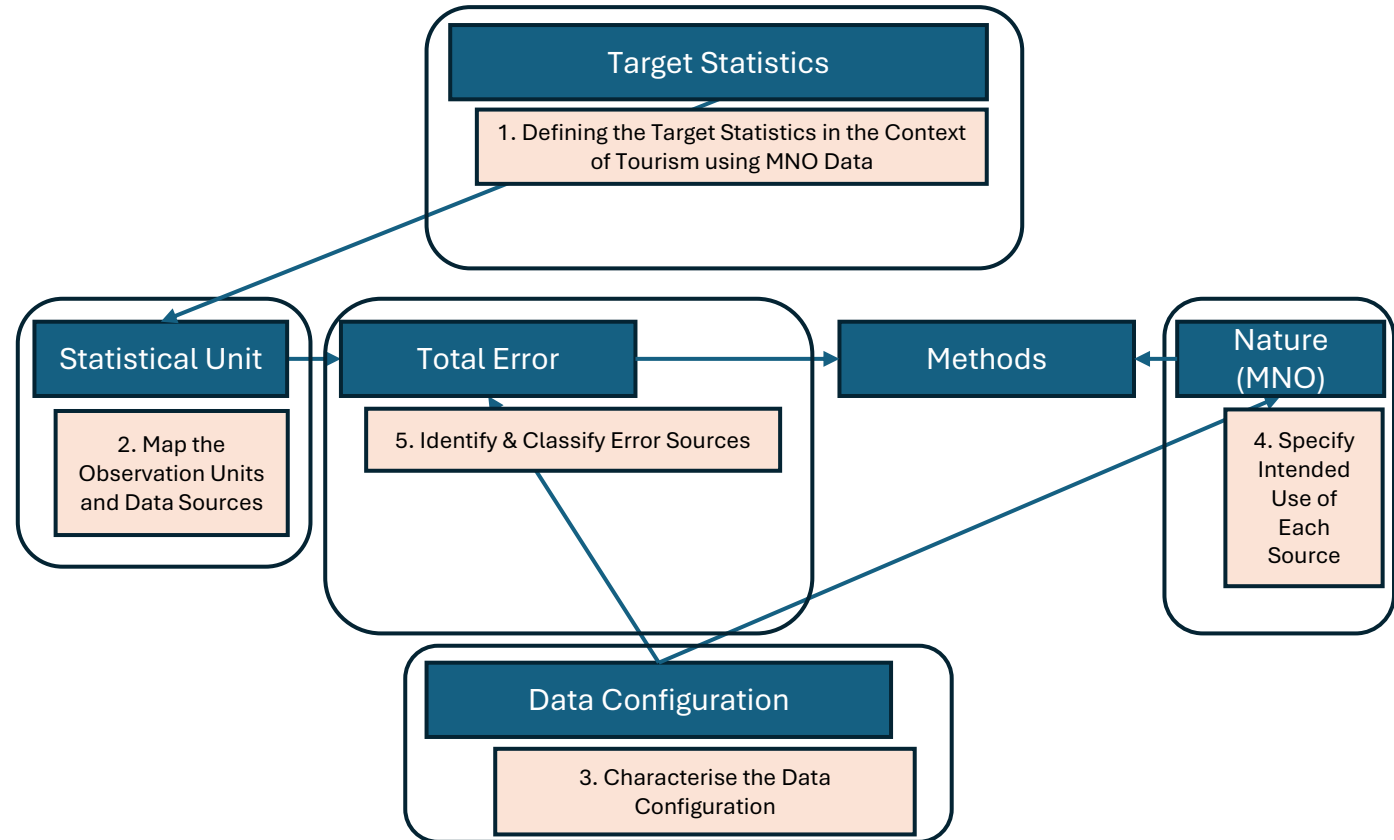
Lost without a signal?

Roadmap for Total Error Framework



Identify and Classify Error Sources - Given the methodological approach at Step 4, one should now systematically identify the potential sources of error introduced by each data stream. Therefore, each error type must be evaluated:

- in the context of its specific source (e.g. MNO logs, survey data),
- with attention to how it may affect the final indicator, and
- reference to statistical principles such as accuracy, coherence.



Roadmap

D3.4

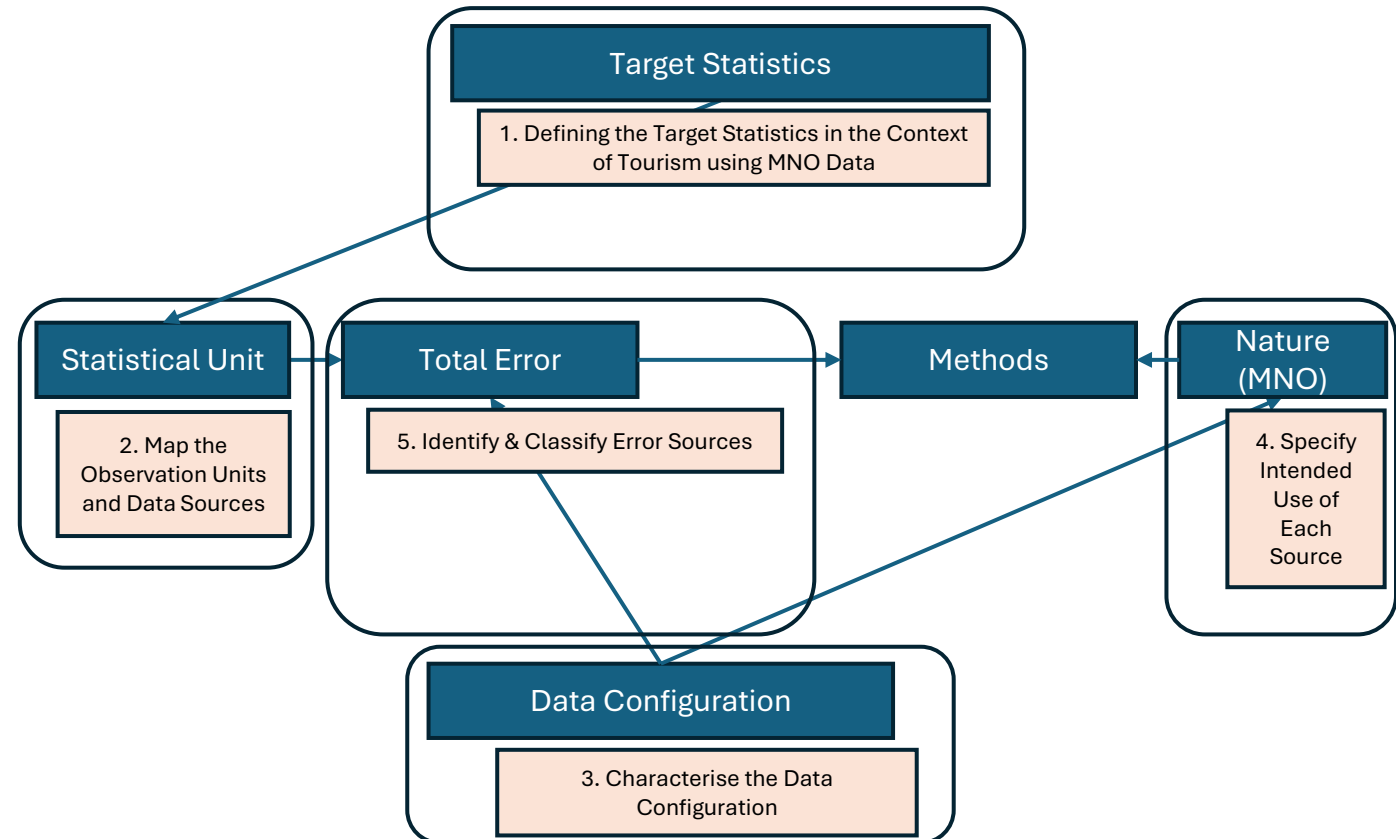
Lost without a signal?

Roadmap for Total Error Framework

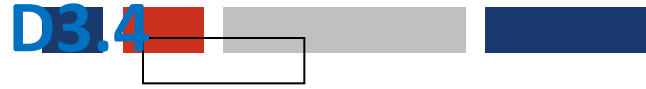
We list some common errors:

- Coverage Error
- Measurement Error
- Misidentification
- Nonresponse Error
- Processing Error or Constraints
- Linkage Error
- Model or Algorithm Error

Always with examples: A tourist classification model trained on summer data may perform poorly for winter visitors, inflating off-season tourism estimates.



Roadmap



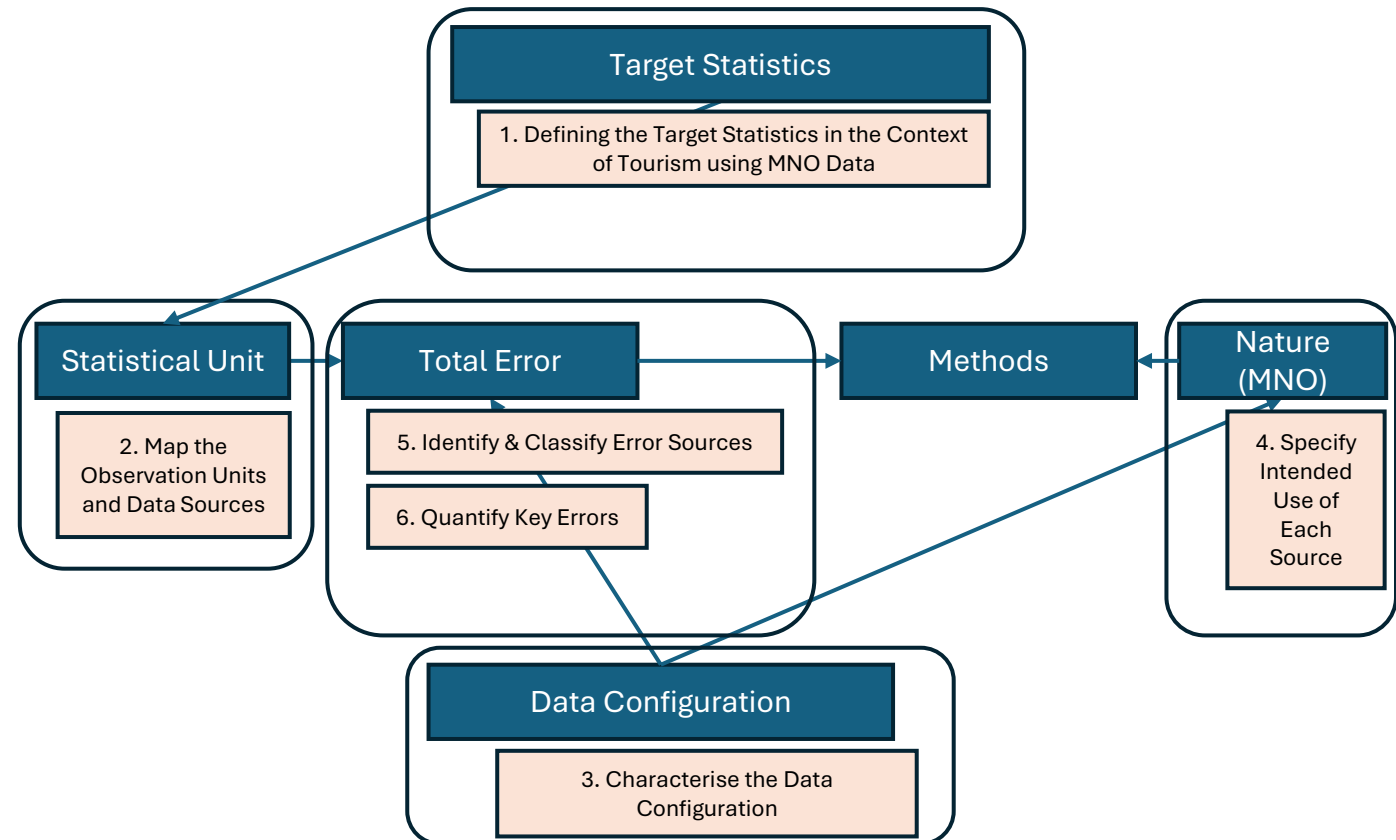
Lost without a signal?

Roadmap for Total Error Framework

Prioritise the Key Errors (If Possible) :

- Sample survey weighting calibrated to MNO total
- Modelling non-MNO variables using MNO features
- Adjusting MNO totals for user selection effects

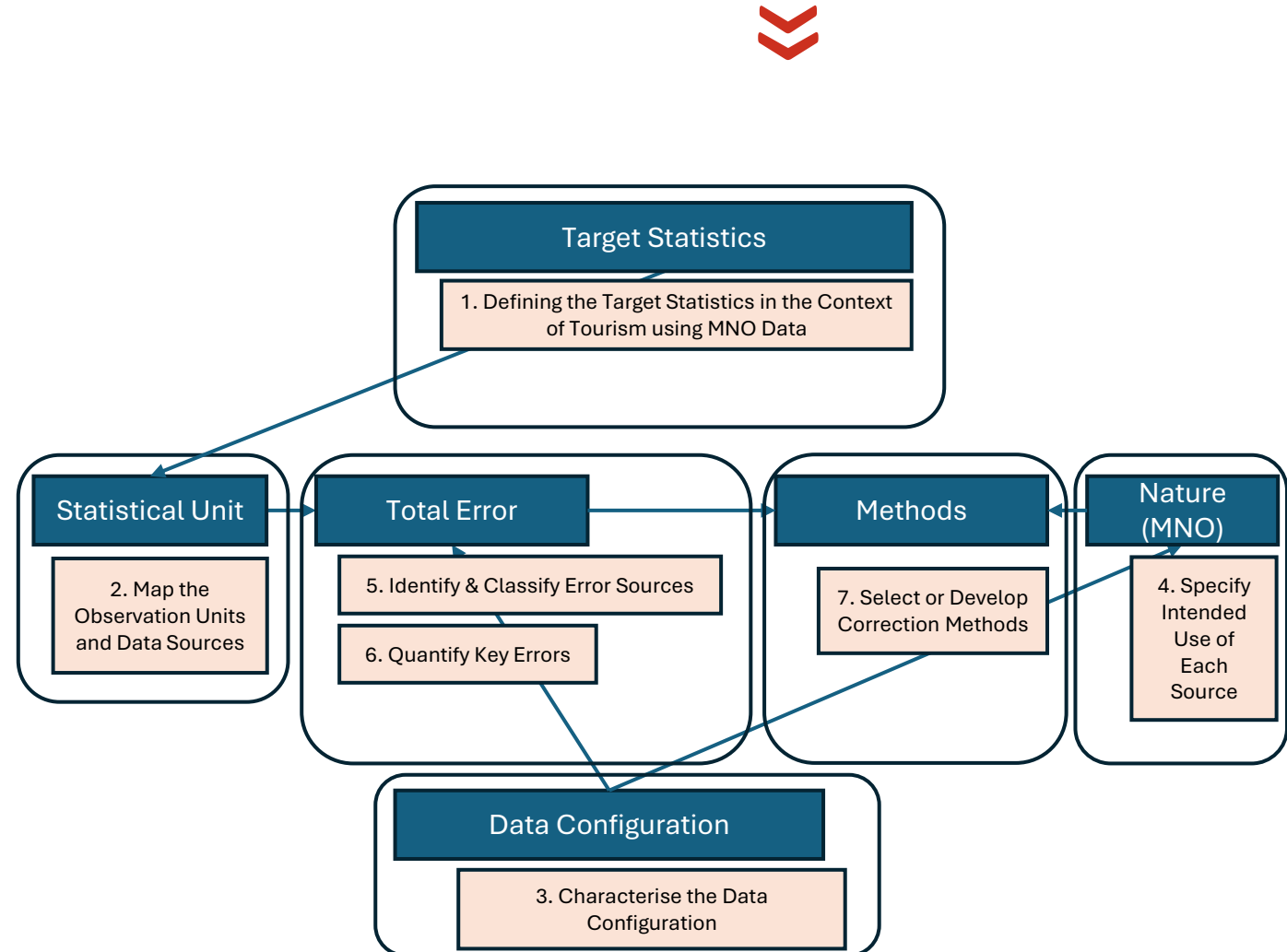
Once potential errors have been classified and their causes clearly understood, the next Step is to assess their impact on the target statistics. Ideally, to determine the direction, magnitude and variability of a given error, one would fix the other aspects of the data...



Implement, Develop Estimation Methods:

- Sample survey weighting calibrated to MNO total
- Modelling non-MNO variables using MNO features
- Adjusting MNO totals for user selection effects

MNO-MINDS D3.2 presents a range of estimation methods for combining MNO macro data with various types of non-MNO data, which are illustrated and discussed for a number of application scenarios. The relevant software tools are documented in MNO-MINDS D3.3.



Lost without a signal?

Roadmap for Total Error Framework

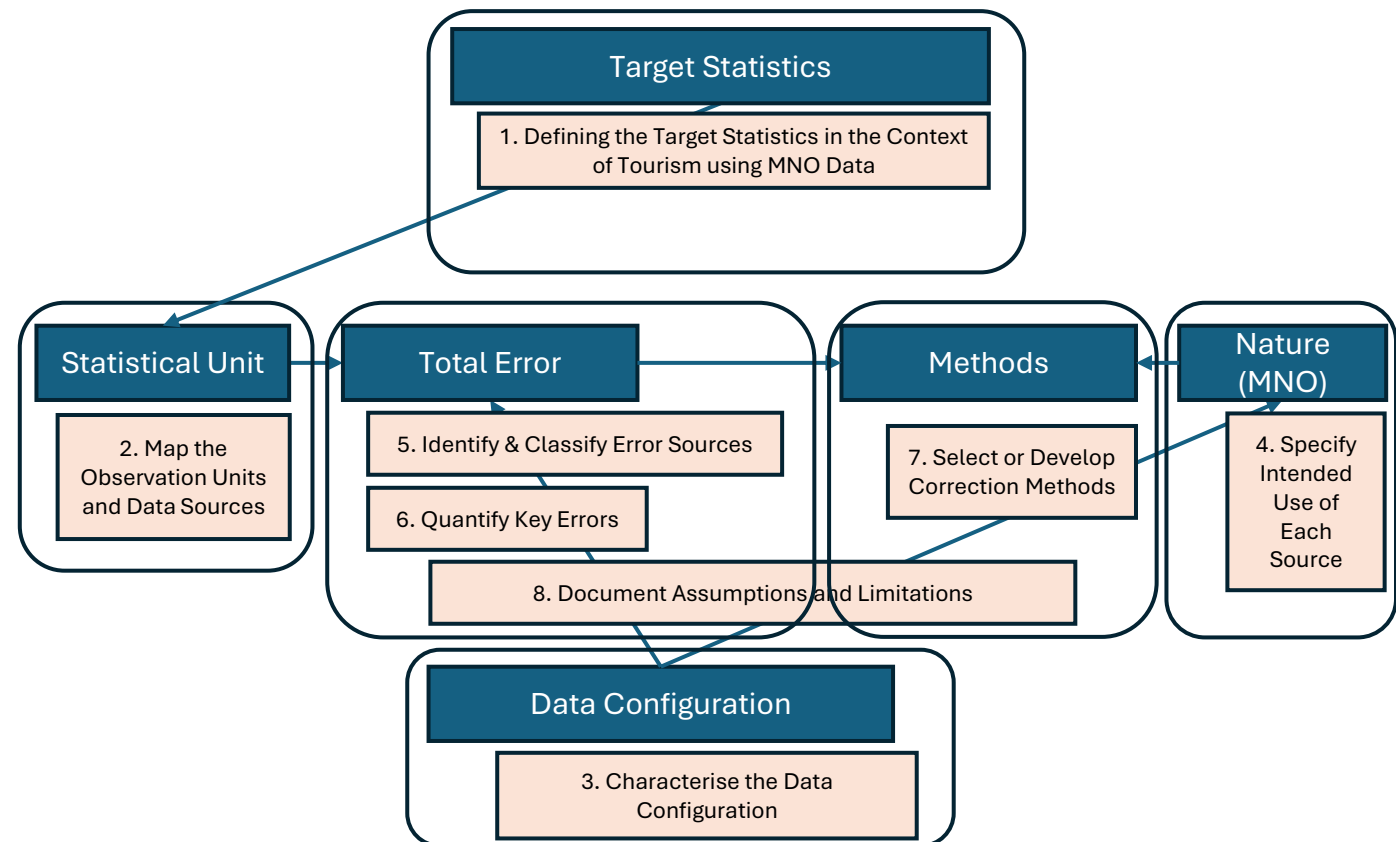


Roadmap

D3.4

Document Assumptions and Limitations:

- Selection Error Adjustment
- Stability of Relationships
- Privacy Constraint
- Temporal and Spatial Granularity
- Dependency on non-MNO Data



Lost without a signal?

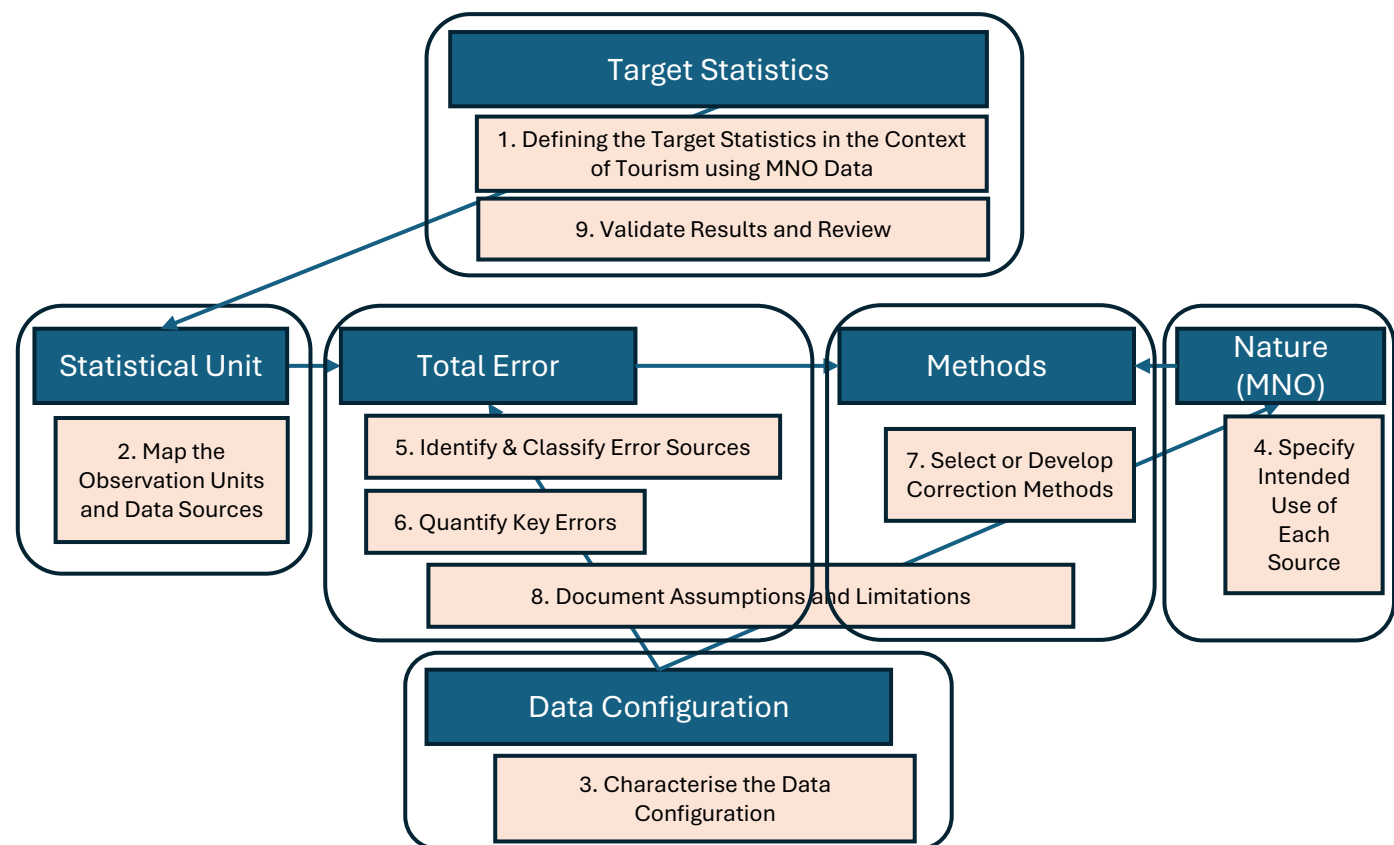
Roadmap for Total Error Framework



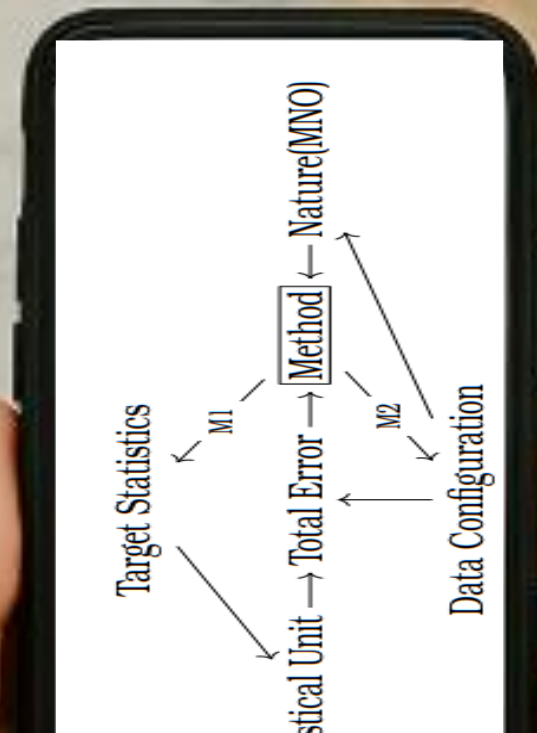
Validate and Review:

- External Validation
 - By alternative statistics – focus on coherence
 - Auditing to assess the error of the final estimates – focus on accuracy
- Review (Peer review within the ESS, User review or consultancy)

Continuous Quality Improvement is focused on incremental enhancements over time. This is the aim of the final Step in the Roadmap, which may lead to iterations of the previous Steps.



ROADMAP FOR TOTAL ERROR FRAMEWORK Questions?



Thank you for your attention!