

## 6. DATA AND INFORMATION TECHNOLOGY

**Lead authors:** J.G. Masek and V. N. Razuvaev

**Contributing authors:** V. Gershenson and P. Ya. Groisman

### 6.1. NEESPI Data Requirements

Observations are an important component of the regional science program and are essential for environmental monitoring. Fortunately, there is a large overlap between the observations needed to support global change research and the observations needed for environmental and natural resource management. The operational monitoring systems necessary to provide the systematic observations are expensive to maintain and hard to justify based on science inquiry alone. No single country can provide the requisite observing systems for regional and global monitoring: international coordination is essential. The observational needs for global change research and resource management are being promoted and coordinated under the international Integrated Global Observing System (IGOS) program<sup>58</sup>.

In general, NEESPI seeks to address the interactions between Northern Eurasian ecosystems, climate, and human activity, using a combined framework of predictive models, long-term in-situ and spatially complete modern remote-sensing observations, and process studies. Ultimately, by developing mature, tested models specific to Northern Eurasia, we hope to develop realistic predictions of how global and regional climate will respond to ecological changes during the next several centuries. In this context, NEESPI observational datasets will be used to:

- parameterize and validate predictive models of ecosystem dynamics and ecosystem-climate interactions;
- provide evidence for long-term trends in Northern Eurasian climate and ecosystems, in response to human and global climate forcing;
- generate new understanding of ecosystem processes that can be incorporated into model physics.

The last point is particularly important, in that existing models used for European or North American ecosystems may neglect processes important in the Northern Eurasian region (e.g. thermokarst activation, forest stand replacement, detailed albedo/snow-cover feedbacks). Thus, identifying and/or creating relevant observational datasets is a vital aspect of the NEESPI science program.

The NEESPI science themes drive specific data requirements. The initiative is formed from three major research thrusts: (1) understanding the biogeochemical cycling, natural and man-induced dynamics, and changes to Northern Eurasian ecosystems, (2) understanding the interactions between the land surface (including terrestrial ecosystems and hydrology) and climate systems, and (3) understanding the linkages between human activities and environmental change. Each of these thrusts has unique requirements for data. Tables 6.1-6.3 present the types of data products of importance to NEESPI; these are discussed in more detail in the sections below.

#### 6.1.1. Terrestrial Ecosystems and Biogeochemical Cycles

---

<sup>58</sup> For the terrestrial component, the Global Observation of Forest Cover and Land Cover Dynamics (GOFC/GOLD) program, which is part of the Global Terrestrial Observing System (GTOS) is developing regional networks of scientists and data users to articulate the observational needs (Townshend et al. 2004).

Of paramount importance are datasets that map the distribution of vegetation composition, structure, dynamics, and biochemistry. These data are essential for parameterizing photosynthetic models of primary productivity. In the forest zone, long-term measurements of forest age and disturbance rate (e.g. forest fire, insect defoliation) will allow the calculation of net ecosystem and biome productivity and the land-atmosphere carbon flux. Observed standing biomass can be used to validate ecosystem productivity models, and to assess the loss of carbon stocks following disturbance. For these purposes, forest inventory data are essential to the NEESPI enterprise. Although access to such inventory data is not uniform across the NEESPI region, the program accepts as a goal the dissemination of sufficient forest ecosystem data to allow parameterization of continental-scale biogeochemical models. Given the importance of methane balance in Northern wetlands, detailed maps of wetland extent and vegetation composition are also important. In the steppe, semi-desert, and desert zones, the primary scientific questions revolve around land degradation, susceptibility of agriculture and pasture land use to changes in climate, and changing land use in response to shifting socio-economic conditions. Addressing these issue require datasets on agricultural extent and land management, as well as socio-economic time series. The potential activation of below-ground carbon in Northern Eurasia necessitates measurements of soil carbon concentration, accumulation, and respiration rate. In tundra and forest-tundra regions, these data can be supplemented with observations of permafrost extent and thermokarst formation in response to measured energy balance.

|                                  |   |
|----------------------------------|---|
| Vegetation Dynamics              | Land-cover type<br>Disturbance history<br>Land-cover change<br>Phenology  |
| Vegetation composition/structure | tree types<br>stand height<br>stand age<br>biomass<br>LAI   |
| Vegetation biochemistry          | fAPAR<br>light use efficiency<br>water use efficiency<br>nutrient limitations<br>CO <sub>2</sub> , methane fluxes |
| Soils                            | soil type<br>soil carbon<br>soil hydrology  |
| Hydrology                        | Carbon export<br>forest ET fluxes   |
| Ecosystem health                 | stand replacement history<br>fragmentation<br>biodiversity metrics  |

Models of carbon exchange fed by observational inputs may be complemented by direct observations of CO<sub>2</sub> and methane fluxes using eddy covariance techniques. There is currently a significant gap in the global FluxNet network within northern Eurasia, and NEESPI should help to coordinate filling this gap. Ideally, such FluxNet observations would be spread across multiple ecosystems (forest, grassland, tundra, etc), and incorporate observations from chronosequences of known age and disturbance history, and sample a

variety of land management regimes. In addition, the emplacement of a small number of “tall” flux towers (100-500m height) would allow validation of regional and continental-scale carbon fluxes.

In addition to understanding land-atmosphere fluxes of carbon, long-term changes in Northern Eurasian ecosystems are of interest for predicting biodiversity and understanding the impacts of regional climate change. Long-term observations of stand replacement, grassland/woodland dynamics, plant and animal inventories, and habitat fragmentation may provide information on ecosystem health in response to changing climate, land use change, and a general anthropogenic load (pollution, water withdrawal, urbanization, timber harvesting, various industrial and agriculture activities).

**6.1.2. Surface Energy and Water Cycles**

The Surface Energy and Water Cycles research element seeks to understand how changes in global climate propagate to Northern Eurasian ecosystems, and, in turn, how ecosystem changes feed back to affect regional and global climate. Observational data are required to parameterize regional and global climate models, to analyze historical trends in surface hydrometeorology, and to evaluate the linkages between terrestrial ecosystems and climate. Meteorological variables are required to determine the transfer of energy and water between the land surface and atmosphere. Surface temperature, air temperature, wind speed, radiation balance, cloudiness, and surface albedo help constrain the sensible and radiative energy balance. Observations of precipitation, relative humidity, and water vapor, combined with temperature and wind speed as a well as direct measurements of evapotranspiration in selected locations help constrain the transfers of water and latent heat.

|                      |  |
|----------------------|--|
| Energy balance       | Cloud fraction, type<br>Surface albedo<br>Surface, air temperature<br>Longwave fluxes<br>Wind speed, direction |
| Water balance        | Relative humidity<br>ET fluxes<br>Precipitation<br>Snowcover, depth<br>River runoff<br>Soil moisture           |
| Atmospheric Dynamics | Barometric pressure<br>Large-scale circulation<br>Topography<br>Aerosol, dust loading                          |

Hydrologic observations are critical to understanding how water availability varies through time, and how it affects terrestrial ecosystems. Surface and subsurface hydrology measurements (runoff, soil moisture, soil permeability, etc) can be combined with meteorological observations (evapotranspiration<sup>59</sup>, precipitation) to constrain the regional water balance. Fluctuations in water availability can be related to observations of ecosystem health and productivity, and to regional climate patterns. The importance of cryospheric processes to the energy and water balance in the region requires measurements of glacier dynamics, snow cover extent, depth, water equivalent, and albedo.

**6.1.3. Land use and socio-economic information**

<sup>59</sup> Evapotranspiration is a calculated variable in most hydrological studies. But, unique observational programs that exist in Northern Eurasia (heat balance and lysimeter networks) allow direct measurements of this variable that may then be used to validate/calibrate the calculations.

One of the most important components of NEESPI is the inclusion of data on socio-economic resources of the region. Changes in the environment, climate, and social systems are closely linked and, in many cases, the wellbeing of regional cultures depend on the provision of natural resources from the environment. Therefore, the NEESPI implementation should provide and use comprehensive information on the state and dynamics of the socio-economic system. This includes:

|  |   |
|--|---|
| Social Information                     | Population and Demographics<br>Health and Life Expectancy<br>Work Activity<br>Education<br>Migration Patterns<br>Income and Wealth<br>Land Tenure |
| Economic Infrastructure and Technology | Employment<br>Industrial Production by Sector<br>Economic Indicators (GDP, etc)<br>Trade Balance<br>Productivity<br>Energy Consumption            |
| Natural Resources                      | Water Resources<br>Agricultural Resources<br>Timber Volume<br>Energy Resources  |

In addition to the present state of resources, we also need to have historical information in order to estimate socio-economic trends and the depletion of resources, key aspects for the decision making process. Using this information in conjunction with data on the environment and climate will allow a comprehensive approach to studying processes that are occurring in the region in all their complexity and feedbacks. Information on resources and socio-economics can be obtained from the reference documents that are routinely prepared by the government institutions of Russia, PRC, The Ukraine, Kazakhstan, and other countries of the region (as well as by the former USSR prior to 1991). Considerable information is currently available via the Internet (e.g., <http://old.priroda.ru>; [http://www.sci.aha.ru/RUS/wab\\_.htm](http://www.sci.aha.ru/RUS/wab_.htm); <http://www.eurasianet.org/eurasianet/resource/>; and others).

## **6. 2. Satellite and In-Situ Data Availability**

### **6.2.1. Satellite Data**

As noted in Chapter 4 of the Science Plan, given the geographic expanse of Northern Eurasia and the relative paucity of certain types of in-situ measurements, much of NEESPI research will rely on satellite remote sensing. The use of remote sensing technologies, carefully calibrated by direct station measurements, may offer a more robust approach for providing this information and filling in the gaps in the in-situ data. Here we review issues related to the availability of various satellite data sets.

**Table 6.4: Satellite Remote Sensing systems relevant for NEESPI science activities.**  
**Italics represent future missions. Resolution in meters unless otherwise specified**

| <b>Coarse Resolution<br/>Passive Optical</b> | <b>Resolution</b> | <b>Operations<br/>Dates</b> | <b>Spectral<br/>Range</b> | <b>Applications</b>   | <b>Data Availability</b>  |
|--|-------------------|-----------------------------|---------------------------|---|---|
| NOAA AVHRR                                   | 1100              | 1978 -                      | VNIR-TIR                  | global vegetation, surface temperature, cloud cover, fire                             | Derived products from USGS EDC, NASA GSFC, UMD                                      |
| RESURS-01 MSU-SK                             | 150               | 1994 -                      | VNIR-TIR                  | global vegetation, surface temperature  | TransparentWorld  |
| Terra MODIS, Aqua MODIS                      | 250-1000          | 2000 -                      | VNIR-SWIR-TIR             | global vegetation, surface temperature, water vapor, clouds, aerosol, fire, snowcover | EOSDIS  |
| SPOT Vegetation                              | 1200              | 1998-                       | VNIR-SWR                  | global vegetation   | CNES  |
| <b>Hi-Res Optical</b>                        |                   |                             |                           |   |   |
| Landsat                                      | 30-90             | 1972 -                      | VNIR-SWIR-TIR             | land-cover/ change, coastal environment   | USGS EDC (paid), UMD Global Land Cover Facility, MSU Landsat.org, Transparent World |
| SPOT XS/Pan                                  | 20                | 1986-                       | VNIR                      | land-cover/ change, coastal environment   | SPOT Image (paid)   |
| Terra ASTER                                  | 30                | 2000 -                      | VNIR-SWIR-TIR             | Land-cover, surface temperature, fire   | EOSDIS  |
| RESURS-01 MSU-E                              | 30                | 1994 -                      | VNIR                      | land-cover/change   | Transparent World   |
| IRS LISS                                     | 25                | 1988 -                      | VNIR                      | land-cover/ change  | Transparent World   |
| SpaceImaging IKONOS                          | <4                | 1999 -                      | VNIR                      | urban, land-cover, validation   | SpaceImaging (paid)   |
| DigitalGlobe Quickbird                       | <4                | 2000-                       | VNIR                      | urban, land-cover, validation   | DigitalGlobe (paid)   |
| <b>Active Microwave</b>                      |                   |                             |                           |   |   |
| RadarSat                                     | 8-100             | 1995-                       | C band                    |   | commercial distribution   |
| ERS  | 25                | 1991-                       | C band                    | land-cover, sea ice, vegetation structure   | ESA (limited, paid)   |
| JERS   | 18                | 1994-1998                   | L band                    | vegetation structure, sea ice   | JAXA, JRC   |
| <i>PALSAR</i>                                | <i>10-100</i>     | <i>2004 -</i>               | <i>L band</i>             | <i>vegetation structure, sea ice</i>  | <i>JAXA (limited)</i>   |
| Envisat ASAR                                 | 30-75             | 2002 -                      | C band                    | land-cover, sea ice, vegetation structure   | ESA (limited, paid)   |
| <b>Passive Microwave</b>                     |                   |                             |                           |   |   |
| Aqua AMSR-E                                  | 5-50km            | 2001 -                      | Passive microwave         | Soil moisture, sea ice, snow hydrology  | EOSDIS  |
| NOAA AMSU                                    | 50-150km          | 1998-                       | Passive microwave         | Soil moisture, sea ice, snow hydrology  | NOAA SAA  |
| DMSP SSM/I                                   | 15-60km           | 1987-                       | Passive microwave         | Soil moisture, sea ice, snow hydrology  | NOAA SAA  |
| <b>Other</b>                                 |                   |                             |                           |   |   |
| ICESAT GLAS                                  | ~100              | 2003-                       | Lidar                     | vegetation structure, topography  | EOSDIS  |
| Terra MISR                                   | 275               | 2000 -                      | VNIR                      | multi-angle vegetation, clouds, aerosol   | EOSDIS  |
| SRTM   | 30-90             | 2000                        | Radar                     | global topography   | USGS EDC  |
| ENVISAT MERIS                                | 300               | 2002-                       | Hyperspectral             | global vegetation, clouds, aerosol, snowcover   | ESA (limited)   |
| EO-1 Hyperion                                | 30                | 1999-                       | Hyperspectral             | vegetation composition, land-cover  | USGS EDC (paid)   |
| Terra MOPITT                                 | 22km              | 2000-                       | Spectrometer              | CO, CH4 concentrations  | EOSDIS  |
| Terra/Aqua CERES                             | 20km              | 2000-                       | Radiometer                | Albedo, clouds, radiation budget  | EOSDIS  |

Table 6.4 provides a brief description of relevant satellite data streams, together with the typical geophysical parameters derived from them. Note that in some cases data

availability may limit the utility of these data streams. For example, very high-resolution optical imagery (IKONOS, Quickbird, air photos) may also be useful for certain ecological applications (tree counts, crown closure estimates, forest migration studies, etc). However, obtaining these data can be costly, and we anticipate that only a limited amount of high-resolution imagery may be purchased for specific NEESPI projects through individual grants.

It should be noted that long-wavelength Active Radar (L-band SAR) data may be extremely important for mapping wetlands, vegetation biomass, and for land-cover in conditions of low solar elevation or high cloud cover. At present, there is no U.S. or Russian source for L-band SAR data. However, archived JERS-1 data and the upcoming PALSAR mission from JAXA may offer an approach to filling this need. For example, the Siberia-2 project has recently released a mosaic of JERS-1 imagery from the late 1990's across Siberia. Passive microwave datasets can be used to examine snow cover and soil moisture, two critical parameters within the study region. These datasets are widely available through European, Japanese, and U.S. research missions (e.g. NASA Aqua, Envisat, etc).

Coarse-resolution optical remote sensing data (e.g. spatial resolution > 100m) is relatively easy to access from NASA, NOAA, ESA, and RKA. Within the United States, MODIS and MISR reflectance data and geophysical products may be downloaded at no cost from Distributed Active Archive Centers (DAACS) or from sites hosted by academic institutions. Similarly, SPOT Vegetation products may be obtained for no cost for research purposes. Archival AVHRR reflectance and NDVI data are also available from 1981 to the present. Thus, there is a ready supply of coarse-resolution optical data for the last twenty years, with more detailed geophysical products available since 1998.

**Table 6.5. Moderate-Resolution Optical Data Sources**

| Dataset                                 | Description  | Source  |
|---|--|---|
| GeoCover 1975-2000                      | Decadal (1975, 1990, and 2000) global cloud-free coverage from Landsat MSS, TM, and ETM+ | UMD Global Land Cover Facility; USGS EDC, MSU Landsat.org |
| NASA LCLUC Archive                      | 370+ Landsat TM and ETM+ images purchased by NASA PI's                                   | NEESPI Project Office                                     |
| Transparent World Archive               | 530 Landsat-7 images, as well as extensive archives of Landsat-5, IRS, and ASTER         | R&D Center ScanEx , Moscow                                |
| USDA Foreign Agricultural Service (FAS) | 360+ Landsat-5 and Landsat-7 images (since 1989)   | USDA FAS  |
| EO-1 Archive                            | 400+ EO-1 ALI and Hyperion images over Northern Eurasia                                  | USGS EDC  |
| Terra ASTER archive                     | TBD ASTER granules over Northern Eurasia   | EDC DAAC  |

Access to moderate-resolution imagery (Landsat, SPOT XS, ASTER, IRS, MSU-E, etc) is more problematic. Unlike coarse-resolution sensors that always gather data, moderate-resolution sensors are typically tasked to acquire particular images of interest. Thus, depending on acquisition priorities, data gaps may occur. Furthermore, most of these sensors are partly commercial (SPOT, IRS) or charge for data distribution (Landsat). As a result, it can be difficult to assemble comprehensive, interannual, continental datasets for research purposes. For example, analysis of existing image databases indicates a persistent gap in coverage for the late 1980's and early 1990's throughout the Russian Far East for both Landsat and SPOT systems. Nevertheless, large archives of free or inexpensive Landsat-type images do exist (Table 6.5). For example, the Global Land Cover Facility (GLCF) at the University of Maryland distributes, free of charge, the GeoCover datasets produced by the

Earth Satellite Corporation. These datasets include cloud-free imagery for most of Northern Eurasia centered on 1975, 1990, and 2000, providing decadal views of land-cover across the region. ScanEx Corporation in Moscow also hosts archives of Landsat, IRS, and RESURS imagery available at low cost through the TransparentWorld interface.

### 6.2.2. In-Situ Data

Reflecting the long history of ecological and geophysical research in the region, there are numerous archival geophysical datasets that have been collected for routine monitoring or for special studies. In many cases, these data have little distribution or visibility outside of the host country or institution. One of the priorities of NEESPI is to identify these datasets and promote their distribution to address the science goals of the NEESPI project.

**Table 6.6. List of meteorological elements observed at the standard meteorological station within the former USSR.**

| 3-HOURLY DATA SET   | DAILY DATA SET  |
|---|---|
| Air temperature<br>Water vapor pressure<br>Dew-point temperature<br>Relative humidity<br>Sea level pressure<br>Station level pressure<br>Air pressure tendency<br>Visibility<br>Total cloud amount<br>Lower cloud amount<br>Cloud genera<br>Height of cloud base<br>Wind speed<br>Wind direction<br>Precipitation<br>Present weather<br>Past weather<br>Surface skin temperature<br>Ground state<br>Atmospheric phenomena | Mean daily air temperature<br>Maximum air temperature<br>Minimum air temperature<br>Daily precipitation<br>Snow depth<br>Snow coverage<br>Characteristics of site<br>Minimum of relative humidity<br>Minimum of surface temperature<br>Wind speed maximum<br>Atmospheric phenomena<br>Atmospheric phenomena duration<br>Daily total and low cloud amount<br>Sunshine duration |

During the NEESPI implementation, different types of in-situ information will be used: hydrometeorological, socio-economic, land-cover, and land-use data. Each data type has particular attributes with regard to collection, archiving, and pre-processing. For example, the system of hydrometeorological observations in within the former Russian Empire and USSR has been established over a lengthy period of time. At the end of the 19<sup>th</sup> century (1891 is considered as the year when the major types of regular observations were finally established; Vannari 1911), the system had already satisfied major needs in the data on environment in he densely populated parts of the continent. Since that time an enormous volume of observations has been accumulated and, relatively recently, most of it was digitized. Table 6.6 shows major meteorological elements that have been observed at the meteorological network within the former USSR (up to 1936 three times per day, then four times per day, and since 1966, 3-hourly). In addition to standard meteorological elements listed in Table 6.6, special networks observe runoff characteristics, soil moisture, and atmospheric radiation. Frequently, these stations are co-located.

The number of the major operational stations over Russia and the former USSR varied from ~100 in the past decade of the 19<sup>th</sup> century to maximum of ~ 3500 in 1985. At that time, however, precipitation was measured at ~11,000 locations. Presently, at the Russian territory there are ~1900 operational stations. The information from these stations is routinely archived at the State Data Fund in several special “archive” formats. However, this

information is not used for research purposed directly, but serves as a baseline source for specialized sub-arrays to address particular tasks (e.g., such WMO programs like GEWEX-GAME-Siberia or ACSYS). It is expected that funded NEESPI activities will take the form of joint research projects, with formal collaboration of the relevant State Data Fund data managers and scientists involved in organization, maintenance, and quality assurance of pertinent observation programs. Such a partnership will ensure the integrity of data used for investigations.

In addition to data available via the State Fund, there are a lot of data arrays accumulated during various research projects, field expeditions/experiments, etc. Many of these datasets are referenced in Section 3.3. These data are a property of Institutions, private firms, but very infrequently the “Principal Investigators”. Access to these data can be very different. As it can be seen from Table 6.6, the scientific potential of the baseline archive is significant and should be utilized. In addition to the direct data flow from meteorological stations to the State Fund, an operational (“real-time”) data flow exists through the Global Telecommunication System (GTS) mostly for weather forecasting needs. These data also can be used for NEESPI purposes but only as “near-real time data”, i.e. in a quality-controlled form. The “near-real time” data flow is, nevertheless, less reliable (also it is available within a month or so) and should be later be replaced.

Ecological datasets from Northern Eurasia include forest inventories and ecological research networks, documenting vegetation structure, species composition, and biomass. For example, within China, the Chinese Ecological Research Network (CERN) consists of 33 research stations covering major vegetation biomes. CERN stations started collecting data in 1988, and include datasets on soil ecology, photosynthetic productivity, biodiversity, and carbon cycling. Another example, the forest inventory of Russia, has long collected regional data on timber resources and forest types across the country. These inventory datasets have very different rules and methods of collection, archiving, and dissemination that also must be accounted for realistic estimates of their availability for the NEESPI needs. Hydrological datasets of interest include river runoff and soil moisture observations. In some cases the geographic scope of the datasets is limited (i.e., soil moisture was primarily monitored in agricultural regions), or the data have not been widely available outside the host country. In some cases, these station observations have been compiled and gridded to create maps of regional climate, hydrological, and ecological patterns. An example of this is the IIASA Land Resources of Russia dataset, which features numerous maps of socio-economic variables, natural resources and climate, and land use compiled from various primary data sources. In cases where original plot-level field data are not available, NEESPI researchers can utilize these existing gridded data products for model inputs.

### **6.2.3. Known Data Gaps**

Several gaps in data coverage have been identified during the NEESPI planning process.

These include:

- moderate/high-resolution optical remote sensing coverage for northeast Siberia during the 1980’s and early 1990’s.
- meteorological, hydrological, forest and ecological inventory data from Northern Eurasia, which exist but (a) not in digital form; (b) are not available for scientific purposes; and (c) have not been widely distributed ;
- direct measurements of atmospheric CO<sub>2</sub> and water fluxes from a permanent flux-tower network;
- socio-economic datasets that provide coherent views of population, demographics, and land-use patterns and their changes.

Where possible, NEESPI will attempt to fill these gaps through the collection of new datasets, or through improved dissemination of existing datasets.



### ***6.3. Data Policy and Management***

The success of the NEESPI project requires the open exchange of data and information among project participants, to the greatest extent allowable by institutional, national, and international regulations. For addressing issues related to data management and policy, a special working group should be created under the NEESPI Project auspices. The group should coordinate the conditions of dissemination of the information products prepared through NEESPI, select the team responsible for update and creation of derived products, and resolve publication policies for use of those products. The end result of this process should be a NEESPI Data Policy NEESPI, to be included with the overall project Implementation Plan.

The first and crucial step in the information policy will be a creation of a comprehensive metadata archive: information about base datasets, information products, and their derivatives. History of stations and their observation practice should be a subject of another metadata archive that will facilitate the quality assessment of the scientific results and secure the NEESPI scientist from “discoveries” that thereafter will be discarded as artifacts due to the inhomogeneities in the data.

To facilitate the exchange of information, a NEESPI Information System should be implemented. NERIN, the Northern Eurasian Regional Information Network, has been proposed as an umbrella network for the exchange of data and information among NEESPI participants. As a forum for data producers and data users, NERIN is helping to formulate, prioritize, and articulate the requirements for satellite and in situ monitoring of the region. The regional network is currently involved in engaging regional scientists and resource managers in the evaluation and validation of new satellite products. The NEESPI Implementation Plan should finalize recommendations for the full NERIN system. In doing so, various factors need to be considered. First, a great many historical datasets exist throughout the region, and care must be taken to ensure that the dataset producer alone is responsible for changes to the data, releasing official versions, and distributing the appropriate version to users. Also, many datasets are not currently available in electronic format, and some institutions across the region do not have reliable access to the internet. To speed access throughout the region, an information architecture should be considered that relies on “mirrored” sites within the United States, Europe, Japan, and countries of Northern Eurasia. Some consideration should be given to partnering arrangements, so that institutions with poor electronic capabilities can collaborate with larger institutions that have good connectivity. Given the importance of Geographic Information Systems (GIS) in allowing data sources to be integrated to further scientific discovery, some emphasis should be directed toward ensuring that NEESPI electronic datasets are available in GIS-compatible formats.

Long-term archiving of datasets is another priority for the NEESPI project. We anticipate that, following the conclusion of the formal NEESPI activities, new NEESPI datasets will still have significant scientific value for the research and applications communities. Datasets created or substantially modified through NEESPI will be permanently archived at appropriate long-term data centers within the host countries. Provisions for this should be documented within both the NEESPI Implementation Plan and Data Policy.

### ***6.4. Data and Information Technology Priorities***

While overall priorities for the NEESPI Research Strategy are given in Chapter 8, the preceding discussion suggests some particular priorities with respect to Data and Technology:

1. ***Creation of a Metadata Archive:*** The first priority for NEESPI should be to catalog existing datasets within NEESPI countries, and gather the metadata from these datasets into a standard form. These metadata may then be used as the basis for an electronic

information system (see item 2 below). Particularly important is to understand the distribution (temporally, geographically) of in-situ monitoring networks, and how the characteristics of those networks have changed through time.

2. **Implementation of a NERIN Information System:** A significant contribution of NERIN would be to allow Earth Scientists to locate diverse data pertaining to Northern Eurasia, to establish a fruitful collaboration with the specialists who have the first hand experience in the observational programs , and, consequently, to promote the open exchange of data products among researchers. A full NERIN Information System would meet these goals. Such an information system should be distributed, to take advantage of existing institutional data warehouses, and should allow for retrieval of various data types and formats (e.g. in-situ plot and station data, national inventories, satellite data and products, etc.)
3. **Creation of a Northern Eurasia Ecological Network:** Given the difficulty in accessing plot-level inventory data from all NEESPI countries, NEESPI should consider supporting a unified ecological network across the region to make independent measurements of vegetation characteristics, soil conditions, and atmospheric fluxes. This approach could merge systematic ecological measurements (e.g. the Chinese CERN approach, long-term biospheric observational stations in Russia, and the international FLUXNET program), new flux tower locations, and validation of satellite data products (e.g. extending the NASA EOS Core Sites). The International Long-Term Ecological Research (ILTER) network could provide a framework for this activity.
4. **Generation of gridded modeling product suites for Northern Eurasia.** While it may not be possible to distribute all national inventories, it may be possible to coordinate the release of coarsely gridded summaries. These summaries, including data on forests, crops, land-cover, water resources, and socio-economics, would be extremely useful for Earth Science modeling, and could obviate the need for release of plot-level data. A model for this activity is the IIASA Land Resources of Russia dataset. However, the NEESPI study area is larger than Russia, and NEESPI should support the harmonization of inventory datasets from across the region.