A Guide for Testing ZIMS PMx Exports

Sara Sullivan and Katelyn Mucha

Please note that this guide is based on PMx Version 1.5.20180324 and ZIMS 3.0 current to 14 April 2018.

SCTi
Software for saving species;
Species Conservation Toolkit Initiative

SPECIES 360

Global information serving conservation.

CONTENTS

OVERVIEW

A Few Reminders

Before exporting your studbook data to PMx:

- Compare the size of your living managed populations between legacy and ZIMS studbooks before exporting to PMx.
  - *Nearly all Managed N differences that we have evaluated have been due to users applying incorrect filters, out of date association lists, or one studbook being more up to date than the other.*
- Make sure your export filters are the same before comparing projects.
- If applicable, apply the same assumptions when exporting from the legacy and ZIMS studbooks.
  - *Note that this means PopLink users must enter MULTs into the *.ped file manually if wanting to include probabilities for multiple potential parents.*
- Check that you are using the most updated version of PMx.
- Always use recently exported ZIMS files.

What to Do When You Have a Question

If you do all of the above and still find differences that concern you when comparing PMx projects created from legacy and ZIMS exports, we request that you take the following actions:

1. Investigate the underlying cause of the difference.
   - Note that some improvements over SPARKS and PopLink exports have been made to ZIMS exports. Therefore, some differences should be expected. Rather than focusing simply on the question of "do the values differ," ask whether or not "the values make sense" using your knowledge of the underlying theory and the software.
2. If you cannot identify the underlying cause or are able to identify it but think it is a bug, please send the following materials to help@vortex10.org (you can also send your questions on the PMxpert Listserv) and support@species360.org:
   - PMx projects (if comparing projects, send both)
   - Exports used to create the PMx projects
   - Copy of the legacy studbook, if applicable
   - Summary of the issue, including what you have done to investigate already, the version of PMx being used, and the date the ZIMS files were exported
   - Screenshots that help to illustrate the issue
   - Example individuals who you know or suspect demonstrate the issue

ENHANCEMENTS

Recent Enhancements

PMx

A number of updates have been made, and these are all reported in the change log. PMx is typically installed on your computer in the default location C:\Program Files (x86)\Species Conservation Toolkit Initiative\PMx. You can find the PMxChangeLog.doc here or any alternative location where you installed PMx on your hard disk. I will highlight specific changes that may be particularly relevant to ZIMS testing:

17 March 2018 (version 1.5)

- Fix made so that Demography tables are updated when a DemSelected is changed. (Previously, Demography was updated only when the project was re-opened.)
- When MULT probabilities add to less than 100%, user is now provided the option to readjust so that it does add to 100% or fill in the gap with UNK. (Previously, UNK was always added.)
- Default move records from birth and death dates will be used if there are no move records provided for some individuals.

> Note also that it has been discovered that PMx (subsequent to version 1.4.6) will not run under Windows XP, because some components of Windows10 cause programs compiled for use in Windows10 to generally not be compatible with Windows XP. To our knowledge, PMx will run under Windows 7, 8, or 10. It might or might not run under Windows emulators on a Mac, depending on which version of Windows is being emulated.

20 September 2017 (version 1.4.8)

- Individuals whose move records show them as neither Alive nor Dead (i.e., they are Lost, Released, or Exported) are now flagged in the Individual table as Lost = true, and their ages are shown as missing values ("---"). They might be coded as Alive = true or Alive = false, depending on what was specified in the primary input file, but they will be treated as no longer in the population. (Previously, their age would have been shown as the age if they were alive as of the Analysis Date, although they always were treated as not in the population.)

6 September 2017 (version 1.4.7)

- Unknown parents can now be coded as "UND" (for "undetermined") as well as "UNK". This change was made for compatibility with the coding in exports from the ZIMS Studbook Module.

17 July 2017

- A bug was in the last release that often would have prevented PMx from opening a file that was saved in that release. This bug has now been fixed.

ZIMS

30 January 2018

- Extra line in census header removed (this was causing the census to start at year 0)

14 April 2018

- Export filter for dates fixed and can now be read by PMx (this was causing PMx to default to the date the project was created as the analysis date)

In development – estimated to be released 27 April 2018

- Move records in the demographic.csv and genetic.csv are now reported in chronological order for individuals with multiple moves

Future Enhancements

PMx

- Update the PMx code to recognize ZIMS for Studbooks in the .csv file headers
- Update the PMx code to recognize Indeterminate/Undetermined birthdate estimates

> Note that once PMx can accommodate the above two changes and Species360 releases updated exports, users MUST update PMx in order to create PMx projects from ZIMS exports.

- Remove extra age classes in the age distribution and life tables
  - Impacts:
    - Number of age classes displayed in the life tables and age distribution
    - Age Outlier reported on the Overview Screen
  - This is not a new issue for PopLink Users. When PMx is determining how many age classes are needed, it is incorrectly including ages of LTF animals as of the analysis date rather than their exit date. This also affects the *Age Outliers* reported on the Overview screen. However, this should not be impacting life table calculations.
- Calculate census and census detail graphs within PMx from imported move records

ZIMS

- Add additional columns to the .ped file export including:
  - House name, UDFs, GAN, tag, band, transponder, birth type, birth location, contraception, taxonomy, estimated conception date, regional association
    - Note: once contraception status is added to the export, PMx will calculate the animal's reproductive status from this information.
- Rename .ped file to .ZIMS file

KNOWING YOUR FILES

Please note that this guide is based on PMx Version 1.5.20180324 and ZIMS 3.0 current to 14 April 2018.

Introduction

At minimum, you must import a *Primary Input file*, which provides the pedigree data to PMx. The Primary Input file from SPARKS contains both pedigree and demographic data, while the Primary Input file from PopLink or ZIMS contains only pedigree data.

	SPARKS 1.6	**POPLINK 2.4**	**ZIMS 3.0**
PRIMARY INPUT	exchange.csv	*.ped	ped.ped
DEMOGRAPHIC INPUT	N/A (included in exchange.csv)	*.prns OR *.demog.csv, *genetic.csv	demographic.csv, genetic.csv
CENSUS	pmxcens.csv	exchcens.txt	exchcens.txt

Frequently Asked Questions

Why does ZIMS export pedigree and demographic data in separate file formats?

Exporting pedigree and demographic data into separate files formats (.ped & .csv) provides the following **benefits**:

- Users can set different demography and genetic filter criteria, if desired, when exporting data for analysis in PMx.
- The demographic.csv and genetic.csv can contain multiple move records per individual, which is not currently possible in the SPARKS exchange.csv.
- Users can more easily manipulate data when troubleshooting or if analyzing data within the exports outside of PMx.

Although ZIMS for Studbooks exports data to PMx in a similar format to PopLink, ZIMS export files differ from both PopLink and SPARKS export files in several ways. Use the following icons to identify information throughout this document that is relevant to PopLink users, SPARKS users, or both.

 This affects PopLink users.

 This affects SPARKS users.

 This affects both SPARKS and PopLink users.

The Primary Input File

Current Data Fields

	CURRENT DATA FIELDS
SPARKS 1.6 (EXCHANGE.CSV)	ID:ID;Sire;Dam;Sex;Selected;BirthType;ReproStatus;Living;DOB;DOBest;DOD; DODest;DeathCodes; Location; LocalID; Social; Program; Country; HouseName; Transpond;TagBand;Tattoo;ColorPhase;Litt erSize; No_Litters;Litter_List; UDFs
POPLINK 2.4 (*.PED)	ID; SireID; DamID; Sex; Selected (T/F); Dead (T/F); Birth Date (blank if unknown); Death Date (blank if not dead or not known); Location; Local ID; Other ID
ZIMS 3.0 (PED.PED)	ID; SireID; DamID; Sex; Selected (T/F); Dead (T/F); Birth Date; Death Date (blank if not dead); Location; Local ID

Frequently Asked Questions

Will ZIMS eventually export additional data fields, such as house name and UDFSs?

Yes, additional columns are currently under discussion with SCTI and will be developed in 2018.

Why does PMx display Lost to Follow Up animals as alive when I create a project from ZIMS and PopLink exports but not SPARKS exports?

PopLink and ZIMS ped files use the field "Dead (T/F)" while the SPARKS exchange.csv uses "Alive (T/F)." LTF animals are coded as Alive = F in SPARKS exports but Dead = F in PopLink and ZIMS exports. For this reason, should expect to see PMx projects created from SPARKS show LTF animals as dead and PMx projects from ZIMS show them as alive.

Why are MULTs and their associated probabilities recorded at the bottom of the ZIMS ped.ped rather than under individual records like the SPARKS exchange.csv?

Unlike SPARKS, MULTs entered in ZIMS must be numbered (e.g., MULT1, MULT2, etc.). When MULTs are numbered in this way, the potential parents of MULTx only need to be identified once in the pedigree file. Currently, all MULTs are identified at the end of the ped.ped. PMx calculations handle these MULT probabilities in the same way as MULT probabilities exported from SPARKS.

Other Minor Changes

The ZIMS ped.ped displays birth and death dates, even if undetermined/indeterminate. In contrast, the PopLink *.ped leaves unknown birth and death dates blank.

The Demographic Input Files

Current Data Fields

	CURRENTLY INCLUDED DATA FIELDS
SPARKS 1.6 (EXCHANGE.CSV)	ID ; DateIn; InEst; InType; DateOut; OutEst; OutType; Location; Notes
POPLINK 2.4 (*DEMOG.CSV AND *GENETICS.CSV)	ID; Sex; Birth Type; Birth Date; Birth Date Est; Entry Date; Entry Date Est.; Entry Age (In Days); Entry Method; Exit Date; Exit Date Est.; Exit Age (In Days); Exit Method; Censored/Complete
ZIMS 3.0 (DEMOGRAPHIC.CSV AND GENETIC.CSV)	ID; Sex; Birth Type; Birth Date; Birth Date Est; Entry Date; Entry Date Est.; Entry Age (In Days); Entry Method; Exit Date; Exit Date Est.; Exit Age (In Days); Exit Method; Censored/Complete

Frequently Asked Questions

Why are the demographic input files in .csv format referred to as "move" records?

Both SPARKS and PopLink users should be familiar with move records. A move refers to an event transaction in which an individual enters or exits the population (as defined by your filters).

E.g., Assume we want to analyze demographic data in ZAA from 1950 to present for an individual born in 1960, exported in 1970, imported back in 1980, died in 1990

Move records would be:
DateIn = 1960, typein = Birth, DateOut = 1970, typeout = Exported
DateIn = 1980, typein = Import, Dateout = 1990, typeout = Death

Why does ZIMS export a demographic.csv and genetics.csv?

Exporting two move files allows you to set different filter conditions for your demographic and genetic analyses, if desired. For example, you may feel that a different set of individuals are a better representation of the demographic or genetic characteristics of the managed population. If you set the same filter conditions for demography and genetics, the demographic.csv and genetic.csv will contain exactly the same data.

> The demographic.csv provides the raw data for PMx to calculate male and female life tables. PMx uses data within the genetic.csv to define the movements of individuals in and out of the genetically managed population. Without this information, PMx would not be able to determine the Managed N correctly.

What is the difference between importing .prn and .csv demographic input files?

The demographic.csv file provides the raw data for PMx to calculate life tables while the PopLink .prn files are text files containing life table statistics already calculated in PopLink.

To date, only minor differences have been observed when comparing life tables calculated in PopLink with those calculated in PMx. You will notice that:

- Because the .CSV contains raw data, you can modify the age class width to different time lengths other than year.
- PMx calculates life tables from the raw data in the .CSV on a daily rather than yearly basis. Therefore, prorating individuals who are at risk for only part of a year is unnecessary, and the number at risk for mortality (Risk Qx) and fecundity (Risk Mx) will be the same when the demographic.csv is imported.

⚠ If you see major differences (e.g., Qx = 0.32 vs. 0.15) when comparing life tables calculated with .PRNs vs the demography.csv, please send your PMx project to help@vortex10.org.

Improvements over Legacy Exports

Use the first known event as the Entry Method and Date

The ZIMS exports have been improved to export all possible information into the PMx export for a more robust analysis. Historically, if an animal's first entry event into the population was unknown, the remainder of the animal's life history was excluded from the .CSVs (even if the animals had a known birthdate). ZIMS exports now use the first applicable entry event without an undetermined date estimate as the *Entry Type* and *Entry Date.*

If all transactions (within the filters) have undetermined date estimates, then ZIMS exports use the first transaction with undetermined date estimates that is within the filter (including birth date, if birthdate is also undetermined estimate) as *Entry Type* and *Entry Date*. These unknown age animals with an entirely unknown transaction histories will be excluded from life table analyses in PMx.

Include move records for unknown age animals

In order to import all relevant move records into PMx, the ZIMS demographic.csv and genetic.csv include data for all selected individuals, including those with undetermined birthdates.

The *demog.csv and *genetic.csv exported from PopLink do not include selected individuals with unknown birthdates. This may have been done in an effort to exclude these individuals from life table calculations. However, PMx already has internal data checks for that. PMx still needs move records for these individuals in order to determine whether or not they remain in the population as of the analysis date. Unknown birthdate individuals being excluded from PopLink .CSVs can cause wrong individuals to be included in the managed N. Lost to Follow Up individuals with unknown birthdates have been particularly problematic. However, this should no longer be a problem when creating a PMx project from ZIMS exports.

Allow multiple move records per individual

It is possible for animals to show in multiple rows of the CSV files if the animal has transferred in/out of the selected institution filters, gone LTF/Return from LTF within the date filters, or was released/recaptured within the date filters. SPARKS exports are unable to record multiple moves for an individual, while ZIMS and PopLink exports can.

Each individuals has one birth transaction

The benefit from this is twofold:

- If an animal in SPARKS or PopLink does not have a birth transaction, they will not show in the PopLink/SPARKS export but will show in the ZIMS export because ZIMS will create a birth transaction if one is missing during migration.
- SPARKS and PopLink pull an individual's birthdate from the master table. If the master birthdate and transactions birthdates are different, then the age at birth will display a negative number in PopLink exports. ZIMS only has one birthdate so this is not an issue

Helpful tip! ZIMS migrates birth events from the transaction grid of SPARKS or PopLink rather than the master table. Since PMx exports from SPARKS and PopLink display an individual's birthdate from the master table, this could cause the birthdates exported from ZIMS to be different.

The Census Input File

	CURRENTLY INCLUDED DATA FIELDS
SPARKS 1.6 (EXCHANGE.CSV)	Year;Total;Male;Female;Unk_sex;Capture;Capture_M;Capture_F;Capture_U; Birth;Birth_M;Birth_F;Birth_U;Stillb;Stillb_M;Stillb_F;Stillb_U;Release;Release_M;Release_F;Release_U;Death;Death_M;Death_F;Death_U;LTF;LTF_M;LTF_F;LTF_U;Import;Import_M;Import_F;Import_U;Export;Export_M;Export_F; Export_U;Wildborn;Wildb_M;Wildb_F;Wildb_U;Captborn;Captb_M;Captb_F; Captb_U;Unkborn;Unkb_M;Unkb_F;Unkb_U;Lambda;Institutions
POPLINK 2.4 (*DEMOG.CSV AND *GENETICS.CSV)	Year; Tot_sex; Male; Female; Unk_sex; Tot_born; Wildborn; Capt; Unk
ZIMS 3.0 (DEMOGRAPHIC.CSV AND GENETIC.CSV)	Year; Tot_sex; Male; Female; Unk_sex; Tot_born; Wildborn; Capt; Unk

Frequently Asked Questions

When will ZIMS exports include census detail data (i.e., number of births, deaths, imports, exports, etc. each year)?

Because it was requested that Species360 prioritize ZIMS exports before the Census Details report in ZIMS was developed, the ZIMS census export currently does not contain the extra fields displayed in the SPARKS census export. However, SCTI has decided that census and census detail data should be calculated within PMx from the move records rather than from an imported census file. SCTI will be working on this update in the coming months.