PAM Time to Failure Transformations Module

Introduction

The Time to Failure Transformations module merges the input data files, prepares the data in the merged file and carries out exploratory data analysis for the analysis and modelling modules. Even though the data available varies between organisations so that the work carried out in the module depends on the organisation, the output of the module is always the same – a master data file that is the input database to the subsequent modules. Failure to prepare the data correctly will lead to an incorrectly structured database and so to invalid results and conclusions.

Input Data

The data available depend on the organisation, industry and asset type, and can be classified as asset data, asset maintenance and failure data, and external data. Table 1 shows example input data.

Table 1

Data Source/Type	Examples
Asset register (asset static data)	Type, use, function, installation date, manufacturer, design specification *
Asset maintenance and failure history	Maintenance and failure data with dates *
Asset operating history	Time in use, number of asset starts *
Asset status	Asset criticalities and redundancies **
Asset costs	Asset maintenance and replacement costs **
Other costs	Costs resulting from asset failure **
Maintenance depot	Location, size of workforce *
Catchment	Type (rural/residential/urban/industrial), population, holiday location, inland/coast *
Season	Weather (rainfall, temperature), month *
Extreme events	Flooding, extreme temperatures *

* candidate predictor variables ** factors applied to models

Output File

The output file is the main input file for the analysis and modelling modules. Since it has one record for each asset-intervention combination, the number of records for each asset is the number of *distinct* maintenance interventions it had (see the section 'Overlapping and Nested Interventions' below).

Data Issues and Work Carried Out

Difficulties in accessing data can arise if the input files are stored in legacy systems that cannot be read easily. Fortunately, this is becoming less of a problem as organisations update their systems and adopt much more rigorous data governance procedures. Other problems can be caused by data stored in silo spreadsheets. Not only are spreadsheets not databases (using spreadsheets as databases can cause data problems), a common problem with silo spreadsheets is that since they are owned by individual people (by definition) and not shared with other people, they tend to be poorly documented and have their own data definitions and formats, so making them difficult for other people to understand and link to other files. Spreadsheets may be appropriate for doing individual local work but they should not be used for storing master corporate data or as substitutes for analytics software.

The work carried out in the module to create the master data repository is described in *Data Preparation, Exploratory Data Analysis and Predictive Asset Management* in PAM Introduction.

Classifying Maintenance Interventions

The asset survival model in **PAM** (see *Asset Survival Models Module* in PAM Modules) requires the event whose failure is being modelled to be defined as a binary variable. The event of interest in **PAM** is asset failure, and so **PAM** defines failure or terminal interventions as repair, refurbishment and replacement interventions, and all other interventions, for example adjustment, lubrication and reset, as non-terminal interventions. This classification enables an understanding to be gained of how an asset's maintenance and failure history affect its current risk of failure.

In addition to defining interventions as terminal or non-terminal, interventions should also be defined as proactive or reactive. Combining the two classifications leads to a new two-by-two classification for interventions as proactive non-terminal, proactive terminal, reactive non-terminal and reactive terminal.

Proactive maintenance includes scheduled maintenance, and some planned short-term and long-term maintenance depending on the priority or urgency of the work. Reactive maintenance includes all reactive work, and planned short-term and long-term maintenance that is very urgent or of high priority.

These definitions of proactive and reactive maintenance interventions must be enhanced in very particular circumstances by considering the sequence of interventions and the periods between them.

Overlapping and Nested Interventions

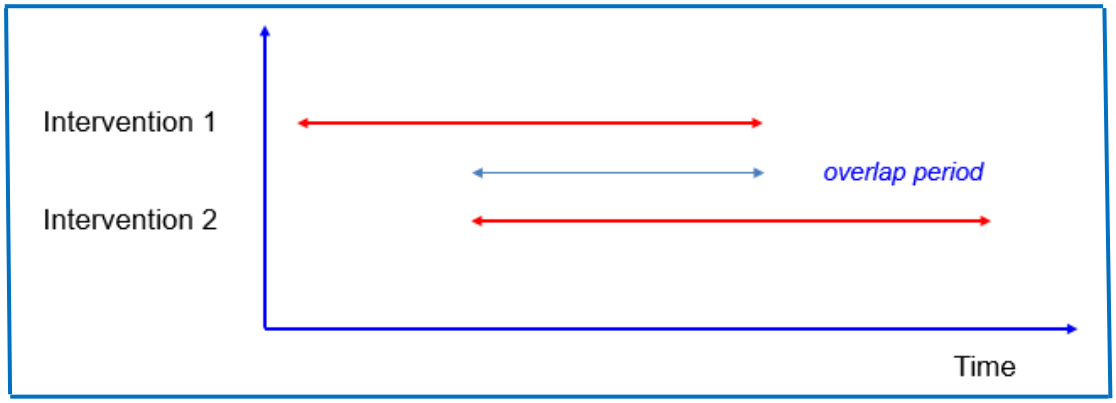
Since **PAM** models the risk of asset failure as a dynamic phenomenon rather than as a static phenomenon, it is essential that the maintenance and failure records of each asset are in the correct chronological order in the output file of the module (the order of the assets is not important). Restructuring and reordering the records for each asset into the correct order is a long, complex and iterative process because the records can have two very serious problems: overlapping interventions and nested interventions (described below). Since the occurrence of nested and overlapping interventions in maintenance databases depends on how the data were recorded, it is possible for the databases to show that an asset had at least one intervention of either type or both types at the same time whereas the input database to the module requires that each asset had one intervention of whatever type at a time.

The run times of the procedures for deduping overlapping and nested interventions can be a large proportion of the total module run time.

Overlapping Interventions

Two interventions for the same asset overlap when the second intervention begins between the start date and end date of the first intervention, and ends after the end of the first intervention. Figure 1 shows the overlap period for two overlapping interventions (assets can have more than two overlapping interventions).

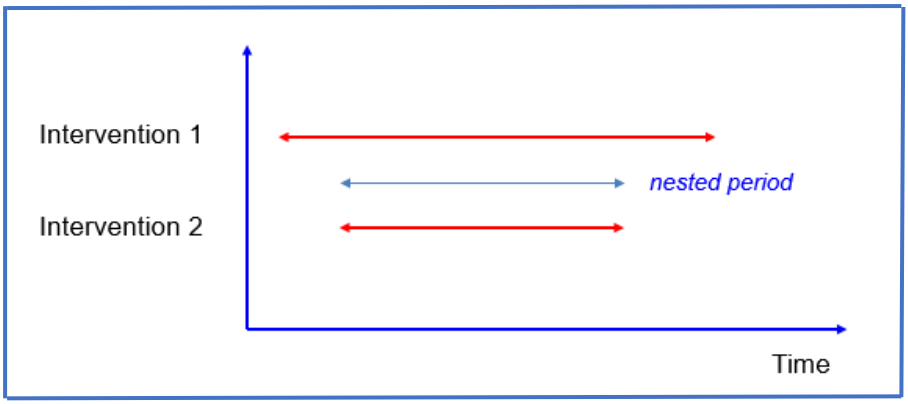
Figure 1



Nested Interventions

Two interventions for the same asset are nested when the second intervention begins between the start date and end date of the first intervention, and ends before the end of the first intervention. Figure 2 shows the nested period for two nested interventions (assets can have more than two nested interventions).

Figure 2



Nested and overlapping interventions are removed separately. The algorithms depend on:

- if the interventions are of the same type (all terminal or all non-terminal) or of different type (terminal and non-terminal)
- if nested interventions are embedded in overlapping interventions, and vice-versa
- the chronological order of the interventions
- the maximum level of the overlapping
- the maximum depth of the nesting
- the minimum interval between interventions for them to be designated as different interventions.

The result of deduping a set of nested and/or overlapping interventions is an intervention that captures the true nature of the underlying interventions and does not coincide with any other interventions.

Interventions of the Same Type

Consider the following scenarios:

- non-terminal interventions that overlap
- non-terminal interventions that are nested

- non-terminal interventions that overlap and are nested
- at least two non-overlapping non-terminal interventions within a short period
- terminal interventions that overlap
- terminal interventions that are nested
- terminal interventions that overlap and are nested
- at least two non-overlapping terminal interventions within a short period.

The example below considers some of these scenarios and raises some of the questions that must be answered when nested and overlapping interventions of the same type are transformed into distinct interventions.

Simplifying Overlapping and Nested Interventions (Same Intervention Type)

Consider the following interventions:

Intervention 1: 26 May to 30 May
Intervention 2: 5 June to 3 July
Intervention 3: 12 June to 20 June
Intervention 4: 1 July to 10 July

Intervention 1 is distinct; intervention 3 is nested in intervention 2; interventions 2 and 4 overlap.

- Interventions 2 begins very shortly after the end of intervention 1.
  - Does this dynamic proximity mean that intervention 2 is part of intervention 1?
  - If they are to be combined, how is it to be done and what are the start and end dates of the combined intervention?
- How should intervention 3 be combined with intervention 2?
  - What are the start and end dates of the combined interventions?
- How should intervention 4 be combined with intervention 2?
  - What are the start and end dates of the combined interventions?

Interventions of Different Types

Consider the following questions:

- what is the type (terminal or non-terminal) of the intervention formed by combining interventions of different type?

- which interventions have priority in combinations of overlapping and nested interventions, and terminal and non-terminal interventions when deduping them?
- if at least one non-terminal event is followed shortly by at least one terminal event, are the non-terminal events warnings of the subsequent terminal events?
- if at least one terminal event is followed shortly by at least one non-terminal event, are the non-terminal events part of the terminal events?

The example below considers these questions when interventions of different types are transformed into distinct interventions. Simplifying interventions of different types is more complicated than simplifying interventions of the same type.

Simplifying Overlapping and Nested Interventions (Different Intervention Types)

Consider the following interventions:

Intervention 1: non-terminal	10 June to 20 June
Intervention 2: terminal	16 June to 24 July
Intervention 3: terminal	1 July to 10 July
Intervention 4: non-terminal	1 August to 3 August

Interventions 1 and 2 overlap; intervention 3 is nested in intervention 2; intervention 4 is distinct.

- How should interventions 1, 2 and 3 be combined?
  - What is the type of the combined intervention?
  - What are the start and end dates of the combined intervention?
- Interventions 2 and 4 do not overlap but intervention 4 begins very shortly after the end of intervention 2.
  - Does this dynamic proximity mean that intervention 4 is part of intervention 2 or that they are separate interventions?
  - If they are to be combined, how does it affect the intervention formed by combining interventions 1, 2 and 3?

