

Metadata Creation Tool Content Template For Data Stewards

Instructions for use: Please complete all sections in each table below under the ‘FIELD CONTENT’ column. The cells associated with each field in the table will automatically expand to accommodate all of the text that is either typed into the ‘FIELD CONTENT’ column or pasted in from other documents. An * next to a field name indicates that it is mandatory for information to be entered into the adjacent ‘FIELD CONTENT’ cell. *Italic text* in the ‘FIELD’ column denotes that additional information can be added if available. Do not alter any of the field labels in any of the tables. Please use either the Metadata Creation Tool User Guide or the Metadata Content Guidance Document to assist you with completing each cell. If you need further assistance, please contact the Technical Assistance Team at ephtmetadata@cdc.gov

I. IDENTIFICATION TAB

A. CITATION PAGE

FIELD	FIELD CONTENT
* CATEGORY	Air quality
* PUBLICATION DATE	January 23, 2009
* TITLE	Space-time Predictions of Air and Deposition Pollutants
URL	
* NATIVE DATASET ENVIRONMENT	Predictive output files (.csv) include the following variables: Date, Longitude, Latitude, CMAQ column, CMAQ row, Predictions, Predictive standard error

B. DESCRIPTION PAGE

FIELD	FIELD CONTENT
* ABSTRACT	A space-time Bayesian fusion model (McMillan, Holland, Morara, and Feng, 2009) is used to provide daily, gridded predictive PM _{2.5} (daily average) and O ₃ (daily 8-hr maximum) surfaces for 2001-2005. The fusion model uses both air quality monitoring data from the National Air Monitoring Stations/State and Local Air Monitoring Stations (NAMS/SLAMS) and numerical output from the Models-3/Community Multiscale Air Quality (CMAQ). Predictive surfaces are provided over the entire CMAQ 12 km and 36 km spatial domains for each year.
* PURPOSE	The predictive surfaces are intended for use by statisticians and environmental scientists interested in the spatial distribution of pollution over daily time periods over 2001-2005. These surfaces could be used in modeling public health – air quality relationships, modeling ecosystem reactions to atmospheric inputs, and determining geographic areas with high pollution levels exceeding threshold levels. As new space-time models are developed, we will evaluate their use for providing improved predictions of air and deposition variables.

SUPPLEMENTAL INFORMATION	
* PROGRESS	Complete
* UPDATE FREQ.	As needed

C. TIME & DATE PAGE

FIELD	FIELD CONTENT
* CURRENTNESS	20090123
* DATE TYPE	Range
* SINGLE DATE	
* MULTIPLE DATES	
Date 1	
Date 2	
Date 3	
* RANGE OF DATES	FROM: January 1, 2001 TO: December 31, 2005

D. GEOGRAPHIC AREA PAGE

FIELD	FIELD CONTENT
* WEST COORDINATE	-128.09
* EAST COORDINATE	-65.47
* NORTH COORDINATE	51.46
* SOUTH COORDINATE	23.10

E. KEYWORDS PAGE

FIELD	FIELD CONTENT
* THEME	ISO
* THEME KEYWORDS	Environment
<i>THEME 2</i>	EPA
<i>THEME 2 KEYWORDS</i>	Air
<i>THEME 3</i>	
<i>THEME 3 KEYWORDS</i>	
* PLACES	United States
* PLACES KEYWORDS	Contiguous 48 states
<i>PLACES 2</i>	
<i>PLACES 2 KEYWORDS</i>	
<i>PLACES 3</i>	

PLACES 3 KEYWORDS	
-------------------	--

F. SECURITY PAGE

FIELD	FIELD CONTENT
* SECURITY CLASSIFICATION SYSTEM	EPA classification system
* CLASSIFICATION	Medium Confidentiality
* SECURITY HANDLING DESCRIPTION	May be shared with EPHT partners
* ACCESS CONSTRAINTS	Access for specific applications within use constraints
* USE CONSTRAINTS	The data are intended for use by statisticians interested in the spatial distribution of daily air pollution over multiple years. Collaboration with EPA in these studies is expected.

II. DATA QUALITY TAB

FIELD	FIELD CONTENT
* PROCESS DATE	20090123
* PROCESS DESCRIPTION	Through a collaborative process, EPA has developed software to fit the fused surfaces. At this time, these surfaces are generated at EPA.
<i>PROCESS DATE</i>	
<i>PROCESS DESCRIPTION</i>	
<i>PROCESS DATE</i>	
<i>PROCESS DESCRIPTION</i>	
<i>PROCESS DATE</i>	
<i>PROCESS DESCRIPTION</i>	
<i>PROCESS DATE</i>	
<i>PROCESS DESCRIPTION</i>	
<i>PROCESS DATE</i>	
<i>PROCESS DESCRIPTION</i>	
<i>PROCESS DATE</i>	
<i>PROCESS DESCRIPTION</i>	
<i>PROCESS DATE</i>	
<i>PROCESS DESCRIPTION</i>	
<i>PROCESS DATE</i>	
<i>PROCESS DESCRIPTION</i>	
<i>PROCESS DATE</i>	
<i>PROCESS DESCRIPTION</i>	
* LOGICAL	The predictive surfaces are based on using two sources of spatial information:

CONSISTENCY REPORT	air monitoring data and CMAQ numerical output for 2001-05. The CMAQ output is produced at EPA (http://www.epa.gov/asmdnerl/CMAQ).
* COMPLETENESS REPORT	Except for PM _{2.5} , predictive surfaces are provided for all days of the year. For PM _{2.5} , predictions cover all days except the first and last day of the year.

III. ENTITY AND ATTRIBUTES TAB

FIELD	FIELD CONTENT
* OVERVIEW	The predictive surfaces are intended for use by statisticians in modeling efforts that require high resolution air quality inputs for the period, 2001-05.
* DETAILED CITATION	<p>Input data The NAMS/SLAMS air quality monitoring data were downloaded from the Air Quality System (AQS) database. The CMAQ numerical output were created from version 4.6 of the model using CBIV mechanism. The output is a 24-hour integrated PM_{2.5} areal average concentration over 12 km and 36 km CMAQ cells. These CMAQ results are based on: (1) the emissions data from the EPA's National Emissions Inventory (NEI) 2002 (developed using mobile emissions model Mobile 6; and (2) daily continuous emissions monitoring (CEM) data for the major NO_x point sources). Further, the meteorological data used for these model results is from Mesoscale Model 5 (MM5) version 3.6.3 simulations (FDDA, Pleim-Xu lsm).</p> <p>The space-time Bayesian fusion model combines the monitoring data and CMAQ output to predict PM_{2.5} through space and time. The model assumes that both the actual monitoring data and the CMAQ data provide good information about the same underlying pollutant surface, but with different measurement error structures. It gives more weight to the accurate monitoring data in areas where monitoring data exists and relies on the CMAQ data and satellite data in areas where no monitoring data is available. The modeling is divided into hierarchical components where each level of the hierarchy is modeled conditional on the preceding levels. To fit the model, a custom-designed Monte Carlo Markov Chain (MCMC) software was used.</p>

IV. DISTRIBUTION TAB

FIELD	FIELD CONTENT
RESOURCE DESCRIPTION	Downloadable Data Files (.csv) containing Predictive Surfaces

DISCLAIMER/LIABILITY	Although these data have been processed successfully on a computer system at the Environmental Protection Agency, no warranty expressed or implied is made regarding the accuracy or utility of the data on any other system or for general or scientific purposes, nor shall the act of distribution constitute any such warranty. It is also strongly recommended that careful attention be paid to the contents of the metadata file associated with these data to evaluate data set limitations, restrictions or intended use. The U.S. Environmental Protection Agency shall not be held liable for improper or incorrect use of the data described and/or contained herein.
CUSTOM ORDER PROCESS	

V. METADATA TAB

FIELD	FIELD CONTENT
* DATE CREATED	200908123
* STANDARD NAME	EPHTN TEMPLATE VERSION 1.1
* ACCESS CONSTRAINTS	Access for specific applications within use constraints
* USE CONSTRAINTS	The data are intended for use by statisticians in modeling efforts that require high resolution predictive spatial fields of air pollution.

VI. CONTACTS TAB

A. MATRIX PAGE

FIELD	FIELD CONTENT
* CONTACT 1 NAME	David M. Holland
* CONTACT 1 TYPE	Primary Statistician
CONTACT 2 NAME	Vasu Kilaru
CONTACT 2 TYPE	Primary Web Site Manager
CONTACT 3 NAME	
CONTACT 3 TYPE	
CONTACT 4 NAME	
CONTACT 4 TYPE	

B. ORIGINATORS PAGE

FIELD	FIELD CONTENT
* PERSON	David Holland
* ORGANIZATION	US Environmental Protection Agency
* TITLE	Mathematical Statistician
USERID	
HOURS	
INSTRUCTIONS	
* PHONE NO. 1	919-541-3126
PHONE NO. 2	
* FAX	919-541-1138
* E-MAIL	Holland.david@epa.gov
TDD/TTY	
* STREET ADDRESS	109 T. W. Alexander Drive, E243-05 (NERL)
* CITY	Research Triangle Park
STATE	NC
COUNTRY	
* ZIP	27711

C. DISTRIBUTORS PAGE

FIELD	FIELD CONTENT
* PERSON	David M. Holland
* ORGANIZATION	US Environmental Protection Agency
* TITLE	Mathematical Statistician
USERID	
HOURS	
INSTRUCTIONS	
* PHONE NO. 1	919-541-3126
PHONE NO. 2	
* FAX	919-541-1138
* E-MAIL	Holland.david@epa.gov
TDD/TTY	
* STREET ADDRESS	109 T. W. Alexander Drive, E243-05 (NERL)
* CITY	Research Triangle Park
STATE	NC
COUNTRY	
* ZIP	27711

D. METADATA CONTACTS PAGE

FIELD	FIELD CONTENT
* PERSON	Vasu Kilaru
* ORGANIZATION	US Environmental Protection Agency
* TITLE	Atmospheric Scientist
USERID	
HOURS	
INSTRUCTIONS	
* PHONE NO. 1	919-541-5332
PHONE NO. 2	
* FAX	919-541-1138
* E-MAIL	Kilaru.vasu@epa.gov
TDD/TTY	
* STREET ADDRESS	109 T. W. Alexander Drive, E243-05 (NERL)
* CITY	Research Triangle Park
STATE	NC
COUNTRY	
* ZIP	27711