Science to Support the USGS Changing Arctic Ecosystems Initiative
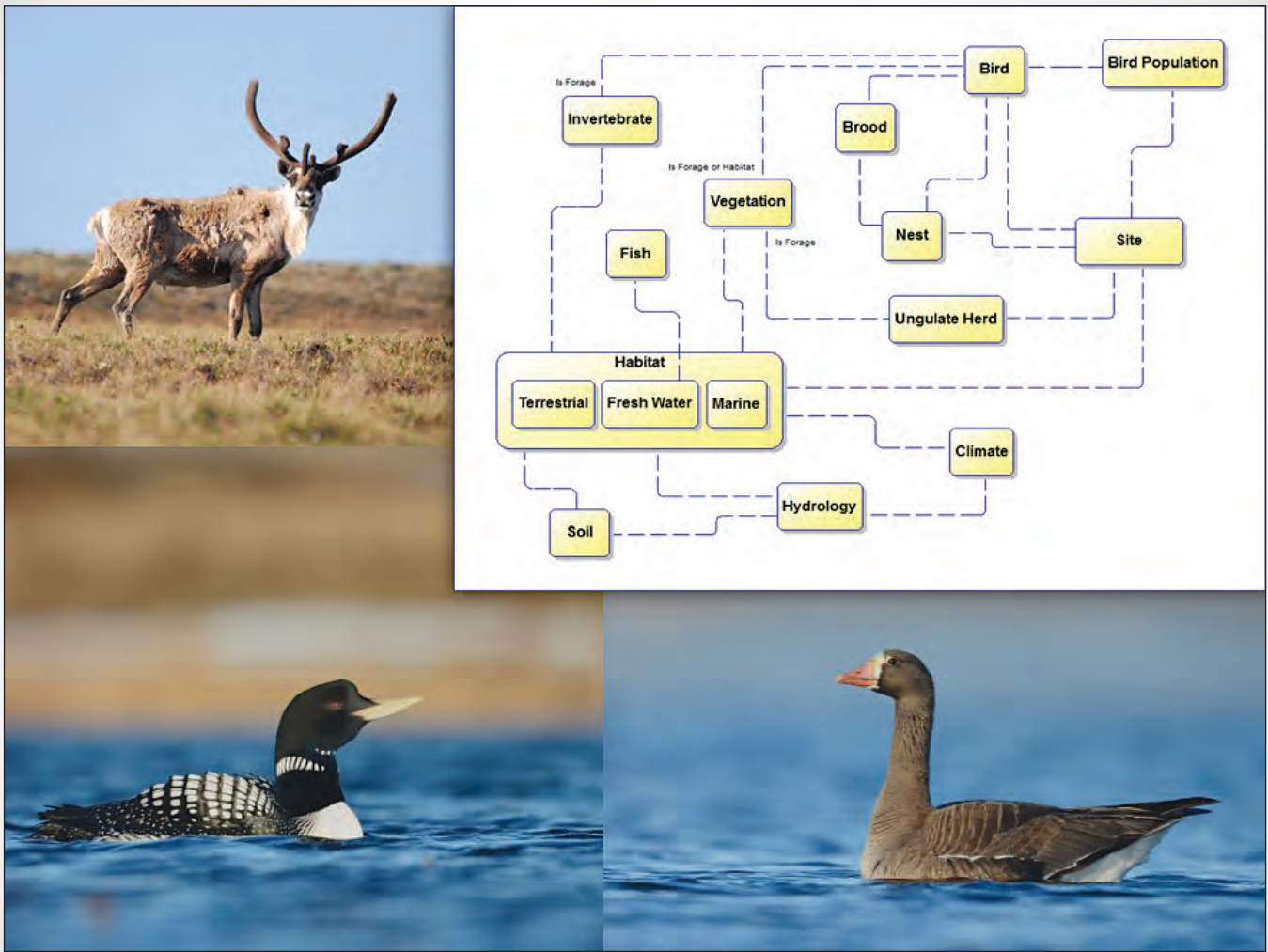
Conceptual Data Modeling of Wildlife Response Indicators to Ecosystem Change in the Arctic

Open-File Report 2015–1148

Conceptual Data Modeling of Wildlife Response Indicators to Ecosystem Change in the Arctic

By Dennis Walworth and John M. Pearce

Science to Support the USGS Changing Arctic Ecosystems Initiative

Open-File Report 2015-1148

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment—visit http://www.usgs.gov/ or call 1–888–ASK–USGS (1–888–275–8747).

For an overview of USGS information products, including maps, imagery, and publications, visit http://www.usgs.gov/pubprod/.

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Contents

Figures

Conversion Factors

Inch/Pound to International System of Units

Multiply	By	To obtain
Length		
inch (in.)	2.54	centimeter (cm)
inch (in.)	25.4	millimeter (mm)
Pressure		
pound per square inch (lb/in^2)	6.895	kilopascal (kPa)

International System of Units to Inch/Pound

Multiply	By	To obtain
Length		
centimeter (cm)	0.3937	inch (in.)
millimeter (mm)	0.03937	inch (in.)
meter (m)	3.281	foot (ft)
meter (m)	1.094	yard (yd)
kilometer (km)	0.6214	mile (mi)
kilometer (km)	0.5400	mile, nautical (nmi)
Area		
square meter (m^2)	0.0002471	acre
Velocity		
kilometer per hour (km/h)	0.6214	mile per hour (mi/h)
Mass		
gram (g)	0.03527	ounce, avoirdupois (oz)

Temperature in degrees Celsius (°C) may be converted to degrees Fahrenheit (°F) as °F = (1.8 × °C) + 32.

Temperature in degrees Fahrenheit (°F) may be converted to degrees Celsius (°C) as °C = (°F – 32) / 1.8.

Conceptual Data Modeling of Wildlife Response Indicators to Ecosystem Change in the Arctic

By Dennis Walworth and John M. Pearce

Abstract

Large research studies are often challenged to effectively expose and document the types of information being collected and the reasons for data collection across what are often a diverse cadre of investigators of differing disciplines. We applied concepts from the field of information or data modeling to the U.S. Geological Survey (USGS) Changing Arctic Ecosystems (CAE) initiative to prototype an application of information modeling. The USGS CAE initiative is collecting information from marine and terrestrial environments in Alaska to identify and understand the links between rapid physical changes in the Arctic and response of wildlife populations to these ecosystem changes. An associated need is to understand how data collection strategies are informing the overall science initiative and facilitating communication of those strategies to a wide audience. We explored the use of conceptual data modeling to provide a method by which to document, describe, and visually communicate both enterprise and study level data; provide a simple means to analyze commonalities and differences in data acquisition strategies between studies; and provide a tool for discussing those strategies among researchers and managers.

Introduction

Large science initiatives are often challenged to effectively expose and document the types of information being collected and the reasons for data collection across what are often a diverse range of research studies engaging different disciplines. Metadata provides data documentation, but often after the study is over, rather than during data acquisition planning. Metadata also is not in a form suitable for diverse investigators to actively participate as a team in data collection decisions to improve the ultimate scale and scope of findings. To document data collection decisions before and during a large research study, a different kind of methodology is needed; one that informs how a data acquisition strategy is informing scientific objectives, supporting data integration, and offering an effective means of communicating that strategy to external partners. Ideally, such a methodology would provide for a visual examination for comparative analysis of what kinds of data will be collected between different studies, and therefore assist in identifying opportunities for enhancing and leveraging data acquisition strategies.

There currently are multiple science initiatives focused on understanding the rapid physical and biological processes in the Arctic to inform critical natural resource and policy decisions. Wildlife habitats in the Arctic are changing rapidly and some of these changes are occurring faster than previously forecasted (Stroeve and others, 2007). For example, sea ice extent and thickness have been progressively decreasing (Overland and Wang, 2013), air temperatures have increased in Alaska over the last 50 years (Stafford and others, 2000), and the melting of the seasonal snowpack has been occurring earlier in northern Alaska, leading to an increase in the growing season by approximately 8 days since the mid-1960s (Stone and others, 2002). These changes in the physical processes of the Arctic are influencing biological drivers of habitat suitability and productivity. In Alaska, several studies have documented changes in wildlife distribution, behavior, and habitats resulting from physical and biological drivers (Flint and others, 2008; Durner and others, 2011; Jay and others, 2011; Ravens and others, 2012; Tape and others, 2013). However, for many wildlife species and habitats in Alaska, the full extent of responses to climate change remains poorly defined.

The U.S. Geological Survey (USGS) Changing Arctic Ecosystems (CAE) initiative focuses research efforts in northern Alaska to identify and understand the links between physical processes, ecosystems, and wildlife population response (Van Hemert and others, 2015). With the fundamental physical change in the Arctic being temperature and its influence on sea ice and permafrost, USGS CAE studies are focusing on ice-dominated ecosystems of Alaska—the marine sea-ice environment, the continuous permafrost tundra areas of northern Alaska, and the discontinuous permafrost landscape of the Arctic tundra-boreal forest ecotone (Oakley and others, 2012; Pearce and others, 2012). As part of this research initiative, the USGS is collecting a wide variety of data to identify and understand the link between physical processes, ecosystems, and wildlife populations (for examples, see Van Hemert and others, 2015). An associated need is to understand how data collection strategies are informing the overall science initiative and enabling communication of those strategies to a wide audience. We sought a data visualization methodology for the USGS CAE initiative that would fill these needs.

We investigated the field of information modeling, or data modeling for methods to accomplish our goals. Information modeling was developed as a software engineering specialization in the 1970s as a means to understand and document data used by organizations, and to design effective business and data management practices that enhanced understanding (Codd, 1970; American National Standards Institute, 1975; Chen, 1976). Information modeling offers well-vetted approaches to documenting data, and visually communicates an understanding of data and systems, relative to an organization's objectives (Zachman, 1987).

For this effort we selected a specific information modeling method that results in the generation of conceptual data models (CDM). Conceptual modeling addresses the overall perspective of data from an organizational point of view. CDMs are an abstraction of data unconcerned with the specifics or physical forms of the data (American National Standards Institute, 1975). Because of their simplicity, CDMs are ideal for communicating general ideas about data and science objectives to a broad audience.

In this report, we describe our approach for applying conceptual data modeling to document, analyze, and describe data produced by scientific research in the Arctic by the USGS CAE initiative. Our primary objectives were (1) to produce a graphical depiction of our early terrestrial ecosystem data collection strategy to facilitate conversations among scientists and research managers about data acquisition, project design, and science products for both current and future research efforts; (2) to develop a unifying methodology to describe the data being acquired at different study locations across the Arctic and make this approach extendable to new studies or sources of information that may be collected in the future; and (3) to facilitate communication with external partners regarding our data strategy.

Methodology

The USGS CAE initiative is collecting numerous types of terrestrial ecosystem data across a vast region of northern Alaska (fig. 1). Examples of data being collected include temperature and precipitation, hydrological patterns, and ice thaw dynamics; wildlife forage characteristics, such as vegetation growth, nutrient availability, and fish and invertebrate abundance; and wildlife population characteristics, such as individual movements, foraging ecology, and reproduction (Flint and others, 2014; Gustine and others, 2014; Uher-Koch and others, 2014). Our methodology was to analyze the types of data being collected across the terrestrial ecosystem portion of the USGS CAE initiative and develop CDMs to document the data for current studies.
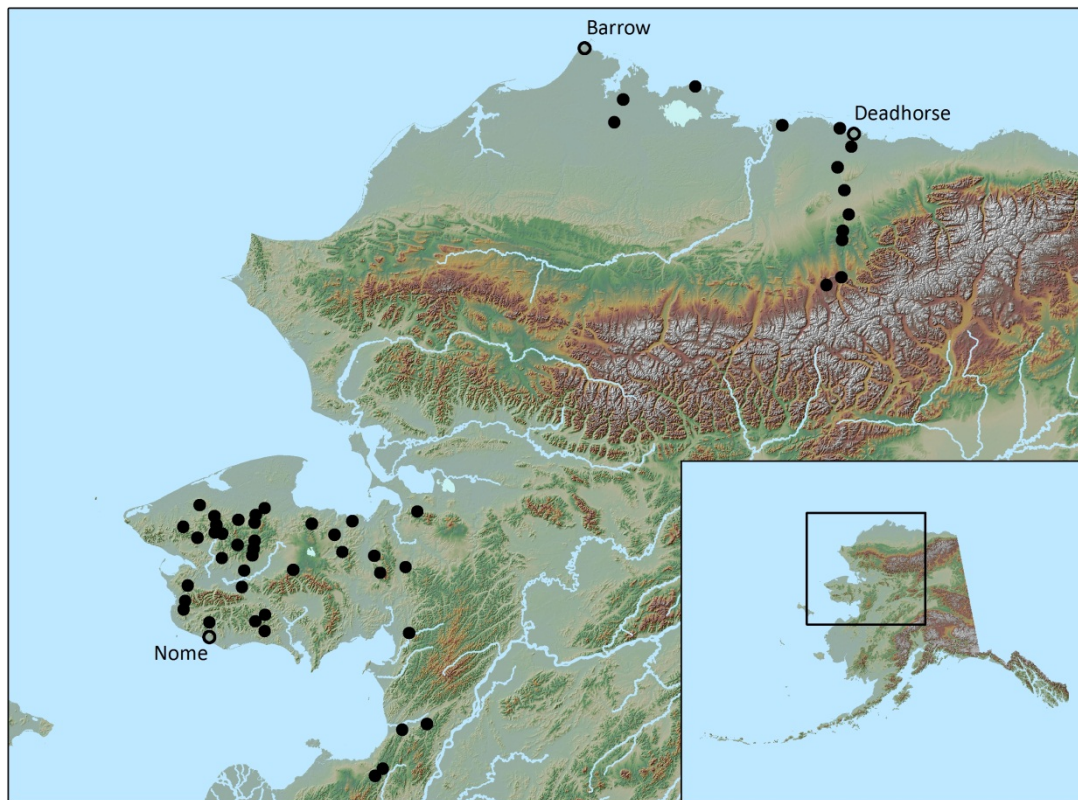
Figure 1. Depiction of U.S. Geological Survey Changing Arctic Ecosystem terrestrial study site locations from selected phase 1 research projects, northern Alaska, 2010–14. Inset shows location of study region within the State of Alaska.

As mentioned, CDMs address data in a generalized way from an organizational or enterprise point of view. CDMs are generally used to identify scope. They identify the main subjects of interest and the relationships between the subjects in a general way, but do not delve into details about the data. Physical forms of the data, collection methodologies, and physical storage are all conceptualized. In a sense, it is a pure "science" perspective of the data.

CDMs contain "entities" which are the main subject areas under study, such as "Bird," "Vegetation," and "Habitat." Relationships between entities are symbolized as lines connecting the entities to portray influence and some degree of dependence. (fig. 2). Relationships are bi-directional and therefore read from the perspective of either entity in relationship to the other. Text along the line may be placed to clarify the nature of the relationship. For example, "Vegetation" can be either "Forage" or "Habitat" to a "Bird" (fig. 2). Other examples of CDM modeling as applied to science applications include genetics (Bornberg-Bauer and others, 2002) and metadata for remote sensing data (National Atmospheric and Space Administration, 2001).
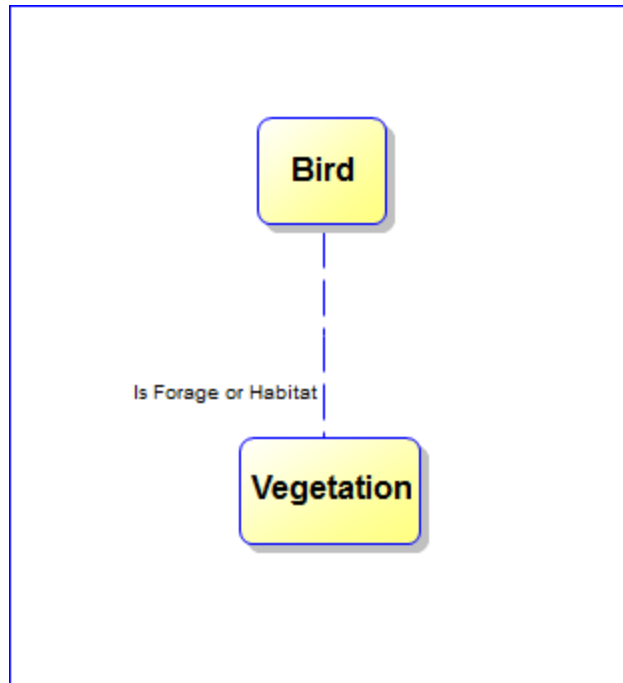
Figure 2. An example of a simple, conceptual data model showing relationship between two data entities.

Process Steps

A number of USGS CAE studies were already underway, therefore the effort was focused on documenting the data currently being collected as part of the initiative. The process of constructing CDMs began with characterizing and organizing the various types of physical and biological data being collected. We developed an approach to understand general attribute groupings, called "data collections." Data collections are groupings of like data. Precipitation, temperature, and wind speed would be examples of "climate" data collections. Interviews with project scientists were conducted to identify entities and data collections from studies taking place across the terrestrial ecosystem study region of northern Alaska (fig. 1). We standardized entity and data collection naming, scope, and definitions across the studies. Standardization facilitates comparative analysis through examining commonality of data across the studies.

Once the CDMs were completed for each study, we developed an overall model of the CAE terrestrial initiative, called an enterprise CDM (Kendle, 2005). The CAE enterprise CDM is essentially an aggregation of the study models. An enterprise CDM is intended to portray the data for a larger domain, such as an organization or subject area as a unified "enterprise" view. In our use, the enterprise CDM portrays the scope of the CAE terrestrial initiative. The enterprise model was integrated with the study models such that standards changes could be addressed at the enterprise level and pushed outward to the affected studies, thus maintaining consistent standards between study models.

A best practices workflow would consist of constructing the enterprise model first, portraying the overall view of the initiative as it was being planned. Then the study models would be constructed as the individual studies were being planned. The CDM models would inherit any entities or data collections from the enterprise model as relevant to the study. As the study identified new subject areas for study and additional data for collection, those additions would be reflected in the enterprise model, thus establishing new standards for future use by other studies. Because this modeling effort occurred after the initiative and the studies were well underway, the process was one of reverse engineering. By first documenting the studies, this enabled an understanding with which to construct the enterprise CDM as a collective of the terrestrial studies of CAE.
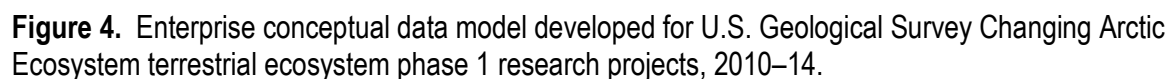
The CDMs were documented using a Computer Aided Software engineering (CASE) tool called PowerDesigner® (business process and data modeling software, SAP®, https://help.sap.com/powerdesigner). Using the tool, CDMs were developed for each study as a graphic model. We developed data dictionaries for each CDM identifying the data standards for naming, scope, and definitions applied to that model (fig. 3).

CDM methodology lends itself to rapid and iterative development throughout a planning phase. In like fashion, the CAE terrestrial CDMs were quickly developed with the most effort spent standardizing the data collections. After the CDMs were completed, we conducted peer reviews of the models using principal investigators of each study to verify that data interpretations and groupings were accurate.

Vegetation

| Description | A general term for plant life and ground cover found at a study site. |

List of Vegetation Data

Name	Comment
Species Composition	Derived data from distance sampling and population estimation statistics. The composition of species observed across a set of point-count surveys or at a single survey location.
Plant Density	Count of plant stems found within a set area as determined by study objectives.
Plant Biomass	Amount of plant mass (in grams). May be quantified for above and or below ground vegetation depending upon study objectives. May be sampled across a season and in relation to experimental

Figure 3. Excerpt of the data dictionary listing the standardized name and meaning of the data collected. See appendix A for complete enterprise data dictionary.

Results

The enterprise CDM developed for the USGS CAE initiative portrays entities relative to the study of terrestrial ecosystem mammals (Bird, Ungulate), breeding success (Brood, Nest), abundance (Bird Population), forage quality (Fish, Vegetation, Invertebrate), habitat type (Habitat and sub-entities: Terrestrial, Fresh Water and Marine, and Vegetation) and physical factors affecting habitat quality (Climate, Hydrology, Soil) (fig. 4). The entity "Habitat" consists of three overall habitat types that are within the USGS CAE initiative (Terrestrial, Freshwater, and Marine). The model depicts entities grouped by physical and biological factors. For example, physical factors of "Climate," "Hydrology," and "Soil" are arranged below the Habitat entity and are considered bottom-up drivers of habitat change. Biological entities related to different classes of wildlife (fish, ungulates, and birds), are interrelated through entities such as "Site" and "Vegetation," but located within separate networks of the CDM. The enterprise CDM in this effort was developed for the terrestrial ecosystem, but the "Marine" habitat is included to demonstrate the overall effort and that entities can be added as place holders for future development. The collective data dictionary of entity and collection definitions is listed in appendix A.

Figure 4. Enterprise conceptual data model developed for U.S. Geological Survey Changing Arctic Ecosystem terrestrial ecosystem phase 1 research projects, 2010–14.

From the terrestrial ecosystem research theme of the USGS CAE initiative, we generated five individual CDMs for the following research studies: (1) Climate Effects on Birds in the Boreal-Arctic Transition Zone, (2) Hydrologic Influences on Arctic Coastal Plain Ecosystems, (3) Distributional Shifts of Molting Black Brant Geese, (4) Effects of Changing Plant Phenology on Arctic Birds, and (5) Climate Effects on Ungulate Populations (see fig. 5 for example; appendix B for models). Each study model uses the same symbology as the enterprise CDM and portrays the entities and data collections relative to that study. A data dictionary was generated for each study.
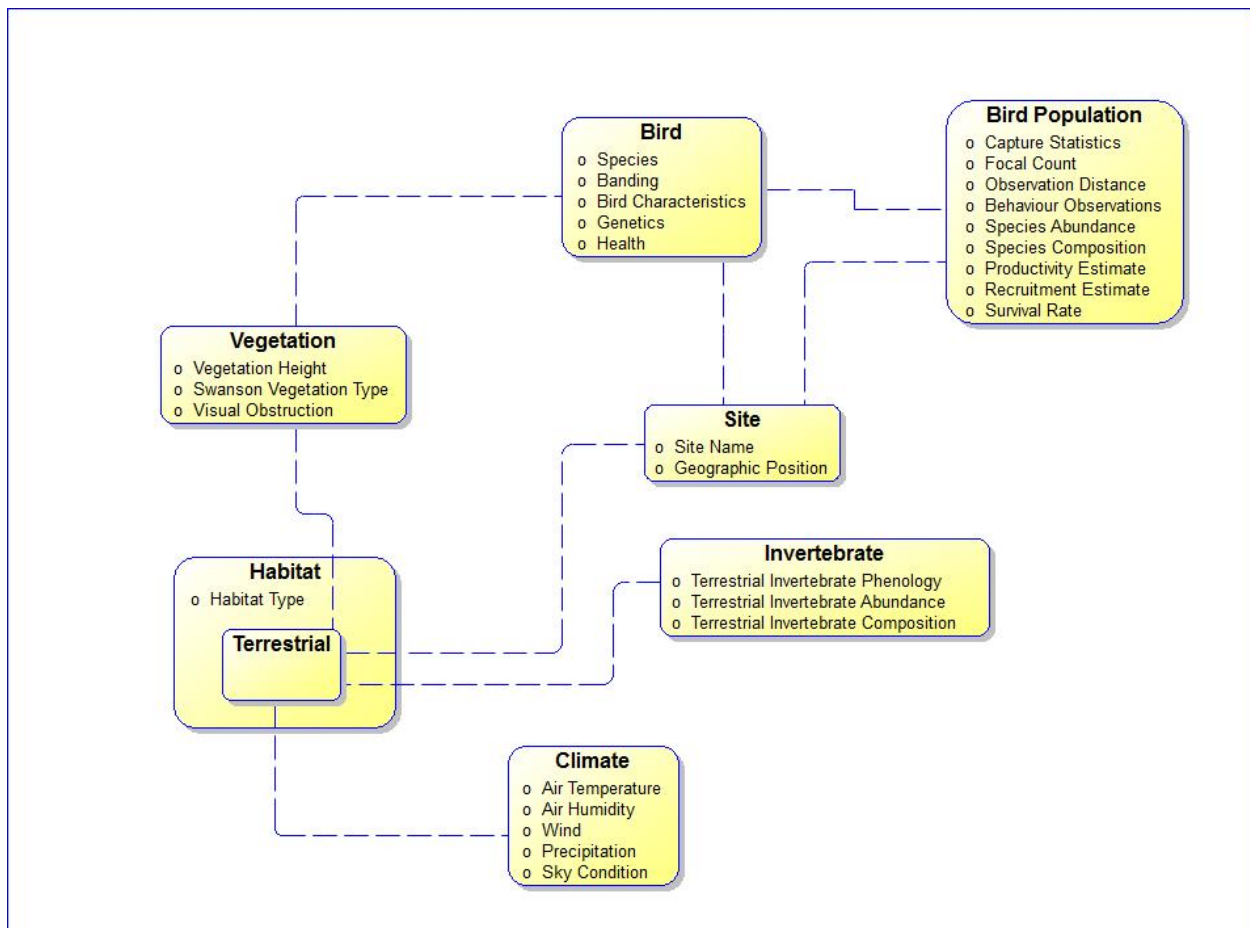

Figure 5. Example of a study Conceptual Data Model for the project, Climate Effects on Birds in the Boreal-Arctic Transition Zone study, taking place in western Alaska near Nome.

Discussion

The final CDMs of different terrestrial ecosystem studies being conducted by the USGS CAE initiative and standardized data collections made it easy to visualize and compare kinds of data being collected between studies and between geographic locations. As expected, some data are commonly collected, such as climate data, whereas other data may be unique to one or two studies. In comparisons, no major deficits in data collection were revealed, but the models did document currently collected data in an easy to visualize form. The value of CDMs would be during the planning phases of studies and initiatives when discussions are initially taking place about project design and the influence of individual study decisions can be seen on the initiative and as compared to other studies and geographic locations.

An example of using a CDM to facilitate discussion was demonstrated when this approach was presented to and subsequently adopted by the Arctic Landscape Conservation Cooperative (Arctic LCC; http://arcticlcc.org), for a project to develop a long term terrestrial climate change observational system for the North Slope of Alaska called the Terrestrial Environmental Observation Network (TEON; http://arcticlcc.org/projects/teon). The CDM was developed for the project early in the planning stage to provide a first level understanding of project scope to researchers from different disciplines and organizations joining the project. The CDM was developed very quickly and has been readily understood by all involved with the project. The CDM also was used as a gateway to more detailed analysis during planning discussion.

Comparisons of CAE study CDMs revealed where the potential for integrating data between studies exists, thus allowing for greater use of the data in analysis and data products. However, the CDMs have limitations in illustrating the compatibility of like data because of their generalized nature. For instance, does "Plant Biomass" really mean the same thing between different studies? Are the attributes the same and are they collected to the same standards? Data modeling at a more detailed level would more clearly reveal the compatibility and potential for integrating like data across studies and disciplines and establishing standards for like data. An application of this detailed analysis has been applied with the CAE Ungulates Forage study and again the Arctic LCC TEON project. In both cases, the subject areas were modeled in greater detail to the attribute level using a method called logical data modeling (LDM); and, the details of the data can be seen and potentially compared with other studies. In the case of Arctic LCC TEON, the tasks of data modeling and project design ran parallel in a collaborative fashion in which one informed the other. The data models were used as a means of collective discussion between researchers about project design and requirements. Conversely, decisions made about the project affecting data were in turn reflected in the data models documenting the data requirements of the initiative. In the CAE Ungulates Forage and Arctic LCC TEON examples, the modeling is being used to design physical implementations for data collection and storage. In this way, CDMs show another potential in being the first step to a more detailed analysis where project planning and design are benefitted by an accompanying modeling effort.

In consideration of future studies within the USGS CAE initiative, this approach also shows potential for data acquisition planning when a new study is modeled and compared to existing enterprise and study CDMs. This may offer opportunities to consider not only what data best inform the study, but also how data may inform other studies and the initiative as a whole. This approach also would reveal the potential for data integration before data collection begins, thus providing the opportunity to design initially for integration, rather than coaxing data together from multiple studies after data are already collected into particular physical forms, often with varying standards.

CDMs are a rapid approach to provide an easy to understand documentation that can facilitate discussion and inform interested parties on overall data acquisition strategies. It facilitates consideration of an individual study's effect to other studies and the initiative as a whole. CDMs can provide an easy means to understand how data from past, present, and future studies intersect, and have shown their potential to facilitate discussion between researchers from multiple organizations and disciplines regarding project scope.

Future Directions

- One of the questions posed during this effort is how the CDMs can facilitate a landscape-based understanding of data acquisition strategies. For instance, in which geographic locations are plant biomass being collected, and what can be inferred by that distribution? Discussed was the possibility of a geographic information system (GIS) tool that could incorporate the information from the CDMs into a geographic-based presentation and the potential for providing that information to the general public. Thus, future enhancements of the modeling approach described here could include the following: develop a GIS-based tool that will allow visualization of the data models and portray the content of the models by study site location. This would provide a visual, geospatial context to comparative analysis of data acquisition strategies across northern Alaska.
- Make CDMs available through a web interface. This would provide a searchable form of public outreach regarding how data acquisition by the USGS and collaborators informs research on climate change in the Arctic. The web tool also would allow researchers to explore data that is being collected by whom and where and further facilitate conversations among scientists and the public.
- Expand the enterprise model to include additional studies as relevant to the USGS CAE initiative and other science initiatives at the Alaska Science Center. In an organization whose primary asset is data, the enterprise model represents a valuable high-level and integrative accounting of the data assets retained by the program and by the organization in terms of research efforts enacted past, present, and planned. This would enable a broader assessment of research data, enabling comparative analysis across a broader spectrum of research data.
- Consider a detailed analysis of high value data crossing multiple studies, with an interest in developing minimum attribute data standards to ensure an adequate level of robustness and integrity in the data collected that will support cross-study and cross-program integration and collaboration of data. These standards can support current and future data integration, as well as aid in the integration of observed, historical, and postulated values.
- Develop data model templates outlining minimum entities needed for future studies.
- Develop study CDMs for future USGS CAE studies thereby starting discussion about data acquisition strategy in regards to program objectives early in the project planning process. Integrate new studies into the enterprise CDM, thereby maintaining an accurate view of study data.

Acknowledgments

References Cited

American National Standards Institute (ANSI), 1975, ANSI/X3/SPARC study group on data base management systems, interim report: FDT, ACM SIGMOD Bulletin, v. 7, no. 2, 140 p.

Bornberg-Bauer, E., and Paton, N.W., 2002, Conceptual data modeling for bioinformatics: Briefings in Bioinformatics, v. 3, p.166–180.

Chen, P., 1976, The entity-relationship model-toward a unified view of data: ACM Transactions on Database Systems, v. 1, p. 9–36, doi:10.1145/320434.320440.

Codd, E.F., 1970, A relational model of data for large shared data banks: Communication of the ACM, v. 13, p. 173–189, doi:10.1145/362384.362685.

Durner, G.M., Whiteman, J.P., Harlow, H.J., Amstrup, S.C., Regehr, E.V., and Ben-David, M., 2011, Consequences of long-distance swimming and travel over deep-water ice for a female polar bear during a year of extreme sea ice retreat: Polar Biology, v. 34, p. 975–984, doi:10.1007/s00300-010-0953-2.

Finstad, G., 2008, Applied range ecology of reindeer (*Rangifer tarandus tarandus*) on the Seward Peninsula, Alaska: Fairbanks, Alaska, University of Alaska-Fairbanks, Ph.D. dissertation.

Flint, P.L., Mallek, E.J., King, R.J., Schmutz, J.A., Bollinger, K.S., and Derksen, D.V., 2008, Changes in abundance and spatial distribution of geese molting near Teshekpuk Lake, Alaska—Interspecific competition or ecological change?: Polar Biology, v. 31, 549–556, doi: 10.1007/s00300-007-0386-8.

Flint, P.L., Whalen, M.E., and Pearce, J.M., 2014, Changing Arctic Ecosystems—Sea ice decline, permafrost thaw, and benefits for geese: U.S. Geological Survey Fact Sheet 2014-3088, 2 p. doi:10.3133/fs20143088.

Gustine, D.D., Adams, L.G., Whalen, M.E., and Pearce, J.M., 2014, Changing Arctic Ecosystems–Resilience of caribou to climatic shifts in the Arctic: U.S. Geological Survey Fact Sheet 2014-3103, 2 p., doi:10.3133/fs2014310.

Jay, C.V, Marcot, B.G., and Douglas, D.C., 2011, Projected status of the Pacific walrus (*Odobenus rosmarus divergens*) in the twenty-first century: Polar Biology, v. 34, p. 1065–1084.

Kendle, N., 2005, The enterprise data model: The Data Administration Newsletter, July 2005. [Also available at http://www.tdan.com/view-articles/5205.]

National Atmospheric and Space Administration (NASA), 2001, Release 6A Implementation Earth Science Data Model for the ECSS Project: National Atmospheric and Space Administration, Technical Paper, 266 p.

Oakley, K., Whalen, M., Douglas, D., Udevitz, M., Atwood, T., and Jay, C., 2012, Changing Arctic Ecosystems—Polar bear and walrus response to the rapid decline in Arctic sea ice: U.S. Geological Survey Fact Sheet 2012-3131, 4 p. [Also available at http://pubs.er.usgs.gov/publication/fs20123131.]

Overland, J.E., and Wang, M., 2013, When will the summer Arctic be nearly sea ice free?: Geophysical Research Letters, v. 40, p. 2097–2101.

Pearce, J., DeGange, T., Flint, P., Fondell, T., Gustine, D., Holland-Bartels, L., Hope, A., Hupp, J., Koch, J., Talbot, S., Ward, D., and Whalen, M., 2012, Changing Arctic Ecosystems—Measuring and forecasting the response of Alaska's terrestrial ecosystem to a warming climate. U.S. Geological Survey Fact Sheet 2012-3144, 4 p. [Also available at http://pubs.er.usgs.gov/publication/fs20123144.]

Ravens, T.M., Jones, B.M., Zhang, J., Arp, C.D., and Schmutz, J.A., 2012, Process-based coastal erosion modeling for Drew Point, North Slope, Alaska: Journal of Waterway, Port, Coastal, and Ocean Engineering, v. 138, p. 122–130.

Reed, B.C., Brown, J.F., and VanderZee, D., 1994, Measuring phenological variability from satellite imagery: Journal of Vegetation Science, v. 5, p. 703-714.

Stafford, J.M., Wendler, G., and Curtis, J., 2000, Temperature and precipitation of Alaska—50 year trend analysis: Theoretical and Applied Climatology, v. 67, p. 33–44.

Stone, R.S., Dutton, E.G., Harris, J.M., and Longenecker, D., 2002, Earlier spring snowmelt in northern Alaska as an indicator of climate change: Journal of Geophysical Research Atmospheres, v. 107, p. 10–13.

Stroeve, J., Holland, M.M., Meier, W., Scambos, T., and Serreze, M., 2007, Arctic sea ice decline: Faster than forecast: Geophysical Research Letters, v. 34, p. L09501, doi:10.1029/2007GL029703.

Tape, K., Flint, P.L., Meixell, B.W., and Gaglioti, B., 2013, Inundation, sedimentation, and subsidence creates goose habitat along the Arctic coast of Alaska: Environmental Research Letters, v. 8, p. 045031, doi:10.1088/1748-9326/8/4/045031.

Uher-Koch, B.D., Schmutz, J.A., Whalen, M.E., and Pearce, J.M., 2014, Changing Arctic Ecosystems—Ecology of loons in a changing Arctic: U.S. Geological Survey Fact Sheet 2014–3093, 2 p., doi:10.3133/fs20143093h.

Van Hemert, C.R., Flint, P.L., Udevitz, M.S., Koch, J.C., Atwood, T.C., Oakley, K.L., and Pearce, J.M., 2015, Forecasting wildlife response to rapid warming in the Alaskan Arctic: BioScience, v. 65, p. 718–728, doi: 10.1093/biosci/biv069.

Zachman, J.A., 1987, The Zachman framework for enterprise architecture: IBM Systems Journal, v. 26, no. 3, p. 276–292.

Appendix A. U.S. Geological Survey Changing Arctic Ecosystems Enterprise Data Dictionary Report

Summary: This model describes at an enterprise level, the data entities and data collections used in the U.S. Geological Survey Changing Arctic Ecosystems Initiative.

Contents

Bird

  Description   An individual bird that is part of a population under study. An individual
                          may be identified with a band, web tag, nest identifier, or hunter identifier.

List of Bird Data

Name	Description
Species	Taxonomic identification of an animal under study. Can be a common name or Latin species name.
Banding	Data collected from a bird during metal banding. Data may include physical attributes of the bird (weight, measurements, fat, skull ossification, age, feather wear, molt status, breeding status), USGS metal band number, color code combination applied to leg, other auxiliary markers (neck collar, radio transmitter, flag band, geolocator), and other data such as date, time, and location of banding.
Bird Characteristics	Other notes relating to general features or observations made at the time of capture and banding.
Genetics	Sample ID number and type of sample collected (blood, feather, swab, tissue, egg).
Health	General health status (normal, deformed, sick) or other sample (swab, blood, tissue, feather) that will be used to determine exposure to disease or contaminants.
Habitat Use	Derived data from studies of marked birds (typically with radio transmitters) of the habitat used during molting or brood rearing. Determined by mapping movements onto land cover images and quantifying use of broad habitat classes.

List of Business Rules

Name	Description
Bird Species Identification	A bird species can be identified using the American Ornithologists' Union four letter alpha species code (see http://www.pwrc.usgs.gov/bbl/MANUAL/bandsize.cfm).

Bird Population

Description   A collection of individual bird data from a study area that is used for scientific research.

List of Bird Population Data

Name	Description
Species	Taxonomic identification of an animal under study. Can be a common name or Latin species name.
Capture Statistics	Number of total birds captured, banded or otherwise marked, number of recaptures from prior years or other banding locations within 1 year.
Gosling Growth Rate	Derived data from repeat captures of same individuals over a season to determine growth rate curve relative to forage quality or other environmental conditions.
Productivity Estimate	Derived data from nest and brood monitoring over a season.
Recruitment Estimate	Derived data from multiple years of capture and banding efforts. The number of juveniles that return to a population in a subsequent year for breeding.
Survival Rate	Derived data from multiple years of capture and recapture or resighting data. The number of birds that survive to be seen on a study area across multiple years.
First Arrival Date	Derived data using several possible methods, including: first recorded date of resighting of an auxiliary marker of a previously marked bird; first observation in spring of a given species; first detection of a previously marked bird using remote tracking devices such as satellite or other radio transmitter previously applied to a bird.
Nest Initiation Date	Date when egg laying was estimated to have started. The data are based on nests observed during laying and those for which the embryo age was recorded during a visit when incubation was underway. Where more than one estimate was available for a nest, the estimate was based on the earliest visit, under the assumption that it is the most accurate.
First Departure Date	Derived data using several possible methods, including: last recorded date of resighting of an auxiliary marker of a previously marked bird; last observation in spring of a given species; last detection of a previously marked bird using remote tracking devices such as satellite or other radio transmitter previously applied to a bird.
Focal Count	Used in distance sampling and population estimation statistics. The number of bird species and individuals recorded at a particular location for a given length of survey time.
Observation Distance	Used in distance sampling and population estimation statistics. Distance at which a focal observation was made during a point-count survey.
Behavior Observations	Used in distance sampling and population estimation statistics. Observed behavior of a species during a point-count survey. Behaviors may include, flying, standing, calling, displaying, foraging.
Species Composition	Derived data from distance sampling and population estimation statistics. The composition of species observed across a set of point-count surveys or at a single survey location.
Species Abundance	Derived data from distance sampling and population estimation statistics. The number of species observed across a set of point-count surveys or at a single survey location.

Brood

Description A group of young (hatch year) birds associated with one or more adults. Typically not found at a nest. A term used after hatch and once chicks are mobile.

List of Brood Data

Name	Description
Brood Size	Count of young (hatch year) birds in a brood.

Climate

Description An observation of weather conditions and light intensity at a location. Observation of climatological effects on habitat.

List of Climate Data

Name	Description
Air Temperature	Measured air temperature, in degrees Celsius.
Air Humidity	Relative humidity expressed as a percentage.
Water Temperature	Measured temperature of surface and subsurface water, in degrees Celsius.
Wind	Wind direction in compass degrees and wind speed, in kilometers per hour.
Precipitation	Measured liquid precipitation in millimeters. May be reported on a daily, monthly, seasonal, or annual basis.
Barometric Preasure	Measured atmospheric pressure, in millibars of mercury.
Sky Condition	Observed sky cover classification.
Light Intensity	Measured incoming solar radiation in watts/m^2, logged every 15 minutes.
Growing Degree Days	Growing degree days estimated for a given day. Growing degree days equals the degrees above 0 degrees C for a Julian date with a minimum temperature greater than 0 degrees C (Finstad, 2008). Calculated from hourly temperatures acquired from Hobo data loggers deployed between May 12 and September 30.

Fish

Description A collection of individual fish samples that make up a population under study. Individuals may be collected or sampled and released.

List of Fish Data

Name	Description
Species	Taxonomic identification of an animal under study. Can be a common name or Latin species name.

Habitat

Description   A landscape area inhabited by a particular species or group of species. Habitat also contains the forage and physical environmental characteristics necessary for the species or group of species to exist.

List of Habitat Data

Name	Description
Habitat Type	A particular definition or description of a landscape area used by a species or group of species. Naming of habitats will vary depending upon research focus and species under study.
Seasonal Phenology Image	Image record of visible landscape changes at timed intervals throughout the growing season. Records changes in thaw, surface water, and vegetation cycles.

Freshwater

Description   A broad habitat classification that includes wetlands, lakes, and rivers.

List of Freshwater Data

Name	Description
Water Body Type	General classification of a water body based on size, depth, shoreline, and physical characteristics (for example, thaw lake, polygonal pond, stream, trough).
Water Body Size	Size characteristics of a water body. May include measures of length and width, volume, depth, radius, and circumference.
Water Temperature	Measured temperature of surface and sub-surface water in degrees Celsius.
Chemical Composition	Measured from samples collected from water bodies. Analytical composition will vary depending upon study objectives.
Environmental Treatment	Descriptions of experimental additions of heat, shade, and/or nutrients to water bodies as part of research studies.

Marine

Description   A broad habitat classification that includes nearshore and offshore habitats.

Terrestrial

Description   A broad habitat classification that includes onshore estuarine, wetland, xeric, mesic, upland, and alpine habitats.

List of Terrestrial Data

Name	Description
Environmental Treatment	Descriptions of experimental additions of heat, shade, and/or nutrients to water bodies as part of research studies.

Hydrology

Description   Hydrology involves the quantifying the presence on the landscape and in soils, as well as the movement of water between soils, surface pools, and the atmosphere. Measurements include surface water extent and discharge, atmospheric fluxes, chemistry, and subsurface flow potential.

List of Hydrology Data

Name	Description
Snow Coverage	Percentage of ground covered by snow on a given date as determined by visual observation across time or by time-lapse cameras.
Snow Depth	Depth of snow in centimeters on a given date, collected at points associated with one snow density measurement.
Snow Density	Measurement of the density of snow at a location. Collected using a snow corer at multiple locations representative of the dominant landscape features.
Water Coverage	Estimated percent of land covered by water at a specific time, as determined by a few sparsely distributed cameras.
Water Depth	Measured depth of a water body. Measured in multiple ways: Small ponds are usually measured with a graduated rod by averaging measurements at three locations. Streams and large lakes are measured with a Level Logger pressure transducer that records pressure in pounds per square inch (ps**i)** every 15 minutes. Stream and lake depths are independently verified during site visits, using a measuring tape or wading rod.
Thaw Depth	Depth of thawed soil measured in centimeters. Measured with a steel probe and measuring tape from the top of the frozen ground to the top of the ground surface. Measured at multiple times and locations across a season. Active layer is collected at multiple sites across one polygon transect, and at other random times and locations associated with experimental needs.
Stream Discharge	Measurement or estimation of streamflow. Determined using a wading rod and pygmy meter during site visits. Interpolated between site visits using stage-discharge rating curves developed during site visits, and increasingly using a model based on Manning's equation, with slope determined from ArcGIS and Manning's roughness determined by solving the equation for instances when discharge was independently verified with a rod and pygmy meter.
Groundwater Discharge	Measurement or estimate of groundwater flow. Presently, only subsurface hydraulic conductivity is being measured (soil property). This data may be combined with lidar slope estimates to predict groundwater discharge. In the summer of 2013, infiltration experiments will be used to quantify flow potential under a realistic precipitation rate.

Invertebrate

Description   Animal species that do not develop a vertebral column. A common forage item for birds.

List of Invertebrate Data

Name	Description
Terrestrial Invertebrate Phenology	Seasonal timing of invertebrate species found on land. May include dates of activity and emergence.
Terrestrial Invertebrate Abundance	Quantity of invertebrates (either collectively or summarized by taxonomic order) collected at a sampling locale. May be summarized across a season to determine timing of emergence and peak activity.
Terrestrial Invertebrate Composition	Species composition of invertebrate species per sample unit or across a season.
Aquatic Invertebrate Phenology	Seasonal timing of invertebrate species found in the water or at the water/terrestrial border. May include dates of activity and emergence.
Aquatic Invertebrate Abundance	Quantity of invertebrates (either collectively or summarized by taxonomic order) collected at a sampling locale. May be summarized across a season to determine timing of emergence and peak activity.
Aquatic Invertebrate Composition	Species composition of invertebrate species per sample unit or across a season.
Phytoplankton Abundance	Quantity of phytoplankton collected at a sampling locale. May be summarized across a season to determine timing of emergence and peak activity.

Nest

Description   A location at which a bird lays and incubates its eggs. For altricial species (for example, Passerines), also the location where adult birds feed and raise their young.

List of Nest Data

Name	Description
Species	Taxonomic identification of an animal under study. Can be a common name or Latin species name.
Nest Characteristics	May include vegetation characteristics (species type and composition), depth, size, vegetation cover (vertical obstruction around the nest), distance to water, number of eggs, and status (active, depredated, hatched)
Nest Success	Derived data from multiple observations of a sample of nests. Proportion of nests that successfully hatch one or more eggs.

Site

Description   A location at which a study has been conducted.

List of Site Data

Name	Description
Geographic Position	A position on the Earth's surface, usually expressed in latitude and longitude. May be an explicit or representative location of a study site.
Physiography	The physical characteristics of the land.
Site Name	Unique name and/or identifier of a site.

Soil

Description   Soils are a heterogeneous mixture of mineral, organic, and decaying plant material, water, and air. Measurements may include soil moisture, temperature, active layer depth, depth of various horizons, and chemistry.

List of Soil Data

Name	Description
Elemental Composition	Concentrations of carbon and nitrogen within a soil sample.
Soil Composition	Measurements of the thickness and type (that is, mineral or organic) of various soil horizons.
Soil Temperature	Measured on the surface and subsurface, in degrees Celsius. For subsurface, measured at 10, 20, 40, and 80 centimeters, every hour throughout the year.

Ungulate Herd

Description   Historical geographic extent of specific Caribou herds. The herd names are: Central Arctic, Western Arctic, and Teshekpuk Lake herds.

Vegetation

Description   A general term for plant life and ground cover found at a study site.

List of Vegetation Data

Name	Description
Species Composition	Derived data from distance sampling and population estimation statistics. The composition of species observed across a set of point-count surveys or at a single survey location.
Plant Density	Count of plant stems found within a set area as determined by study objectives.
Plant Biomass	Amount of plant mass (in grams). May be quantified for above and/or below ground vegetation depending on study objectives. May be sampled across a season and in relation to experimental treatments to determine trends and peak in mass.
Plant Phenology	A stage of vegetative growth as defined by Finstad (2008).
Nutrient Quality	Amount of nutrients (typically measured as percent or concentration of carbon and nitrogen) in a sample of plant material.
Vegetative Light Reflectance	The Normalized Difference Vegetation Index (NDVI) as estimated for the MODIS satellite platform from a UNISPEC-SC handheld Spectrometer (Reed and others, 1994).
Dominant Species	Identification of a species that is the dominant available forage species at a given time during the growing season.
Vegetation Height	Measured height of the tallest plant at a site.
Swanson Vegetation Type	Plant community type as defined by Swanson and others (1985) for the Seward Peninsula of Alaska.
Visual Obstruction	Amount of a site (nest or other) that is visually obscured by vegetation as viewed from a set distance away.

Appendix B. U.S. Geological Survey Changing Arctic Ecosystems Conceptual Data Models for Selected Phase 1 Research Projects (2010–14)
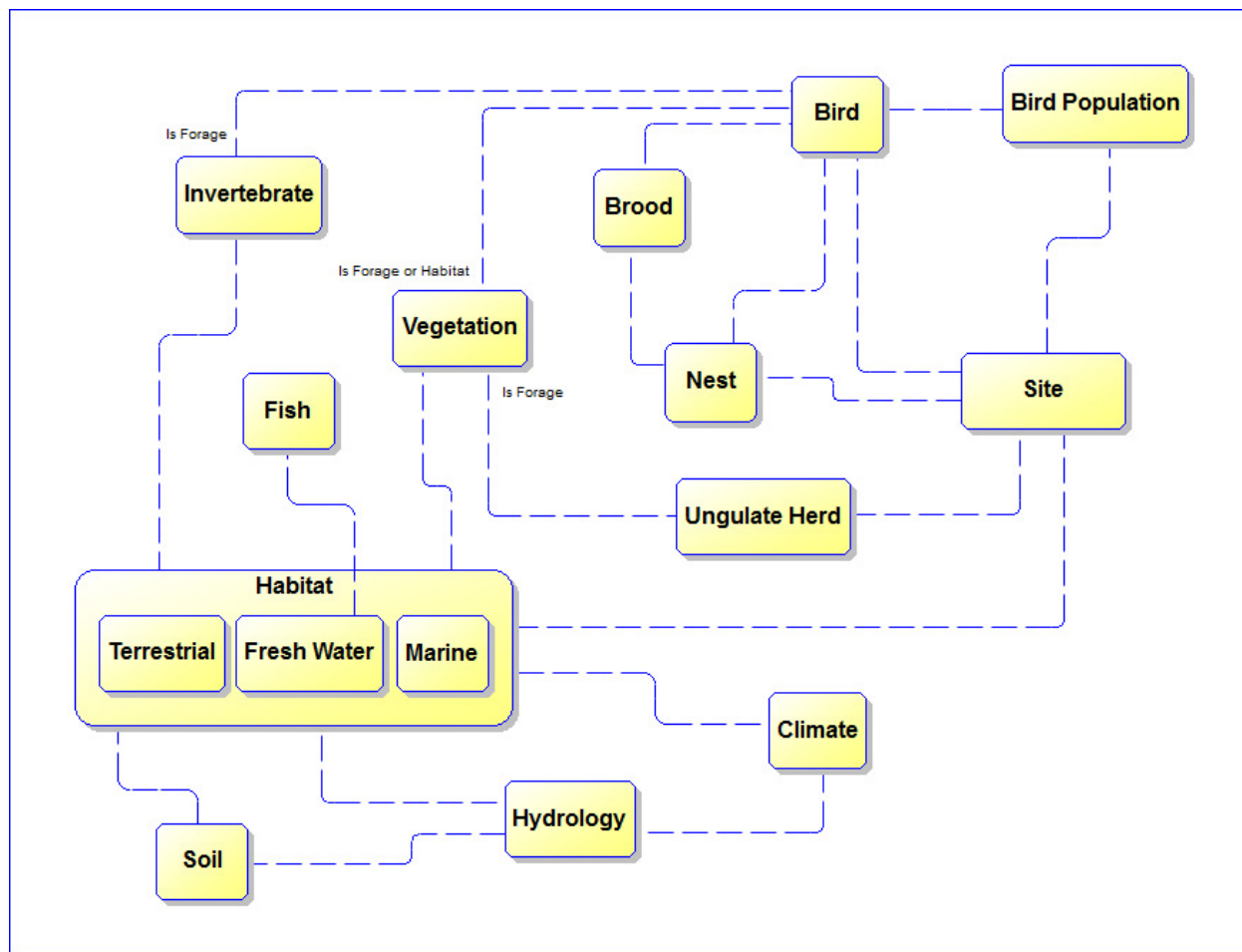

Figure B1. Enterprise conceptual data model developed for the U.S. Geological Survey Changing Arctic Ecosystems initiative.
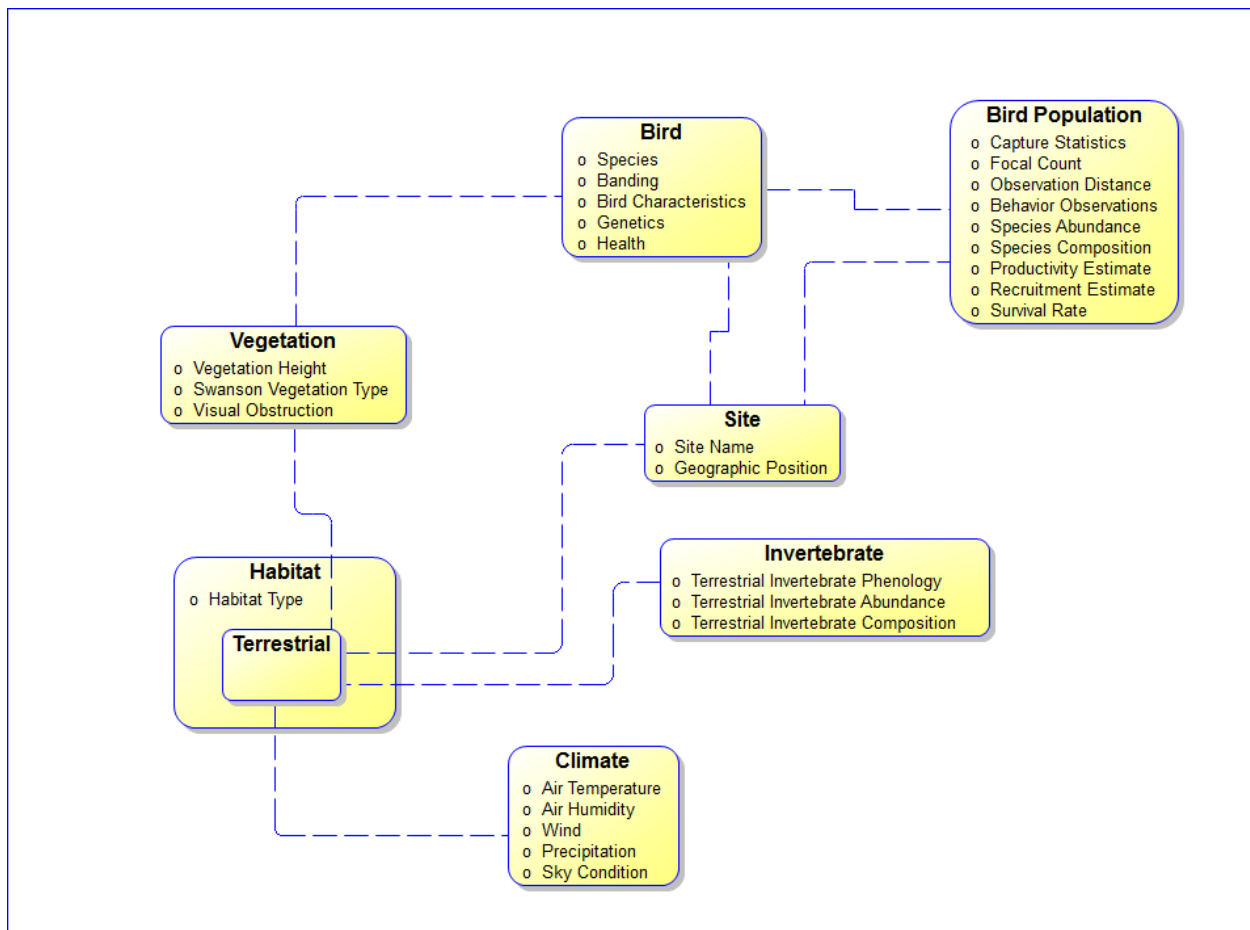
Figure B2. Conceptual data model for the Climate Effects on Birds in the Boreal-Arctic Transition Zone study taking place in western Alaska near Nome.
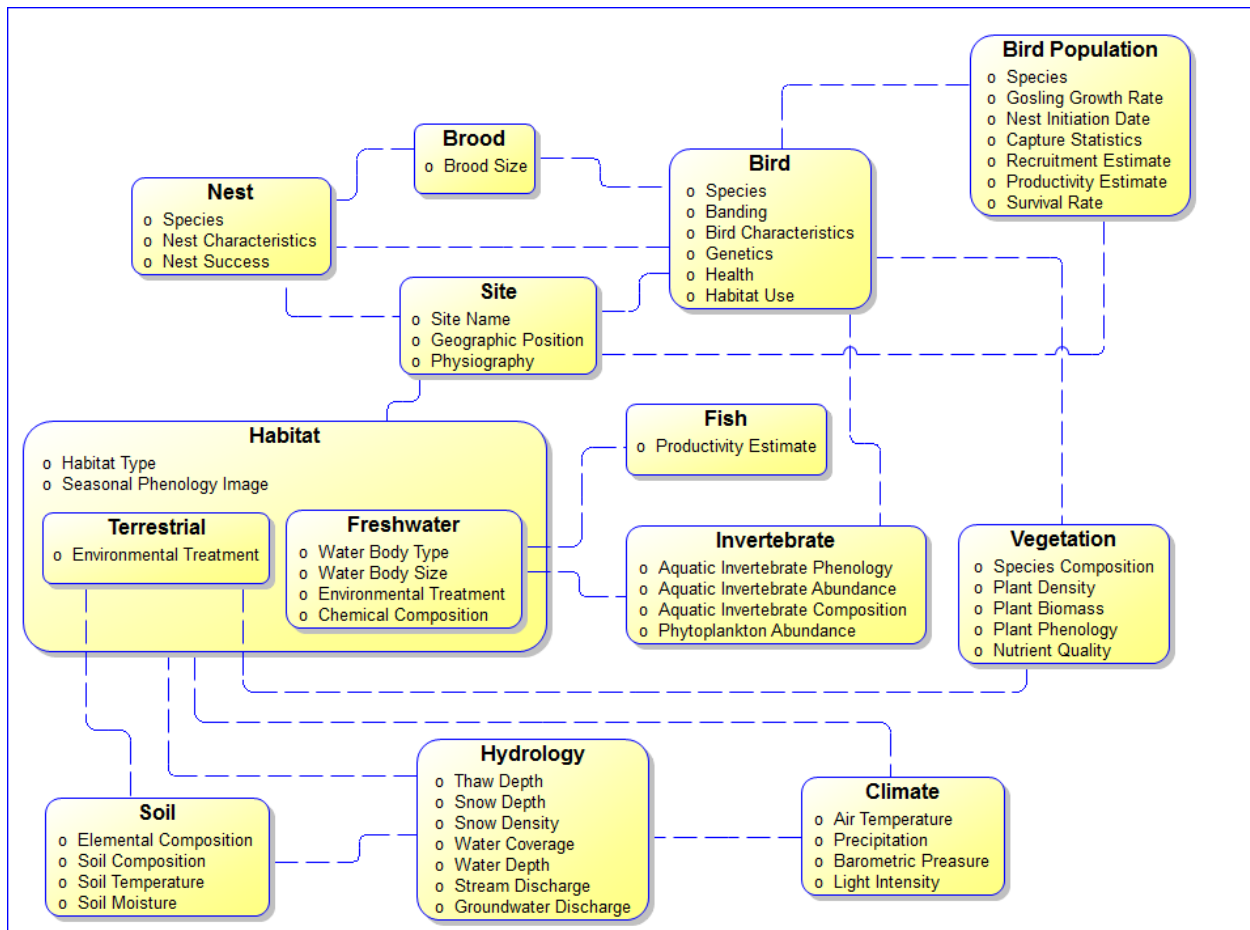
Figure B3. Conceptual data model for the Hydrologic Influences on Arctic Coastal Plain Ecosystems study taking place on the Arctic Coastal Plain of Alaska near Barrow.
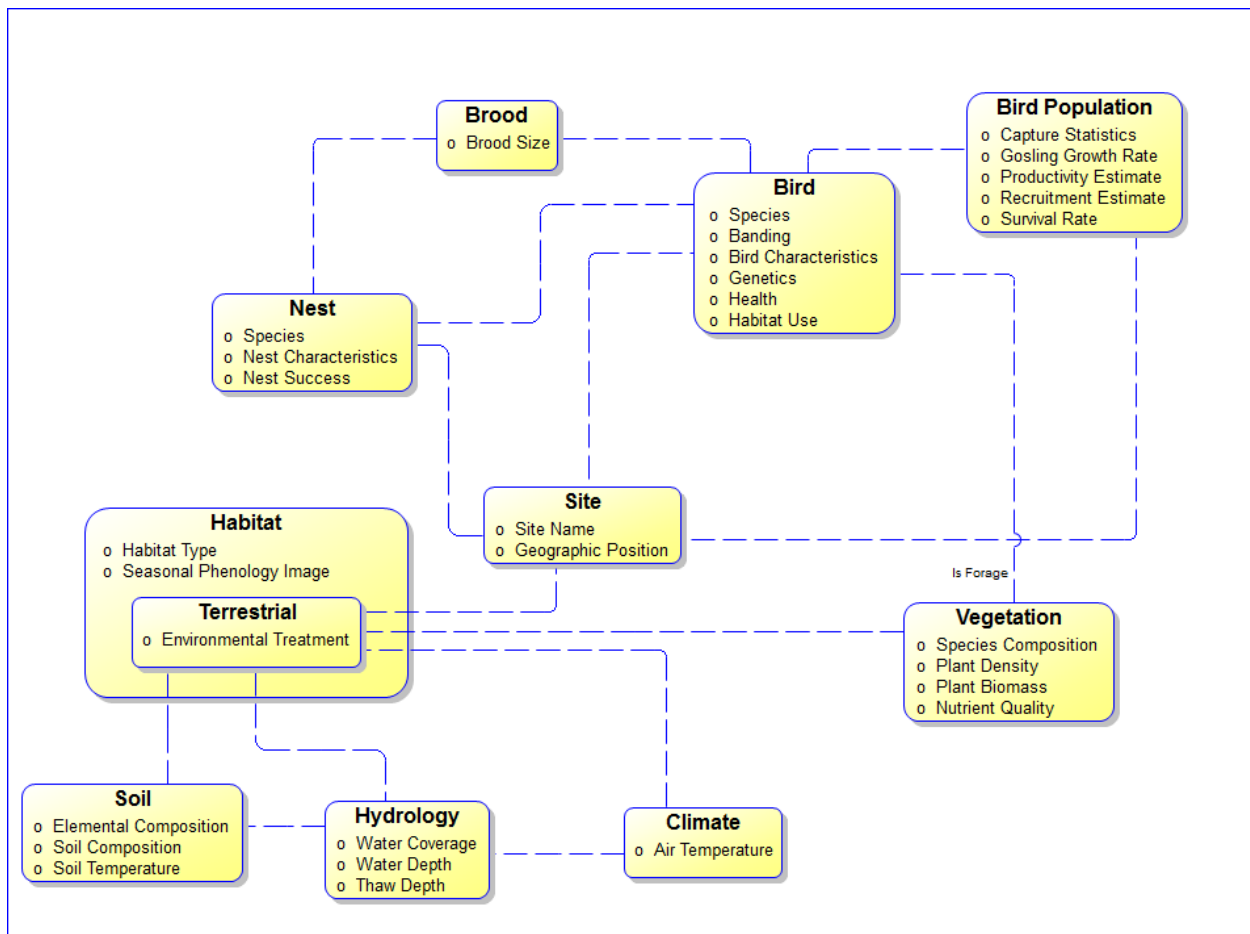
Figure B4. Conceptual data model for the Distributional Shifts of Molting Black Brant Geese study taking place on the Arctic Coastal Plain of Alaska near Barrow.
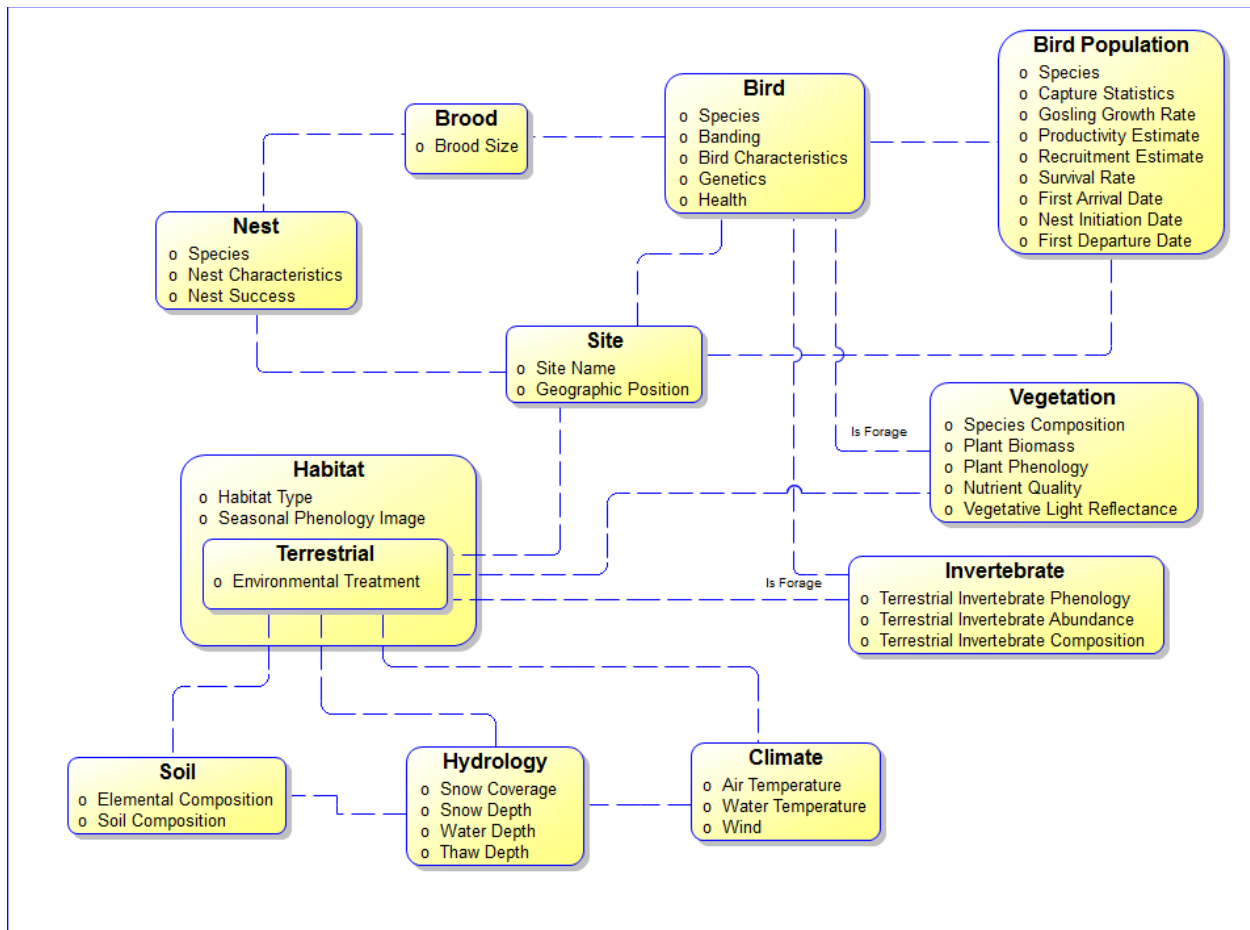
Figure B5. Conceptual data model for the Effects of Changing Plant Phenology on Arctic Birds study taking place on the Arctic Coastal Plain of Alaska near Barrow.
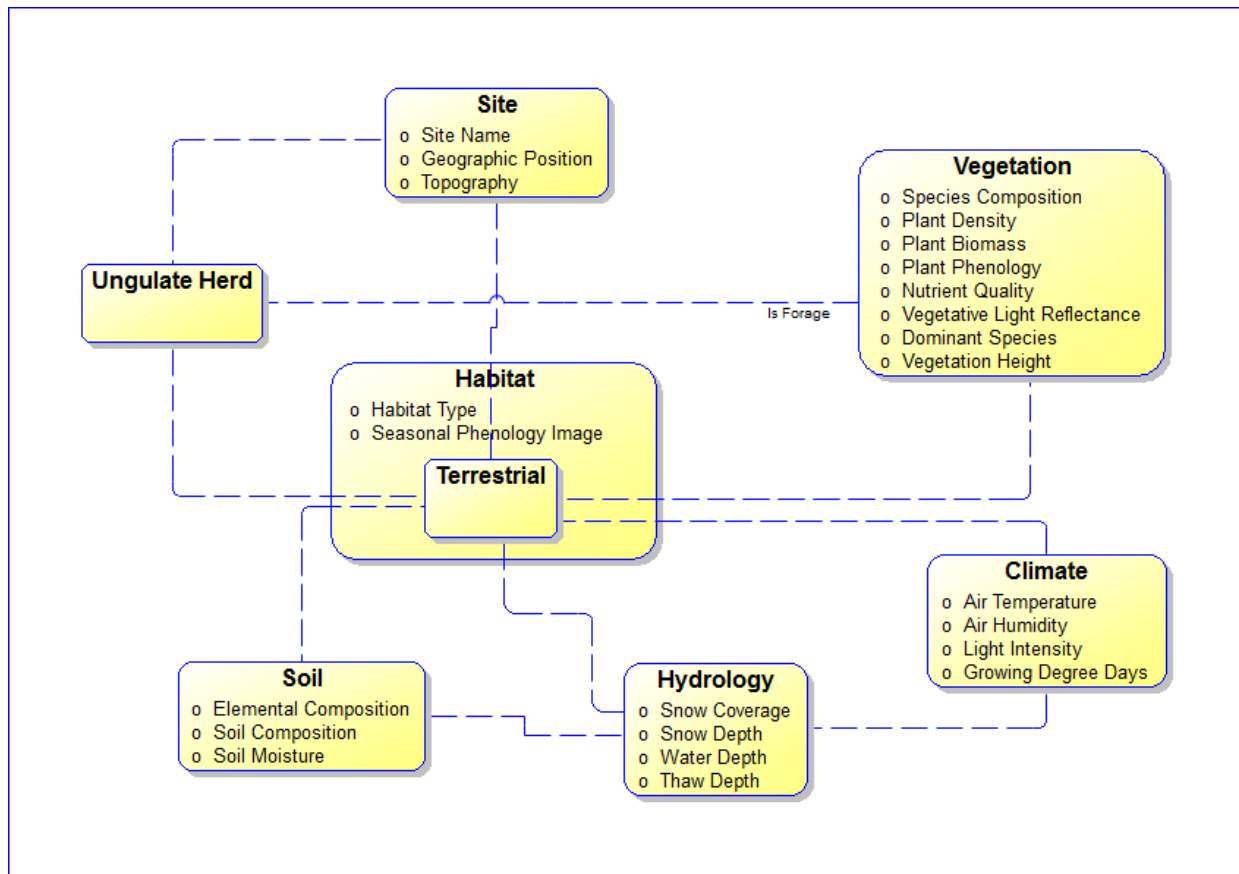
Figure B6. Conceptual data model for the Climate Effects on Ungulate Populations study taking place along a transect from the Arctic Coastal Plain of Alaska south to the Brooks Range near Deadhorse.

Walworth and Pearce—**Conceptual Data Modeling of Wildlife Response Indicators to Ecosystem Change in the Arctic**—Open-File Report 2015–1148

USGS