ERDC/EL TR-12-32

Environmental Laboratory

US Army Corps of Engineers®
Engineer Research and Development Center

Dredging Operations and Technical Support

Preparing Inputs for New Applications of the TernCOLONY Simulation Model

Casey A. Lott, Jonathon J. Valente, Richard A. Fischer, and Colin J.R. Sheppard

December 2012

Approved for public release; distribution is unlimited.

ERDC/EL TR-12-32
December 2012

Preparing Model Inputs for New Applications of the TernCOLONY Simulation Model

Casey A. Lott

Coastal and Waterways Program Coordinator
American Bird Conservancy
1209 Shenandoah Drive
Boise, ID 83712

Jonathon J. Valente and Richard A. Fischer

Environmental Laboratory
U.S. Army Engineer Research and Development Center
3909 Halls Ferry Road
Vicksburg, MS 39180

Colin J.R. Sheppard

Lang, Railsback, and Associates
250 California Ave.
Arcata, CA 95521

Final report

Approved for public release; distribution is unlimited.

Prepared for U.S. Army Corps of Engineers
Washington, DC 20314-1000

Abstract

The TernCOLONY simulation model is an individual-based model of Least Tern reproduction that was developed to better understand how reservoir operations (and other management activities) affect Least Tern breeding populations on large rivers. This document provides practical guidance for those wishing to implement this model in a new region. Specifically, this document provides guidance concerning how to prepare the tabular inputs that are necessary to import required site-specific data on habitat, flows, and bird populations into the relational database that manages TernCOLONY's inputs and outputs. This document is a technical, step-by-step guide to fulfilling this basic objective. For more information on the model itself, please visit the model's website at http://www.leasttern.org.

DISCLAIMER: The contents of this report are not to be used for advertising, publication, or promotional purposes. Citation of trade names does not constitute an official endorsement or approval of the use of such commercial products. All product names and trademarks cited are the property of their respective owners. The findings of this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

DESTROY THIS REPORT WHEN NO LONGER NEEDED. DO NOT RETURN IT TO THE ORIGINATOR.

Contents

Abstract ii

Figures and Tables iv

Preface v

Unit Conversion Factors vii

1 Introduction 1

Purpose 1

Major inputs 1

2 Step 1: Defining the Spatial Extent and Properties of a New Model Application 3

Background and methodological concerns 3

Naming the management area and reaches within the management area 5

Creating a thalweg point file to specify the spatial extent of the management area 6

3 **Step 2: Developing Sandbar-Specific Time Series Inputs for Peak Daily Flows** 8

Background and methodological concerns 8

Providing peak daily flow inputs 8

4 Step 3: Classifying Annual Flow Inputs into Different Breeding Season Water Year Types 10

Background and methodological concerns 10

General properties of flow inputs 12

Categorizing flow inputs by water year types 12

5 Step 4: Providing Site-Based Inputs that Describe Sandbar Nesting Habitat 13

Background and methodological concerns 13

General properties of habitat inputs 15

Providing inputs for sites 15

Providing inputs for cells 17

6 Step 5: Specifying Inputs that Describe Properties of the Bird Population 18

Background and methodological concerns 18

Specifying prior distributions of nesting terns 19

References 20

Glossary 21

Report Documentation Page

Figures and Tables

Figures

Figure 1. The thalweg point file above provides spatial reference for the entire management area of the Arkansas River below Keystone Dam. 5

Figure 2. The thalweg point file for the Arkansas River, zoomed in to the scale of the channel, near a potential nesting sandbar. 5

Figure 3. Example classification of annual flow inputs to water year type categories from Lott and Wiley (2012). 11

Tables

Table 1. Naming the management area. Users must provide names for each new management area in the Management Area table. 3

Table 2. Naming reaches within each management area. Users may specify reaches within each management area if they desire to summarize information about subsets of sites. 3

Table 3. The Thalweg Point table contains xy coordinates for each point; their distance downstream from the dam; and the reach that they belong to. 4

Table 4. The Thalweg Point Group table links thalweg points with the management area, habitat input set (see below), and the period of time to which they refer. 4

Table 5. Daily flow table. This table contains the actual sandbar-specific peak daily flow data. 8

Table 6. Flow Input Set table. This table provides a list of the different sets of annual flow inputs that are available as peak daily flow time series inputs to the model. 9

Table 7. The Water Year Category table lists all possible water year type categories for each management area/flow input set combination and their frequency within the historic record. 9

Table 8. The Water Year table simply assigns each annual flow input to one of the categories listed in the Water Year Category Table. 12

Table 9. Habitat Input Set table. This table provides a list of the different sets of habitat inputs that must be available as primary inputs to the model. 14

Table 10. Site Type Table. This table lists the types of nesting sites that can be included in the model. 14

Table 11. Site table. 14

Table 12. Data necessary to complete the input table “tblCell” in the IBM database. 15

Table 13. Tern Previous Site table. 18

Preface

The TernCOLONY simulation model is an individual-based model of Least Tern reproduction that was developed to better understand how reservoir operations (and other management activities) affect Least Tern breeding populations on large rivers.

The model was collaboratively developed by American Bird Conservancy (ABC) and Lang, Railsback, and Associates (LRA). Dr. Richard Fischer of the U.S. Army Engineer Research and Development Center (ERDC), Environmental Laboratory (EL), provided feedback throughout model development.

This effort was supported by funding from the U.S. Army Corps of Engineers Dredging Operations and Technical Support (DOTS) and the Dredging Operations and Environmental Research (DOER) programs. The Program Manager for DOTS and DOER at the ERDC is Dr. Todd Bridges, Senior Research Scientist for Environmental Science.

This document is the third in a series of three documents related to the TernCOLONY model. The first document in the above-mentioned series is the TernCOLONY model description (Lott et al. 2012). The TernCOLONY model description follows the "ODD" (Objectives, Design Concepts, and Details) protocol of Grimm et al. (2006) as updated by Grimm et al. (2010) for describing individual-based models. This protocol starts with general and conceptual characteristics; followed by the detail needed to make the model reproducible. The model has been extensively verified and rigorously tested using pattern-oriented modeling approaches (Grimm and Railsback 2005) and standard analytical techniques such as sensitivity analysis and parameter calibration. These tests and procedures are reported in Lott et al. (2012b). TernCOLONY is the first complex individual-based model that is completely accessible via the web at http://www.leasttern.org. Additional documentation of the model's software and code is also available at http://www.leasttern.org.

The habitat and flow inputs underlying the first regional application of TernCOLONY (on the Arkansas River below Keystone Dam) were docu-

mented previously in Lott and Wiley (2012), which described research on Least Tern nesting habitat measurements that was funded by the U.S. Army Engineer District, Tulsa.

The authors would like to thank Merrie Morrison, Danny Cunningham, and David Pashley of American Bird Conservancy for administrative support during this project. Technical review of the manuscript for ERDC was provided by Dr. Richard A. Fischer and Mr. Jonathon J. Valente, EL.

The primary contact for TernCOLONY is Casey Lott of American Bird Conservancy at clott@abcbirds.org. The ERDC points of contact are Dr. Richard A. Fischer (502-315-6707; Richard.A.Fischer@usace.army.mil) or the focus area manager of the Dredging Operations and Environmental Research Program, Dr. Todd Swannack (601-634-2068; Todd.M.Swannack@erdc.usace.army.mil). This technical note should be cited as follows:

Lott, C.A., J.J. Valente, R.A. Fischer, and C.J.R Sheppard. 2012. *Preparing model inputs for new applications of the TernCOLONY simulation model*. ERDC TN DOTS-EXX. Vicksburg, MS: U.S. Army Engineer Research and Development Center. http://el.erdc.usace.army.mil/dots/doer/doer.html.

During the preparation of this work, Antisa Webb was Chief of the Ecological Resources Branch; Dr. Edmond Russo was Chief, Ecosystem Evaluation and Engineering Division.

Commander of ERDC was COL Kevin J. Wilson. Director of ERDC was Dr. Jeffery P. Holland.

Unit Conversion Factors

Multiply	By	To Obtain
acres	4,046.873	square meters
cubic feet	0.02831685	cubic meters
feet	0.3048	meters
inches	0.0254	meters
miles (U.S. statute)	1,609.347	meters
miles per hour	0.44704	meters per second
square feet	0.09290304	square meters

1 Introduction

Purpose

This short document is an explicit, step-by-step guide to preparing the full suite of model inputs that are necessary for any new application of the individual-based simulation model of Least Tern reproduction known as TernCOLONY (Lott et al. 2012, in preparation). It was written for a very narrow audience: those preparing a new application of this model. More general information on the model; its development and testing; model applications to specific management problems; and a running version of the model itself, is available at http://www.leasttern.org. The model is run from a web-accessible graphical user interface (**GUI**), which pulls information from a *Postgres* relational database that stores model inputs and outputs. The database and GUI were both designed so that new applications of TernColony could be rapidly deployed by appending sets of model inputs for new areas into the existing project database. Consequently, this document describes in detail the type of information that should be included in every field of each of the 13 database tables that hold model inputs. An Excel document with multiple worksheets, titled TernColonyInputs.xls, is a companion to this technical report. The Excel file provides empty versions of each of these tables, as well as examples of the type of information that should be included in each field. Using the instructions in this document, those wishing to apply the model in new areas should use this template to provide new input data to the TernCOLONY database administrator (DBA). The DBA then assimilates these new data into the database, updates the copy of the database on the server where the model is stored, and the new application of the model can immediately be run from the existing GUI at http://www.leasttern.org. For this process to be as seamless as possible, users must follow the instructions in this document very carefully. A glossary is provided at the end to define commonly used terms. Terms defined in the glossary are listed in bold type upon their first use.

Major inputs

The model requires five major sets of inputs. This document breaks each of these sets into its component database tables and explains how these are related. Recognizing that each model application will have its own unique set of circumstances, each of the sections on different model inputs

also provides some ideas about HOW the information that is necessary to populate each of these tables might be assembled. The five different sets of model inputs are:

1. Inputs that define the spatial extent of a model application.
2. Sandbar-specific time series inputs for peak daily flows.
3. A scheme for classifying **annual flow inputs** into **breeding season water year types**.
4. **Site**-based sandbar nesting habitat inputs.
5. Inputs describing various properties of the bird population.

This document focuses on formatting the information that is required to run a new application of TernCOLONY into tabular inputs. Lott and Wiley (2012) is a detailed example of the field, GIS, and analytical methods that were used to create model inputs for a single application of TernCOLONY on the Arkansas River below Keystone Dam. However, the details of how flow, habitat, and bird data sources might be acquired or processed prior to being input into this database will depend on data availability, which will be specific to each model application. For example, if LIDAR surfaces are available for an area, these could be used to generate the types of digital elevation models that form the backbone of habitat inputs. If these data are not available, a choice will need to be made about the most efficient way to collect elevation data for model inputs – which will vary by area. In some cases, field topographic surveys, such as the type that were performed in Lott and Wiley (2012) may be feasible. In other cases (e.g., areas like the Lower Mississippi River where the size and abundance of sandbars would make this type of measurement infeasible), LIDAR data collection may be the most cost-effective way to acquire elevation data.

This document describes only the *type* and *properties* of the data that are necessary to run the model and the format in which they must be provided. It is not prescriptive of how these data must be acquired, since there are many good reasons why this might vary by application. Consequently, we recommend preparing application-specific plans for adapting the model to new areas. Application-specific documents should address any **calibration** steps that may be necessary to apply TernCOLONY in new areas (Lott et al. 2012).

2 Step 1: Defining the Spatial Extent and Properties of a New Model Application

Background and methodological concerns

The first version of TernCOLONY was developed for the exploration of Least Tern population dynamics on rivers, when all potential nesting sites are either naturally or mechanically created sandbars. For this application, we defined a "**management area**" as "the complete list of all potential nesting sites for one subpopulation of terns." Other definitions of subpopulations are possible (e.g., all of the potential beach nesting sites in New Jersey); however, the current version of the model is specific to sandbar nesting sites on rivers. If users wish to include off-channel nesting sites near rivers (such as sand-pits, salt flats, industrial areas, or rooftops), or if they wish to apply the model in non-riverine habitats, they should contact the model's developers. Tables 1-4 below (and in the TernCOLONYInputs.xls file) define the properties of a "management area."

Table 1. Naming the management area. Users must provide names for each new management area in the Management Area table.

Table Variable	Description
manCode	Short name for graph axis labels
manName	Longer name for selection menus in GUI
manNotes	Longer text description with details
manCoordinateSystem	xy coordinate system used for spatial locations

Table 2. Naming reaches within each management area. Users may specify reaches within each management area if they desire to summarize information about subsets of sites (e.g., all sites above or below a tributary). This is not mandatory. If no reaches are specified, the database administrator will create a single reach with the same name as the management area.

Table Variable	Description
manCode	Short name for management area
rchCode	Short name for graph axis labels
rchName	Longer name for selection menus in GUI
rchNotes	Longer text description with details

Table 3. The Thalweg Point table contains xy coordinates for each point; their distance downstream from the dam; and the reach that they belong to.

Table Variable	Description
tempThalwegPointID	Unique ID for each point (assign any number)
thgTitle	Name of thalweg point group
rchCode	Short name of the reach associated with each point
thwPointX	X spatial coordinate
thwPointY	Y spatial coordinate
thwMilesFromDatum	Distance from the dam or other datum (in miles, with 5 decimal precision).

Table 4. The Thalweg Point Group table links thalweg points with the management area, habitat input set (see below), and the period of time to which they refer.

Table Variable	Description
manCode	Management Area Code
plHabitatInputSetCode	Habitat Input Set Code
thgTitle	Name for a set of thalweg points

For each management area, users must provide xy coordinates for a set of points (spaced at 100ft increments) along the **thalweg**, which is the deepest part of the river channel. These points should match the coordinate system specified in the management area table (Table 1). This set of points specifies the spatial extent of the management area and is used to locate objects in the model. Figure 1 shows the thalweg point file for the Arkansas River below Keystone Dam management area, which spans from Keystone Dam (red symbol), just upstream of Tulsa, Oklahoma, to the backwaters of Webbers Falls Reservoir further downstream. This point file was created by first tracing a line that followed the apparent thalweg in a set of aerial photos and then converting this line into a series of points at 100ft increments. Alternatively, a line file for the thalweg could be collected by attaching a GPS unit to a boat that would then follow the thalweg across the entire management unit. A number of different routines in ArcGIS, or other GIS software platforms, are available to convert lines into points at evenly spaced increments. Figure 2 shows the thalweg point file, zoomed in at the scale of the channel.

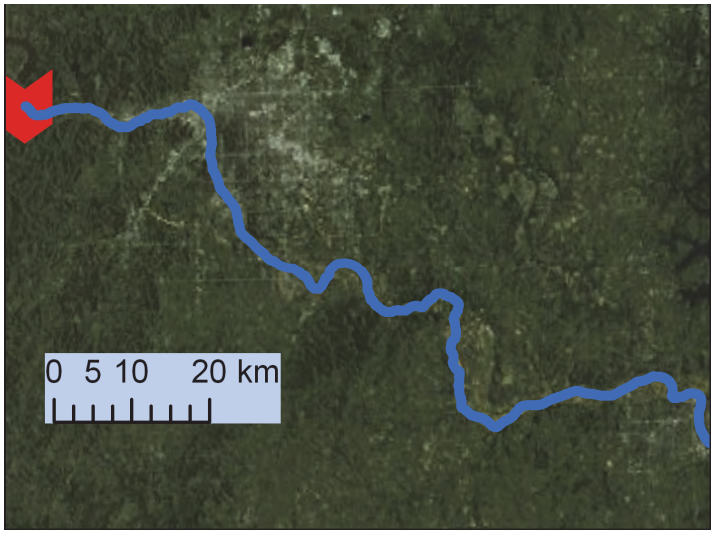

Figure 1. The thalweg point file above provides spatial reference for the entire management area of the Arkansas River below Keystone Dam.

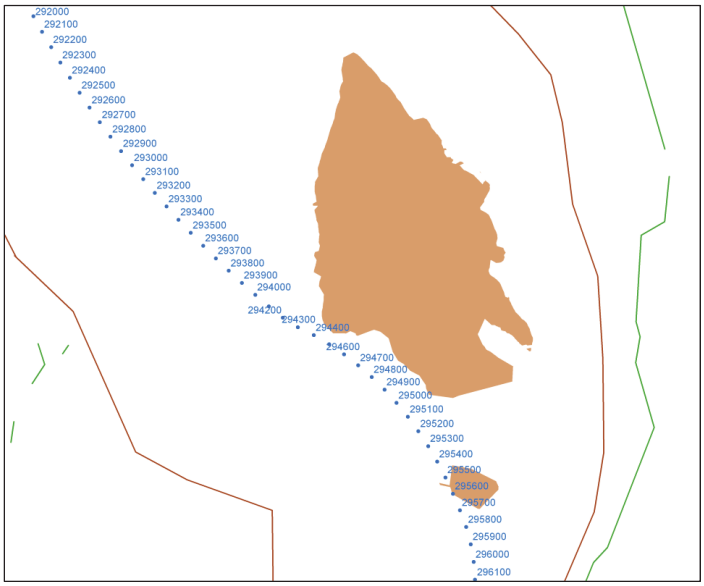

Figure 2. The thalweg point file for the Arkansas River, zoomed in to the scale of the channel, near a potential nesting sandbar (the orange shape). The bank of the active channel is delineated by the brown line and small patches of tall trees and forest edges are indicated by green lines. The blue text indicates the distance of each point downstream from Keystone Dam at the head of the management area (in ft.).

Naming the management area and reaches within the management area

This table introduces a convention that is common in many TernCOLONY database tables: there are three different ways to name things. A short name

field, called "…Code" is used to label data graphs in the GUI. Short names are necessary so that graph axes labels are legible. Entries in this field should be <20 characters. A longer name field, called "…Name" provides a little more detail. This name is the one that is used to select a management area from drop down lists that contain all management areas in the GUI. Entries in this field should be <50 characters. The final naming field, called "…Notes" allows users to specify more detail (of whatever type they'd like) to describe the management area. This information will be available as "additional information" in several areas within the GUI. Entries in this field can be up to 1000 characters. The final field in this table specifies the spatial coordinate system that applies to each of the xy coordinates that are provided in other tables for this management area. Each user should specify a coordinate system based on his/her full name in GIS software packages. The TernCOLONYinputs.xls spreadsheet should be reviewed for example inputs for each of these fields and for example inputs for all other fields that are referenced in this document.

This table introduces another convention that will be applied throughout this document. Fields that are listed in bold and italicized type indicate fields where entries must match exactly the entries in another database table. For example, the entry for the ***manCode*** field in the reach table must exactly match the short name for one of the management areas listed in the ***manCode*** field of the management area table. In the ***manCode*** field within the reach table, users specify the management area to which each reach name applies. The reach table has the same three different types of naming fields for reaches that were present for management areas (with 20, 50, and 1000 character limits).

Creating a thalweg point file to specify the spatial extent of the management area

In the tempThalwegPointID field, users must assign each record a unique numeric ID. The actual values for this code do not matter here, and these values will be reassigned by the database administrator. The entry in the ***thgTitle*** field should match exactly a value from this field in the thalweg point group table (See Table 4). Each point should also be assigned to the reach it is in (e.g., all points upstream of a reach boundary will be in reach "A" and all points downstream of that boundary will be in reach "B"). The entry in the ***rchCode*** field should match exactly a value from this field in the Reach table (See Table 2). The xy coordinates for each thalweg point should match the coordinate system specified in the Management Area

table (Table 1). The value in the thwMilesFromDatum field should indicate the distance (in miles) from the head of the management area (in this case, the upstream end). A number of routines are available in different GIS platforms to assign these values. The authors recommend retaining five decimal places of precision.

Within one management area, users may have more than one set of habitat inputs. For example, habitat inputs may be available for 1994 and 2010. If there has been considerable channel change between these two dates, users may wish to create a new thalweg point file for each time period. This table links different habitat input sets and thalweg point files together. The entry in the ***manCode*** field should match exactly a value from this field in the Management Area table (See Table 1). The entry in the ***plHabitatInputSetCode*** field should match exactly a value from this field in the Habitat Input Set table (See Table 9). The thgTitle field is where users specify a name for each specific set of thalweg points (e.g., "Arkansas River below Keystone Dam 1994" or "Arkansas River below Keystone Dam 2010"). This is how alternative sets of thalweg points for the same management area can be distinguished.

3 Step 2: Developing Sandbar-Specific Time Series Inputs for Peak Daily Flows

Background and methodological concerns

TernCOLONY requires time series inputs of site/sandbar-specific peak daily flows (see Lott et al. 2012). Sandbar-specific flows are required by the model since: 1) peak daily flows tend to decrease downstream from dams in periods with no runoff; 2) flows tend to increase downstream from dams in periods of runoff; and 3) flows increase downstream of major tributaries or other flow inputs (municipal or agricultural water returns). Whenever possible, regional flow models should be used to provide these inputs. However, if these models provide only inputs for mean daily flows, how much these models may underestimate tern nest or chick flooding risk should be evaluated (e.g., mean daily flows may not flood nests, but peak daily flows might). When regional flow models are not available to predict peak daily flows at all sandbar locations, Lott and Wiley (2012) provide simple methods for interpolation of flow data to sandbar locations using available flow data from gauges. Sandbar specific peak daily flows are stored in Table 5.

Table 5. Daily flow table. This table contains the actual sandbar-specific peak daily flow data.

Table Variable	Description
manCode	Management area code
plFlowInputSetName	Flow Input Set Name
wyrYear	Year of annual flow input
sitCode	Code for site that flow applies to
flwPeakFlow	Peak daily flow (in cfs)
flwJulianDay	Julian Day
flwDate	Date (e.g., 12/12/2002)

Providing peak daily flow inputs

The entry in the ***manCode*** field should match exactly a value from this field in the Management Area table (See Table 1). The entry in the ***plFlowInputSetName*** field should match exactly a value from this field in the **Flow Input Set** table (see Table 6). The entry in the ***wyrYear***

field should match exactly a value from this field in the Water Year table (see Table 7). The entry in the sitCode field should match exactly a value from this field in the Site table (see Table 11). The value in the flwPeakFlow flow field is a peak daily flow value that users must provide for each year/site/date combination (in cubic feet per second). The next two fields list the date by Julian Day (so that seasonal patterns in flows can be compared among years, regardless of what year they belong to) and by full date, so that flows are associated with an exact date in an exact year.

Table 6. Flow Input Set table. This table provides a list of the different sets of annual flow inputs that are available as peak daily flow time series inputs to the model.

Table Variable	Description
manCode	Management area code
plFlowInputSetCode	Short category name for graph axis labels
plFlowInputSetName	Longer name for selection menus in GUI
plFlowInputSetNotes	Longer text description with details

Table 7. The Water Year Category table lists all possible water year type categories for each management area/flow input set combination and their frequency within the historic record.

Table Variable	Description
manCode	Management area code
plFlowInputSetName	Name of flow input set
wycCode	Short category name for graph axis labels
wycName	Longer name for selection menus in GUI
wycNotes	Longer text description with details

4 Step 3: Classifying Annual Flow Inputs into Different Breeding Season Water Year Types

Background and methodological concerns

For each simulated **breeding season**, TernCOLONY requires daily time series inputs of site-specific peak daily flows. These time series may be generated from empirical data or flow models (see step 2). In examining long-term empirical datasets, it is clear that common patterns of seasonal flow variation during the breeding season recur below most dams. This is the function of interactions between frequently occurring patterns of regional precipitation and runoff and the rule curves that specify how dam releases occur in response to reservoir storage levels and inflow rates that are also a consequence of these regional patterns. A variety of methods might be used to identify commonly occurring patterns in seasonal flow variation and then to assign specific annual flow inputs to one of these categories. Whichever method is used, the authors strongly recommend creating categories of commonly occurring seasonal flow patterns that can then be used to test different hypotheses about interactions between flows and the bird population, since seasonal variation in flows can have such a strong effect on nesting sandbar availability and flooding mortality. For example, years with consistently low flows tend to play out very differently for Least Terns than years with mid-season flooding. For the Arkansas River, Lott and Wiley (2012) classified each annual flow input time series to one of three specific breeding season water year types (Figure 3). Since seasonal flow patterns will vary by management area, this analysis will need to be conducted separately for each management area. It should be noted, however, that the model does not REQUIRE that annual flow inputs be classified to different breeding season water year types. If users wish to skip this step, the database administrator will assign all annual flow inputs to the default water year type of "not specified." Information on water year type categories is stored in table 7 and the membership of any one annual flow input to a specific water year type category is stored in Table 8.

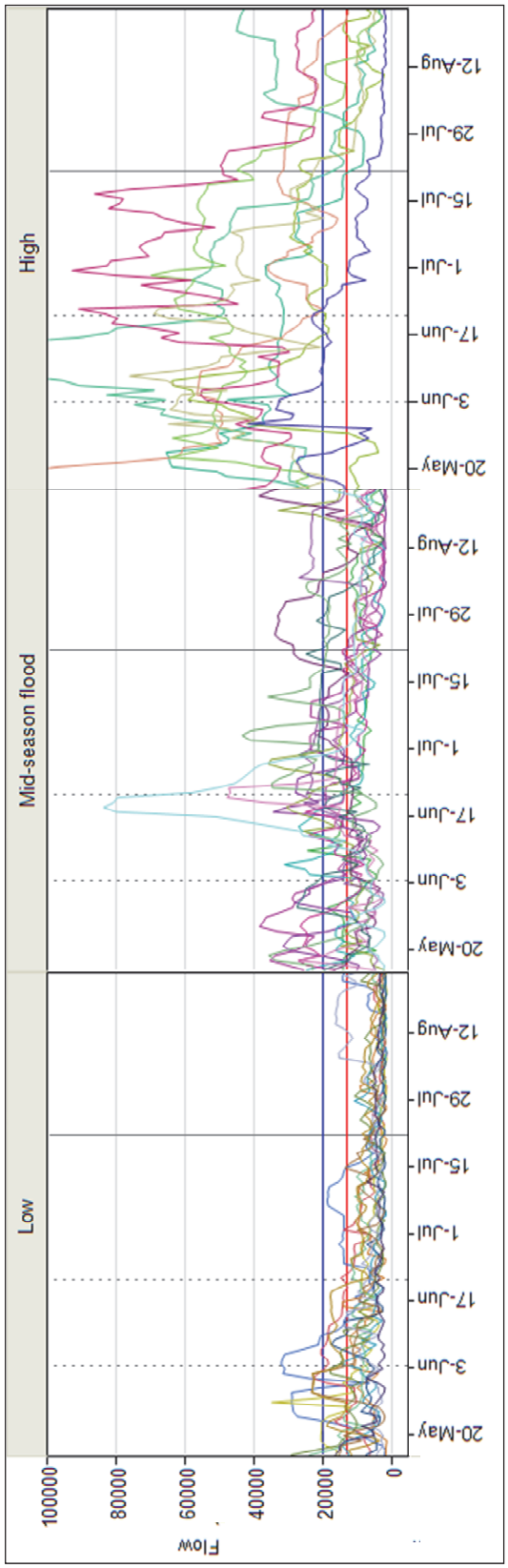

Figure 3. Example classification of annual flow inputs to water year type categories from Lott and Wiley (2012). Panels represent the three common breeding season water year types below Keystone Dam during the post-dam era (1977-2008): low water years (37.5% of years), mid-season flooding years (37.5% of years), and high water years (25% of years). Lines within each panel represent variation in time series of peak daily flows for individual years within each type. See Lott and Wiley (2012) for criteria used to classify each year to a specific water year type. The red horizontal line indicates peak hydropower flows of 13,000 cfs that regularly occur during the breeding season. The Blue horizontal line indicates moderate flood control releases of 20,000 cfs. The dashed vertical lines indicate the dates by which 25% (June 3rd) and 75% (June 21st) of all nests are typically initiated during years with no major floods. The solid vertical line at July 21 indicates the typical last date for nest initiation in a year without major floods.

Table 8. The Water Year table simply assigns each annual flow input to one of the categories listed in the Water Year Category Table (above).

Table Variable	Description
manCode	Code for management area
plFlowInputSetName	Name for Flow Input Set
wycCode	Code for water year type category
wyrYear	Identifier of the water year origin (must be integer)

General properties of flow inputs

The entry in the ***manCode*** field should match exactly a value from this field in the Management Area table (See Table 1). This table has the same three different types of naming fields for reaches that have been present in other tables (with 20, 50, and 1000 character limits).

Categorizing flow inputs by water year types

The entry in the ***manCode*** field should match exactly a value from this field in the Management Area table (See Table 1). Within one management area, users may have more than one set of flow inputs. For example, flow inputs may be available for a time series of years before a dam was in place or after a dam became operational (see Table 6). Users may wish to classify annual flow inputs to different water year types for these different sets of annual flow inputs. The entry in the ***plFlowInputSetName*** field should match exactly a value from this field in the Flow Input Set table (see Table 6). The Water Year Category table has the same three different types of naming fields for reaches that were present for management areas or reaches (with 20, 50, and 1000 character limits). The last field in this table requires users to report the proportion of annual flow inputs (among all annual flow inputs in a flow input set) that belong to each water year category.

The entry in the ***manCode*** field should match exactly a value from this field in the Management Area table (See Table 1). The entry in the ***plFlowInputSetName*** field should match exactly a value from this field in the Flow Input Set table (see Table 6). The entry in the ***wycCode*** field should match exactly a value from this field in the Water Year Category table (see Table 7). The wyrYear field must be an integer value indicating the actual year that the annual flow input time series relates to (e.g.., 1977, 2008).

5 Step 4: Providing Site-Based Inputs that Describe Sandbar Nesting Habitat

Background and methodological concerns

The model requires detailed information about the spatial distribution and quality of potential nesting sites for Least Terns. Within each management area, users may wish to have more than 1 set of habitat inputs, representing different conditions within the management area during different time periods (Table 9). Two main database tables store information about potential nesting sites. In TernCOLONY a "site" is defined as "a single location where terns may nest." When sandbars are the only site types in a model application (Table 10), a site could be described as "a discretely defined sandbar where a group of terns might form a colony" (Table 11). A "cell" is "a square patch of habitat within a site, used to represent spatial variation." Model terns evaluate cell quality and select a nesting location from the potential list of all cells at a site (Table 12). More information about how the model defines habitat quality is available in Lott et al. 2012. Sandbar habitat inputs are updated each day, as the model runs, to reflect the effects of daily flow variation on habitat availability (due to exposure or inundation) and quality (which is affected by the elevation difference between the waterline and habitat cells, otherwise known as "freeboard," which changes daily in response to increases or decreases in flows).

Habitat inputs may be prepared using a combination of field and/or GIS procedures, which may vary according to model application (see Overview). The bare minimum information that must be provided for each application includes: 1) information concerning variation in sandbar elevations among and within sites; 2) location of features that affect model tern colony and nest site selection (e.g., large trees, the river's bank, patches of sandbar vegetation) (Figure 2); and 3) models for how flows affect waterline elevations at each site. Lott and Wiley (2012) provide examples of how this information was provided for the Arkansas River below Keystone Dam management area. However, alternative approaches to the same problems of data collection may be better suited to data availability in other situations. A Python script was developed to take GIS inputs and translate them into text file outputs that are properly formatted to match TernCOLONY database tables (the script and its documentation are available at

http://www.leasttern.org). The authors encourage users to check out this time-saving resource before they begin the process of preparing habitat inputs for a new management area.

Table 9. Habitat Input Set table. This table provides a list of the different sets of habitat inputs that must be available as primary inputs to the model.

Table Variable	Description
manCode	Management area code
plHabitatInputSetCode	Short category name for graph axis labels
plHabitatInputSetName	Longer name for selection menus in GUI
plHabitatInputSetNotes	Longer text description with details

Table 10. Site Type Table. This table lists the types of nesting sites that can be included in the model.

Table Variable	Description
plSiteTypeName	Name for unique Site Type

Table 11. Site table.

Table Variable	Description
manCode	Management area code
rchCode	Reach code
plHabitatInputSetCode	Habitat input set code
plSiteTypeName	Site type name
sitPointX	X spatial coordinate
sitPointY	Y spatial coordinate
sitCellLength	Size of cells within site (in feet)
sitCode	One way of naming sites, must be unique for a given combination of management area and habitat input set
sitNearestCorpsRM	Another way of naming sites, can be the same for a site that is in the same location between two different years/habitat input sets
sitMilesFromDatum	Distance from datum (in miles, with 5 decimal precision)
sitFlowToElevA	Flow to elevation conversion coefficient A
sitFlowToElevB	Flow to elevation conversion coefficient B
sitFlowToElevC	Flow to elevation conversion coefficient C
sitFlowToElevD	Flow to elevation conversion coefficient D
sitFlowToElevE	Flow to elevation conversion coefficient E
sitFlowToElevF	Flow to elevation conversion coefficient F
sitMinSuitableFlow	Lowest daily maximum flow likely to occur during nest incubation

Table 12. Data necessary to complete the input table "tblCell" in the IBM database.

Table Variable	Description
sitCode	Site code
celPointX	X spatial coordinate
celPointY	Y spatial coordinate
celCFSAtPoint	Minimum flow at which site will be inundated
celDist2Bank	Distance to river bank
celDist2ForestEdge	Distance to nearest forest edge
celDist2Veg	Distance to nearest vegetation (use 9999 to indicate that no vegetation patches exist within the distance analyzed)
celMaxElevDiff	Elevation above sitPeakHydroFlow
celVegetation	Whether or not the cell contains vegetation (1 for true, 0 for false)
celElev	Maximum cell elevation

General properties of habitat inputs

For each management area, users may wish to include more than one set of habitat inputs, representing habitat conditions during different time periods (e.g., 2002 versus 2008, post-flood conditions versus degraded habitat conditions, potential restoration sandbars). This table lists each unique habitat input set by name. The entry in the ***manCode*** field should match exactly a value from this field in the Management Area table (See Table 1). This table has the same three different types of naming fields for reaches that have been present in other tables (with 20, 50, and 1000 character limits).

Currently, TernCOLONY includes only two site types, river sandbars or constructed sandbars. If users desire to include other site type categories (e.g., sand pits, rooftops), they should contact the model's developers.

Providing inputs for sites

The entry in the ***manCode*** field should match exactly a value from this field in the Management Area table (See Table 1). The entry in the ***rchCode*** field should match exactly a value from this field in the Reach table (See Table 2). The entry in the ***plHabitatInputSetCode*** field should match exactly a value from this field in the Habitat Input Set table (See Table 9). The entry in the ***plSiteTypeName*** field should match exactly a value from this field in the Site Type table (See Table 10). A single xy coordinate for

each site is necessary to display the site as a single point in the model's graphical user interface. TernCOLONY's default spatial resolution for habitat cells is 6 ft. by 6 ft. This resolution was chosen because it is small enough to accurately represent spatial variability in the distribution of vegetation and elevations that is typically found on river sandbars. It is also fairly typical of the minimum distance between Least Tern nests. Users should enter a value of "6" in the ***sitCellLength*** field, unless they'd like to use a cell size other than this default. If this is the case, they should contact the model's developers. The following three fields represent three different options for naming individual sites. In previous versions of TernCOLONY, the authors used the ***sitCode*** field to link GIS files for habitat data inputs to the database (e.g., if a site was originally named ARM_475_5 in the field, the authors used this as the ***sitCode*** name in the database). The authors have retained these names to preserve links between the database and the GIS data in case the GIS data ever needs to be revisited. However, once the GIS data were explored, two different ways of naming sites were found to be more useful. First, the authors named sites by the Corps of Engineers River Mile that they were closest to (with 2 decimals of precision to distinguish sites that were close to each other), using the field ***sitNearestCorpsRM***. This approach was desirable because River Miles are commonly used for local reference. However, since many modern channels have changed considerably since river miles were specified, Corps of Engineers River Miles are frequently inaccurate for measuring distances. Therefore, the naming field ***sitMilesFromDatum*** was included and the thalweg point file was used to name sites by their distance downstream from the dam at the head of each management area (with 2 decimal places). This allows sites to be listed by their distance downstream from the dam, which helps in analyses designed to understand the effects of dam releases on tern populations.

The six fields called sitFlowToElevA-sitFlowToElevF contain coefficients from log-quadratic equations that specify the relationship between flows (x) and water surface elevations (y) at each sandbar. Ideally, these equations would be constructed from empirical data collected at each site, which would factor in the local effects on this relationship and control for the reality that stage-discharge relationships are always affected by local channel planform and geometry. In the absence of such information, Lott and Wiley (2012) provide details on how to construct these relationships via interpolation for stage-discharge relationships recorded at gauges. It is recommended that developers of new applications of TernCOLONY

contact the authors directly for help with this step, since the authors have developed several scripts that speed the generation of these data.

The ***sitMinSuitableFlow*** requires site-specific estimate for the peak daily flow that is likely to occur during full hydropower production. On many regulated rivers this flow recurs with daily or weekly frequencies during the Least Tern breeding season. Nests that are placed at elevations that are inundated at this flow have virtually zero chance of surviving. This flow can be estimated via interpolation by examining hydrographs at distant river gauges to identify regular peaks associated with hydropower production (See Lott and Wiley 2012).

Providing inputs for cells

The entry in the ***sitCode*** field should match exactly a value from this field in the Site table (See Table 11). The xy coordinates must be reported in the coordinate system defined in the ***manCoordinateSystem*** field in the Management Area table. All other fields in this table represent habitat covariates that must be calculated for each cell. Each of these covariates is defined in Lott et al. 2012. Python scripts that were developed to calculate covariate information for each of the habitat cells at a site, based on a standard set of GIS inputs, are available upon request from the first author of this report. These scripts will likely need small amounts of revision for use in any new management area. Required GIS inputs to calculate these covariates are: 1) digital elevation models for sandbars; 2) polygons that delineate the extent of vegetation on sandbars; 3) line files adjacent to sandbars that indicate the location of the river's bank (e.g., active channel margin); and 4) line files that indicate the location of large trees both inside and outside of the channel. See Lott and Wiley (2012) for methods on collecting these data.

6 Step 5: Specifying Inputs that Describe Properties of the Bird Population

Background and methodological concerns

Least Terns are known to show variable amounts of fidelity to previous breeding sites (Thompson et al. 1997). Therefore, to incorporate the behavior of site fidelity into colony site selection, the model requires tabular inputs that represent the location (and relative abundance) of breeding terns from previous years. This information is stored in Table 13. This allows the model to bring terns back during spring migration to locations in each management area near where they bred the previous year. However, not all terns are site-faithful and the model also allows a fraction of the tern population to be "dispersers." These individuals are assigned to random arrival locations within the management area prior to site selection (see Lott et al. in preparation). TernCOLONY uses a default value of 20% of the population as dispersers, based on extensive model testing versus patterns (Lott et al. 2012). Users who wish to use another value for this parameter should contact the model's developers. Similarly, TernCOLONY includes a set of default parameters that determine the phenology of the Least Tern breeding season (e.g., arrival dates, peak nesting dates, last possible nesting dates, etc.). These dates were chosen, after thorough literature search, to represent the range of dates that tend to be similar across riverine breeding areas from Oklahoma to North Dakota (Lott et al. 2012a). Users who wish to use different values for date parameters should contact the model's developers.

Table 13. Tern Previous Site table.

Table Variable	Description
manCode	Management area code
plHabitatInputSetCode	Habitat input set code
tpsPointX	X spatial coordinate
tpsPointY	Y spatial coordinate
tpsNearestThalwegPointID	Identifier of nearest thalweg point (see Table 3)
tpsFraction	Fraction of adult terns nesting at the site
tpsDefault	Is this set of previous site data the default?

Specifying prior distributions of nesting terns

The entry in the ***manCode*** field should match exactly a value from this field in the Management Area table (See Table 1). Within one management area, users may have more than one set of habitat inputs. For example, habitat inputs may be available for 1994 and 2010. If there has been considerable channel change in the distribution or relative abundance of breeding terns within the management area between these two dates, users may wish to create a new "Tern Previous Site" table for each time period. The entry in the ***plHabitatInputSetCode*** field should match exactly a value from this field in the Habitat Input Set table (See Table 9). The xy coordinate fields must match the coordinate system specified in the Management Area table (Table 1). These represent a single point for a colony location. Since nests are spread out within a colony, this is obviously an imprecise measure. In case xy coordinates from field data do not overlap thalweg points (as is regularly the case), users must specify the thalweg point closes to each site. This should match exactly the identifier for a thalweg point that was provided in the field ***tempThalwegPointID*** in table 3. The tpsCount and tpsFraction fields are where users report the number of adult terns (or nests, if these are the field data that are available) for each site (and their proportion of the total for a single survey). The values in the field tpsFraction must sum to 1 for all records with the same combination of manCode and plHabitatInputSetCode. It should be noted, tpsFraction is used by the model to distribute adult arrivals throughout the management area, given a user-specified population size. If more than one set of "Tern Previous Site" datasets are available for a single habitat input set, placing a value of "1" in the tpsDefault field makes this set the default. A value of "0" indicates that a set of locations/abundances is not the default for a specific habitat input set; however, users will have the option of selecting this set of locations/abundances through the GUI.

References

Grimm, V. and S. F. Railsback. 2005. *Individual-based modeling and ecology*. Princeton, NJ: Princeton University Press.

Grimm, V., U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. Heinz, G. Huse, A. Huth, J. U. Jepsen, C. Jørgensen, W. M. Mooij, B. Müller, G. Pe'er, C. Piou, S. F. Railsback, A. M. Robbins, M. M. Robbins, E. Rossmanith, N. Rüger, E. Strand, S. Souissi, R. A. Stillman, R. Vabø, U. Visser, and D. L. DeAngelis. 2006. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling* 198:115-296.

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, G., Giske, J. and S.F. Railsback. 2010 The ODD protocol: a review and first update. *Ecological Modelling* 221: 2760-2768.

Lott, C.A. and R.L. Wiley. 2012. *Effects of dam operations on Least Tern nesting habitat and reproductive success below Keystone Dam on the Arkansas River*. ERDC/EL CR-12-4. Vicksburg, MS: U.S. Army Engineer Research and Development Center.

Lott, C.A., S.F. Railsback, and C.J.R. Sheppard. 2012. *TernCOLONY 1.0 model description*. ERDC/EL CR-12-3. Vicksburg, MS: U.S. Army Engineer Research and Development Center.

Lott, C.A., S.F. Railsback, C.J.R. Sheppard, and M.C. Koohafkan. *The development and testing of TernCOLONY 1.0: A web-accessible individual-based model of Least Tern reproduction*. (In preparation). Vicksburg, MS: U.S. Army Engineer Research and Development Center.

Thompson, B. C., J. A. Jackson, J. Burger, L. A. Hill, E. M. Kirsch, and J. L. Atwood. 1997. Least Tern: Sterna antillarum. *Birds of North America*:1-32.

Glossary

Annual flow input- A single-time series of peak-daily flows during the breeding season.

Breeding season water year type- One of three categories that group water years by type: 1) low water for whole season; 2) moderate floods early, with low flows for rest of season; or 3) High water for whole season.

Breeding season- This is the temporal extent of a single **model run**. It is inclusive of the entire time period between the arrival of the first adult Least Tern in the management area and the date by which the last tern chick fledges (see **nesting dates** and **last fledging date**).

Calibration- The process by which model parameters are adjusted so that model outputs most closely resemble observable phenomenon or vital rates from real tern populations.

Cell- A cell is a square patch of habitat within a site, used to represent spatial variation in habitat characteristics.

Flow input set- Groups of water years associated with one set of flow inputs (e.g., observed post-alteration time-series [1977-2008], simulated quick-flood routing scenario [1977-2008]).

Freeboard- The elevation difference between the waterline and a location on a sandbar, usually expressed in 100ths of U.S. survey feet.

GUI- Graphical user interface for the TernCOLONY model. This is available at http://www.leasttern.org.

Habitat input set- A groups of sites associated with a set of habitat conditions at one point in time (e.g., excellent habitat conditions [2009], simulated degraded habitat conditions [2005], potential created sandbars).

Habitat quality index- This is a value from 0-1 that is updated for each habitat cell on each day of the breeding season. It is the product of logistic functions for five habitat variables that affect both Least Tern colony site

selection and nest site placement within a colony in the model. See Model Description for more details. The use of different versions of this habitat quality index should be discussed with the model's developers and would require considerable testing to implement.

Last fledging date- A model parameter that determines the last date on which chicks can successfully fledge. It is used by the model to determine the last dates that adults can re-nest. The standard value for this parameter is September 1, based on the assumption that ILT begin fall migration around this date. However, final hatch dates for some datasets suggest that this date could be 1-2 weeks later if late-hatching chicks receive parental care through fledging. Users may wish to increase this value.

Management area- A collection of sites that represents all of the potential nesting areas for one subpopulation of terns. Many different definitions of subpopulations are possible. For river applications of TernCOLONY, the authors specifically define a management area as a collection of sites that are all affected to some degree by the operations of a single large dam.

Minimum suitable flow- Below this threshold flow (in cfs), all tern nests are expected to fail due to inundation (e.g., there has not been a single 3-week period during Least Tern incubation with flows less than 13,000 cfs on the Arkansas River below Keystone Dam). Therefore, all portions of sand-bars that are inundated at flows of 13,000 cfs are considered unsuitable for successful reproduction. This flow will differ by management area and can only be determined from exploration of historic hydrographs. This flow can be defined by site or for the whole management area. It is used in the model as a benchmark to express site-specific elevations as a "freeboard" value that can be directly compared across sites.

Model run- A single simulated breeding season (see **scenario** and **replicate** for two main types of model runs).

Nesting dates- Four different dates define the Least Tern breeding season). These include: 1) the date on which the first nest in a population is initiated; 2) the peak nesting date; 3) the date on which the last nest is initiated; and 4) the last possible fledging date. It should be noted that the first three dates have been specified based on empirical data on nest initiation dates from several locations within the Interior Population (with

default values of May 17, June 3, and July 18). See **last fledging date** for additional detail. Users may wish to modify these defaults. Since changing these parameters will also affect when terns arrive in the model (see Model Description), **annual flow inputs** should always be provided for at least 30 days prior to the first nesting date and 30 days after the last fledging date.

Reach- A pre-defined subset of locations within the management area where factors affecting tern habitat or population processes may be different from other subsets of locations (e.g., a major tributary may form the boundary between two reaches). Reaches are used only for summary of model outputs and have no direct effect on the operations of the model itself.

Replicate- A model run with an identical set of parameter inputs as another model run, where only the random number seed varies among runs. Due to the amount of **stochasticity** in TernCOLONY, the same set of parameter values may generate different outcomes.

Scenario- A single simulated breeding season with a unique list of parameter values (not including the random number seed). More than one replicate can be performed for a single scenario. See also Replicate.

Site- A "site" represents a single location where terns may nest. Version 3.3 of TernCOLONY has only two site types: river sandbars and constructed sandbars. However, future versions could include other types such as sand and gravel pits, salt flats, rooftops, reservoir islands, or coastal beaches. Sites comprise a collection of **cells** that represent how habitat variables vary over space within the site.

Stochasticity- Randomness in the model. Often, variation among individuals (e.g., duration of the incubation period) or in binary processes (e.g., whether or not a predator discovers a nesting site on any one day). TernCOLONY uses a random number generator to select values from distributions that describe the magnitude (and shape) of variability in several Least Tern state variables or event-driven processes. By adding this variability to the model, virtual tern populations can vary similarly in their response to different conditions as real tern populations.

Thalweg- The deepest part of the channel with the strongest current. In TernColony, the thalweg is represented as a set of points, evenly spaced at 100-ft increments, from the dam to the downstream boundary of the management area. These points are used to locate model terns in the upstream-downstream direction.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*
December 2012

2. REPORT TYPE
Final report

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE
Preparing Inputs for New Applications of the TernCOLONY Simulation Model

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Casey A. Lott, Jonathon J. Valente, Richard A. Fischer, and Colin J.R. Sheppard

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Coastal and Waterways Program Coordinator
American Bird Conservancy
1209 Shenandoah Drive, Boise, ID 83712;
Environmental Laboratory
U.S. Army Engineer Research and Development Center
3909 Halls Ferry Road, Vicksburg, MS 39180;
Lang, Railsback, and Associates
250 California Ave., Arcata, CA 95521

8. PERFORMING ORGANIZATION REPORT NUMBER
ERDC/EL TR-12-32

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The TernCOLONY simulation model is an individual-based model of Least Tern reproduction that was developed to better understand how reservoir operations (and other management activities) affect Least Tern breeding populations on large rivers. This document provides practical guidance for those wishing to implement this model in a new region. Specifically, this document provides guidance concerning how to prepare the tabular inputs that are necessary to import required site-specific data on habitat, flows, and bird populations into the relational database that manages TernCOLONY's inputs and outputs. This document is a technical, step-by-step guide to fulfilling this basic objective. For more information on the model itself, please visit the model's website at http://www.leasttern.org.

15. SUBJECT TERMS
Least Tern
Least Tern reproduction
Habitat and flow datasets
TernCOLONY
ODD protocol for models
Reproduction model

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON: Tanya Beck**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER *(include area code)***
Unclassified	Unclassified	Unclassified	Unclassified	32	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. 239.18