Long Island Sound RCA Water Quality Model Preliminary Calibration

New York City Department of Environmental Protection

DEP LIS-HWQMS Project

Contract: BEPA-LIS-HWQMS; PIN: 82619BEPALIS

November 22, 2024

Contents

Tables

Figures

Attachments

PREFACE

Through a cooperative agreement, the New York City Department of Environmental Protection (DEP) and the U.S. Environmental Protection Agency (EPA) have been funding the development of a hydrodynamic and water quality model, and a Graphical User Interface/Decision Support Tool (GUI/DST) for Long Island Sound. The effort is intended to build upon, update, and improve earlier modeling efforts used for water management and Clean Water Act compliance efforts required under the 2000 LIS Total Maximum Daily Load and help guide future watershed management, planning, compliance and assessment activities using recent water quality and environmental data and the best available science. The models will also support development of management strategies at system-wide (New York Bight, New York Harbor, and LIS) and regional (e.g., LIS or New York Harbor) spatial scales.

This report, "Long Island Sound RCA Water Quality Model Preliminary Calibration," is a project deliverable that reports on an initial effort to calibrate the RCA water quality model. Data for calendar years (CY) 2005-2006 was used to perform preliminary calibration because CY2005-2006 is a data rich time-period. The objective of the preliminary calibration report is to present a model calibration approach that can be reviewed, approved, and applied to the full model calibration time-period (i.e., CY2005-2014).

What makes this document worth reading is that it provides information about the status of the LIS water quality model calibration work. This document has been thoroughly reviewed by a peer review group referred to as the Model Evaluation Group (MEG) and staff from the U.S. Geological Survey, EPA, and DEP. Reviewer comments were addressed in a response to comments memo prepared by HDR (Consultant) and reviewer comments will be addressed in the full model calibration report. DEP has accepted responses to reviewer comments even though the reviewer comments are not addressed in this report. For example, one commenter recommended adding additional spatial regions for model performance assessments into the East River and the Battery. In response, HDR indicated that they are developing a model skill assessment procedure for MEG review that will include the recommended additional regions and be presented in the full model calibration report. In other cases, DEP has accepted responses that effectively say "this issue will be addressed in the full model calibration report." For example, one reviewer recommended presenting sediment flux results with paired plots of model results for DO, NH4, NO3/NO2 & PO4 in the bottom layer of the water column to help readers better understand coupled water column-sediment flux results. In response, the consultant agreed to address this in the full model calibration report.

Attachment 6 presents a collection of submitted reviewer comments and Consultant responses to reviewer comments. Moreover, a watermark with the word "Partial" has been added to the pages of this report to reflect the state of this report. This information is presented to provide context for reading this report.

Gregory Wilkerson, P.E., Ph.D.
Director, Watershed Planning and Modeling
Bureau of Environmental Planning and Analysis (BEPA)
NYC Environmental Protection

(O) 718-595-4145 | GWilkerson@dep.nyc.gov

1 Introduction

The New York City Department of Environmental Protection (DEP) and United States Environmental Protection Agency (EPA) are funding the development of a coupled hydrodynamic/water quality model of Long Island Sound (LIS) to replace the existing System-wide Eutrophication Model (SWEM). SWEM used a combination of the Estuarine Coastal and Ocean Model (ECOM) and the Row-Column Advanced Ecological System Modeling Program (RCA). The current project is using the Regional Ocean Modeling System (ROMS) hydrodynamic model coupled with the RCA water quality model for the updated LIS model (HDR, 2021).

The modeling study area includes LIS, tidal Connecticut River, tidal Thames River, Peconic Bay, Gardiners Bay, Block Island Sound, East River, Harlem River, Hudson River, downstream tidal reaches of several NJ tributaries, Newark Bay, NY/NJ Harbor, Jamaica Bay, Raritan Bay, NY Bight, and coastal ocean from Cape May NJ to Nantucket Island MA. The study area and model grid are presented in Figure 1.

The calibration of the water quality component of ROMS-RCA is a multi-step process that includes model preliminary calibration (this report), calibration, validation, and a post-audit. The process began with model testing using October 1994 through September 1995 data or water year 1995 (WY95) inputs from the SWEM model setup (HDR, 2022). The second step is a preliminary calibration of RCA using calendar years 2005 and 2006 (CY05-06), which is the subject of this report. The preliminary calibration time-period was selected as it is a data rich time-period for the ROMS hydrodynamic and RCA water quality model calibration, and also provides the LIS Model Evaluation Group (MEG) an opportunity to review the preliminary calibration during the full model calibration process to the CY05-14 time-period. After completing the model calibration, the coupled ROMS-RCA hydrodynamic/water quality model will be validated for two time-periods (CY03-04 and CY15-18) and undergo post-audit modeling for CY19-22.

It should be noted that the RCA water quality model preliminary calibration does not represent the final model calibration and that model calibration is still in progress. The preliminary model calibration presented does represent a good level of model-data comparison for most parameters both from a visual qualitative perspective and from a quantitative skill assessment perspective. Additional comparisons to observed data for the remaining model calibration years (i.e., CY07-14) is warranted at this time. The remaining calibration years represent a wider range of environmental variability (e.g., river flows, meteorology) for testing of the current set of model coefficients. The goal of calibration is for the model to represent the overall levels of modeled parameters for multiple years at locations throughout the model study area. This report describes the data available for skill assessment, skill assessment metrics, results of the skill assessment, and next steps for further improvement of the model calibration for the full CY05-14 model calibration time-period.

2 RCA

RCA is a three-dimensional generalized water quality modeling computer code developed by HydroQual (now HDR) for application to marine and freshwater systems. The development of RCA has its origins in the mid-1980's but can trace its lineage back to the USEPA-supported water quality model WASP, which was developed in the early 1970's by HydroQual's predecessor firm, Hydroscience, Inc. RCA solves general mass balance equations for water quality variables of interest.

RCA evaluates the fate and transport of conventional and toxic pollutants in surface waterbodies in one, two, or three dimensions. The RCA computer code uses finite difference techniques to simulate the time-varying processes of advection and dispersion, while considering point and diffuse mass loading, boundary exchange, and linear and non-linear losses and production. Information concerning the advective and dispersive transport fields is usually provided to RCA by an offline hydrodynamic model. In the LIS application, RCA has been incorporated into the ROMS code to run simultaneously with the hydrodynamic model. ROMS and RCA were first linked for an application in Chesapeake Bay (Testa et al., 2014).

RCA permits the user to provide site-specific kinetic subroutines to model the contaminants and water quality variables of interest. Kinetic subroutines have been developed, which permit RCA to model coliforms, pathogens, BOD/DO, simple and advanced eutrophication, wetland systems, and toxic contaminants. In the LIS application, the advanced eutrophication kinetics have been applied with the addition of a total suspended solids (TSS) state-variable. Table 1 lists the state-variables applied in the LIS model.

In addition, the advanced eutrophication kinetic subroutine has been constructed to link to a sediment nutrient flux subroutine. This permits the coupling of the water column and sediment bed, to account for the deposition of particulate organic matter, its diagenesis in the sediment bed, and the resulting flux of inorganic nutrients with the overlying water column, and sediment oxygen demand. The sediment nutrient flux subroutine also accounts for the effects of bioturbation on dissolved and particulate mixing in the sediment bed.

Attachment 1 contains a draft RCA User's Manual that presents the water quality and sediment flux models kinetic formulations, description of the various kinetic constants and parameters, and typical input ranges for the constants.

Table 1. State-Variables in LIS RCA Model

State-Variable	Abbreviation
Phytoplankton Group 1 (winter)	PHYT1
Phytoplankton Group 2 (summer)	PHYT2
Phytoplankton Group 3 (spring)	PHYT3
Refractory Particulate Organic Phosphorus	RPOP
Labile Particulate Organic Phosphorus	LPOP

State-Variable	Abbreviation
Refractory Dissolved Organic Phosphorus	RDOP
Labile Dissolved Organic Phosphorus	LDOP
Total Phosphate	PO4T
Refractory Particulate Organic Nitrogen	RPON
Labile Particulate Organic Nitrogen	LPON
Refractory Dissolved Organic Nitrogen	RDON
Labile Dissolved Organic Nitrogen	LDON
Total Ammonium Nitrogen	NH4T
Nitrite + Nitrate Nitrogen	NO23
Biogenic Silica	BSI
Available Dissolved Silica	DSI
Refractory Particulate Organic Carbon	RPOC
Labile Particulate Organic Carbon	LPOC
Reactive Particulate Organic Carbon	RePOC
Refractory Dissolved Organic Carbon	RDOC
Labile Particulate Organic Carbon	LDOC
Reactive Particulate Organic Carbon	ReDOC
Algal Exudate	EXDOC
Aqueous SOD	O2EQ
Dissolved Oxygen	DO
Total Suspended Solids	TSS

3 Model Calibration Data

3.1 Grab Sampling

Grab sampling in this context are samples collected in the field and sent to a laboratory. Grab sampling data are generally the primary type of data used for eutrophication model calibration due to the number constituents and processes the model is attempting to reproduce. Continuous sampling is generally not practical or possible for most of the constituents that are required. Grab sampling data are used to assess how well the model reproduces these constituents. It is important that the model reproduces these constituents because they quantify the growth of algae, water clarity, and dissolved oxygen concentrations. The following data sources provide grab sample data.

- The DEP Harbor Survey (HS) data set (Harbor Water Quality | NYC Open Data (cityofnewyork.us)) includes dissolved oxygen (DO), nitrite+nitrate (NO23), ammonium (NH4), orthophosphate (PO4), total Kjeldahl nitrogen (TKN), dissolved silica (DSI), total phosphorous (TP), total suspended solids, chlorophyll-a (chl-a), total organic carbon (TOC), dissolved organic carbon (DOC), and 5-day biochemical oxygen demand (BOD5). With the exception of DO, which are collected at the surface and bottom, the samples are collected at the surface. Samples are collected on a monthly basis from November through April and on a weekly basis from May through October. The data were used for model calibration in western LIS, East River, Harlem River, Hudson River, Kill Van Kull, Jamaica Bay, and NY/NJ Harbor (see Figure 2). The stations are divided into primary stations, secondary stations, and stations not considered for model calibration (generally located in small tributaries). The preliminary calibration focused on the primary and secondary stations with skill assessment statistics calculated at the primary stations.
- The Connecticut Department of Energy and Environmental Protection (CTDEEP) collects samples for biogenic silica (BSI), chl-a, PO4, NH4, NO23, particulate carbon (PC), particulate nitrogen (PN), particulate phosphorus (PP), DSI, total dissolved nitrogen (TDN), total dissolved phosphorus (TDP), TSS, and 30-day BOD (BOD30) (Water Monitoring Data Availability (ct.gov)). Surface and bottom data are collected and these data were used for model calibration. CTDEEP has monitoring locations throughout LIS (see Figure 3). The stations where grab sampling occurs are year-round stations and were considered primary stations for model calibration.
- The New Jersey Harbor Dischargers Group (NJHDG) dataset includes DO, TSS, TKN, carbonaceous BOD5 (CBOD5), NH4, nitrite (NO2), Nitrate (NO3), total phosphorus (TP), PO4, chl-a and DOC in the Hudson River and northern NJ water bodies. Only stations in the Hudson River and Raritan Bay are considered primary stations (see Figure 4) and have not yet been used for model calibration.
- The Interstate Environmental Commission (IEC) has seasonal monitoring stations (see Figure 5) in western LIS (Western Long Island Sound Monitoring Program | Interstate Environmental Commission (iec-nynjct.org)). The majority of the samples are DO and chl-a collected weekly during the summer. In 2014, IEC began bi-weekly surface sampling for BOD5, TSS, NH4, NO23, PN, PO4, TDN, TDP, PP, DOC, PC, DSI and BSI. They are all considered secondary stations and have not yet been used for model calibration.

3.2 *In Situ* Data and Vertical Profiles

In situ data are collected in the field and require the use of probes, meters, or other sampling gear (e.g., Secchi disk). The use of probes allows measurements at multiple depths within the water column, not just the surface and bottom. *In situ* data were also used for model calibration. *In situ* data are available from the following sources and were used for model calibration.

- DEP Harbor Survey collects DO, Secchi depth, and occasionally photosynthetically active radiation (PAR).

- CTDEEP collects vertical profiles of DO, chl-a, and PAR at their summer stations (Figure 3).

3.3 Continuous Monitoring

The Long Island Sound Integrated Coastal Observing System (LISICOS) program (UConn's Long Island Sound Observatory) has monitoring buoys that collect continuous DO measurements that were used for model calibration (see Figure 6). During 2006, data were available from March through December at Execution Rocks, FB01, FB02, FB03, and the western LIS stations. These data were provided by Dr. James O'Donnell from the University of Connecticut.

There are significant portions of this 2006 DO data set where the data do not vary as expected either due to DO probe malfunctions or other causes. At this point, the 2006 DO data have not been used for the preliminary calibration but are used for qualitative comparison purposes to guide model calibration.

3.4 Algal Primary Production

Two data sets have been found for algal primary production. Goebel et al. (2006) measured primary production and community respiration at seven CTDEEP stations (A4 through I2) during 2002 and 2003. Collins et al. (2013) conducted extensive measurements at CTDEEP station I2 during 2010. Although these data are from different years than the CY05-06 time-period, they were used for qualitative comparisons to evaluate how the model output compares to the available data.

3.5 Sediment

Balcom et al. (2007) deployed benthic chambers to measure SOD at two locations during 2005 and three locations during 2006. These were the only sediment data found for the preliminary calibration period. Mazur et al. (2021) collected data in 2016 and 2017 at five locations in LIS at LISCOS stations (Figure 6). Measurements included solid phase organic carbon and organic nitrogen as well as ammonium, nitrate, denitrification, phosphate, and sediment oxygen demand fluxes. Model results were compared to these data to determine if the model is reproducing the relative magnitude of the data. Water quality improved between 2005 and 2016-2017, so the model may over-estimate the measured fluxes. However, the sediment often takes decades to change substantially, so the measurements are informative. HDR is aware that other sediment data sets exist and will use the data when they become available.

4 Model Inputs

HDR developed a ROMS Hydrodynamic Model Inputs and RCA Water Quality Model Load Development Approach memorandum (HDR, 2023). This document is a living document and was updated for the RCA preliminary calibration. This document should be referenced for the approach used to develop the RCA model loading inputs. Figure 7 provides a loading summary used in the current model preliminary calibration to provide context to the model behavior. Additionally, the model constants used for this preliminary calibration are presented in Attachment 2 (water quality) and Attachment 3 (sediment).

Beyond loading and constants, RCA requires additional inputs that affect the behavior of the model. A few of the important model inputs are described here to provide further insight into the model and the approach taken for model calibration.

4.1 Light Extinction

The original RCA model used a light extinction formulation that assigns a base light extinction coefficient that is added to an additional term based on calculated chl-a levels. The base light extinction coefficient accounts for factors aside from chl-a that limit light penetration into the water column including, TSS, color, and water itself. With the addition of the TSS state-variable to the LIS RCA model a revised formulation was developed. In Chesapeake Bay, Xu et al. (2005) developed an empirical function using TSS, chl-a, and salinity. Salinity was used as a surrogate for colored dissolved organic matter (CDOM) since CDOM data was not available. The idea being that CDOM is associated with freshwater so lower salinity would correspond to higher CDOM.

A similar approach was used to develop a light extinction formulation in LIS. A multiple linear regression was performed using light extinction coefficients calculated from vertical photosynthetic active radiation (PAR), TSS, chl-a, and salinity data for LIS. The relationship was not strong, so the formulation was modified to include a spatially variable base light extinction coefficient, which improved the relationship.

The resulting equation is:

$$K_e = K_{ebase} + 0.017 \times Chla + 0.009 \times TSS - 0.015 \times Sal$$

where K_e is the total light extinction coefficient and K_{ebase} is the spatially variable base light existing coefficient. A minimum K_e of 0.3/m was also used together with this equation. This formulation has been applied in the LIS RCA model.

The base K_e (K_{ebase}) was developed in concert with the equation above to better fit available data. Figure 8 presents the currently applied spatially variable K_{ebase}.

4.2 Net Settling (VSNET)

RCA does not include sediment resuspension. As a way to parameterize the process of resuspension, the model uses a constant named Vsnet that has a value between 0.0 and 1.0. A value less than 1.0 indicates that only a fraction of the organic material that settles to the sediment is incorporated into the sediment. The fraction that does not get incorporated remains at the bottom of the water column and can be advected to other model segments or upward vertically. In areas known to be muddy, Vsnet is set at 1.0. In sandy or gravelly areas, the value can be set to less than 1.0. SWEM had little variation of Vsnet in LIS with most areas set to 1.0.

Poppe et al. (2000) developed a map of sediment types in LIS that guided the assignment of Vsnet. Eastern LIS is comprised of mostly sandy to gravelly sediments indicating that limited organic material is incorporated into the sediment. Much of this area can be set to a Vsnet less than 1.0. This may help push some organic matter from eastern LIS to the central and western LIS to reduce the DO concentrations in these areas. The western LIS is siltier, so it would be expected that the majority of the organic material that reaches the bottom is incorporated into the sediment. Figure 9 presents the currently applied Vsnet.

4.3 Reaeration

The model is currently applying a reaeration oxygen transfer coefficient based on wind speed from the North American Regional Reanalysis model inputs and the following equation.

$$K_L = 0.728 \times WIND^{0.5} - 0.317 \times WIND + 0.0372 \times WIND^2$$

where WIND is wind speed in m/s and K_L is the oxygen transfer coefficient (m/d).

4.4 Phytoplankton Predation

CTDEEP zooplankton abundance data were reviewed for the period of 2003-2018. Three copepod species (meso-zooplankton) tended to dominate the abundance: *Temora longicornis* at cooler temperatures (<10°C), *Acartia hudsonica* at mid-range temperatures (5-20°C), and *Acartia tonsa* at higher temperatures (>20°C). Some literature (Calbet, 2008, Lopez et al., 2013) suggests that it is the micro-zooplankton and not the larger meso-zooplankton that are more important phytoplankton grazers with the exception of *T. longicornis* in the spring. The CTDEEP data for micro-zooplankton is more limited and does not show a strong seasonal pattern except for the western sound where biomass is higher during the summer. In general, there is a decreasing trend in zooplankton abundance from west to east with seasonal peaks in winter/spring and summer.

The approach for assigning phytoplankton predation in the original RCA model was to apply a base zooplankton grazing rate (at 20°C) that is adjusted by temperature. The adjustment is a constant raised by the power of temperature minus 20°C such that the grazing rate is lower than the base rate below 20°C and greater above 20°C. In SWEM, this approach was replaced by assigned monthly rates that varied according to zooplankton abundance data. This step function of rates had higher rates during the late winter and spring and lower rates during the summer, which is essentially the opposite of the original RCA approach.

While the SWEM approach is reasonable for a model calibration of one year, it does not account for annual variability, nor does it account for potential changes to phytoplankton abundance that may occur due to nitrogen controls.

A new phytoplankton predation formulation was employed in the LIS RCA model based on an approach applied in the Chesapeake Bay Model (Cerco and Noel, 2004). This predation formulation includes the activity of zooplankton, filter-feeding benthos, and other pelagic filter-feeders including planktivorous fish. This new formulation accounts for additional predation beyond zooplankton predation as used in the original RCA model, and as applied in the LIS RCA model accounts for predation by all potential predators. Since zooplankton are not modeled explicitly in the RCA model, this new formulation was chosen to account for total phytoplankton predation.

The new predation formulation is represented with a clearance/filtration rate F (m^3/g predator C/day), predator biomass M (gC/m^3), and the phytoplankton biomass B (gC/m^3).

The new predation equation is $PR = F \times B \times M$. Since specification of predator biomass both spatially and temporally is not possible, the predator biomass is estimated as a fraction of the phytoplankton biomass ($M = \alpha \times B$). The equation below presents the new phytoplankton predation formulation. Since α and *F* are not readily known, an empirical constant δ is used in the formulation and is determined via model calibration. A δ value of 0.1 is currently being used in the RCA model calibration.

$$PR = \alpha \times F \times B^2 \quad or \quad \delta \times B^2$$

where: *PR* – phytoplankton predation (g biomass C/m^3/day)

F – clearance/filtration rate (m^3/g predator C/day)

B – phytoplankton biomass (g biomass C/m^3)

δ – product of $\alpha \times F$ (m^3/g biomass C/day)

This new quadratic phytoplankton predation formulation allows predation to still be a function of phytoplankton biomass but with faster cycling of organic nitrogen, phosphorus and carbon due to predation.

4.5 Algal Respiration

The model is currently using an algal respiration approach with a basal rate plus an additional rate associated with algal production as show in the following equation.

$$k_{pr}(T) = r_g \times G_P + r_b \times \theta_{PR}^{T-20}$$

Where k_{pr} is the total respiration rate, r_g is the respiration rate associated with growth, G_p is phytoplankton production rate, r_b is the basal respiration rate and Θ_{PR} is a temperature coefficient. A model input value of 0.05/day is used for r_b, a value of 0.28 is used for r_g, and a value of 1.047 if used for the temperature coefficient.

4.6 Algal Stoichiometry

Particulate carbon (PC), particulate nitrogen (PN), particulate phosphorus (PP), and biogenic silica (BSI) data along with chl-a data were analyzed to estimate algal stoichiometry for use in the RCA model. The data used for this analysis were from CTDEEP at the year-round monitoring locations. Cross-plots of PC to PN, PC to PP, PC to BSI, and PC to chl-a were developed, and linear regressions fit through the data. The slope of the regressions represents the algal stoichiometry and were analyzed by season and annually. Results from these analyses varied by station but provided guidance in setting the algal stoichiometry inputs in the model during calibration. The model inputs used are presented in Table 2 for conditions when excess nutrients are available (i.e., not limiting algal growth).

A variable algal nutrient stoichiometry formulation was also used in the RCA model that increases the nutrient ratios as the nutrient concentrations become limiting to algal growth. This formulation represents the process known as variable stoichiometry, which allows phytoplankton to adjust their stoichiometry when nutrients become limiting to phytoplankton growth. Additional details of the variable algal nutrient stoichiometry are provided in Attachment 1 (Draft LIS RCA Model User’s Manual).

Table 2. LIS RCA Model Algal Nutrient Stoichiometry

Parameter	Algal Group 1 Value (Winter)	Algal Group 2 Value (Summer)	Algal Group 3 Value (Spring)
Carbon to Nitrogen Ratio (C/N)	6.5	6.5	6.5
Carbon to Phosphorus Ratio (C/P)	50	50	50
Carbon to Silica Ratio (C/Si)	4	8	6

The RCA model also requires the assignment of algal carbon to chl-a (C/Chla) ratios to convert from the model calculated phytoplankton carbon to chl-a. A new variable C/Chla ratio formulation was employed in the LIS RCA model based on an approach applied in the Chesapeake Bay Model (Cerco and Noel, 2004). This approach varies the C/Chla ratio as a function of the light extinction coefficient with higher C/Chla ratios occurring with less light extinction with depth (i.e., clearer water). Analysis of the CTDEEP data showed C/Chla ratios ranging from 16-80 based on the seasonal cross-plot analyses (40-50 on an annual basis) and up to 200-300 on a point-to-point calculation. In addition, the C/Chla ratio showed an increasing ratio with lower light extinction coefficients. The variable C/Chla ratio equation implemented in the RCA model is presented below.

$$C/Chla = a + b \times exp(-c \times K_e)$$

where: a = 30

b = 130 (winter and spring algal groups), 200 (summer algal group)

c = 1.2

This variable C/Chla ratio approach allowed better representing of the west to east chl-a variability observed in LIS.

5 Model Calibration and Skill Assessment

5.1 Water Quality Time-Series

Model versus grab sampling time-series data figures are presented to help visually compare the model’s ability to produce the magnitude and timing of changes in water quality concentration over an annual cycle. These figures provide a qualitative assessment of the preliminary calibration. Nine representative CTDEEP and nine representative DEP spatially distributed stations have been chosen to demonstrate the model’s ability to reproduce conditions during the CY05-06 preliminary calibration time-period. Model output was saved hourly but is presented as a 24-hour moving average to make the figures easier to review. The moving average reduces the additional output variability captured by the model. The review of this subset of stations allows for a more concise description of the preliminary calibration. Additional figures for all of the stations for CY05-06 can be found in Attachment 4.

5.1.1 Salinity and Temperature

The salinity and temperature preliminary calibration was documented in the hydrodynamic model preliminary calibration memorandum (HDR, 2023). However, HDR has implemented suggestions made by the MEG since that report was issued. The changes included the following revisions.

- Use of the Northeast Coastal Ocean Forecast System (NECOFS) model temperature and salinity output for assigning ocean boundary conditions in the ROMS hydrodynamic model.
- Assigning spatially variable bottom roughness lengths (ZoB) of 0.002 meters throughout the model grid with adjustments to 0.05 meters in the East and Harlem Rivers, and 0.05 meters on the eastern end of LIS where sand and gravel bottom types are present.
- Use of a minimum model water depth of 2.5 meters referenced to the North American Vertical Datum of 1988 (NAVD88), which is close to mean sea level (MSL) in the model study area.
- Various revisions to freshwater river and point source inputs.
- Use of Jerlov Water Type III light extinction coefficients to improve model calibration to bottom water temperatures.

Since changes in temperature affect rates and changes in salinity affect density stratification, a brief update of the current state of the temperature and salinity calibration is included.

Figures 10 through 17 present the model versus temperature and salinity comparisons at the nine CTDEEP and nine DEP stations for CY05 and CY06. The application of NECOFS derived temperature and salinity model boundary conditions resulted in additional temperature and salinity vertical stratification in LIS in addition to improvements in the overall model temperature calibration. Surface and bottom salinity in LIS is over-estimated by about 2 practical salinity units (psu) at most locations during spring and summer months. Model sensitivity runs have indicated that the assigned model salinity boundary conditions are the primary factor controlling salinity levels in LIS. Further model boundary condition revisions will be investigated as the model is calibrated to the full time-period. The model does calculate differences in temperature and salinity between CY05 and CY06, which in turn drives differences in water quality between these two years.

5.1.2 Phytoplankton

5.1.2.1 Chlorophyll-a

Chl-a concentrations are an indicator of algal biomass. Chl-a grab sample data are available from the CTDEEP monitoring program and DEP Harbor Survey Program. Figures 18 through 21 present examples of model-data time-series comparisons for the CTDEEP and DEP data from CY05 and CY06. These figures show nine representative CTDEEP stations and nine representative DEP stations for the preliminary calibration periods.

Figure 18 shows the chl-a calibration for CTDEEP locations in 2005. More westerly stations are shown on the left side of the page and more easterly stations are shown on the right side of the page. The data show higher chl-a concentrations in the western LIS with the higher chl-a concentrations measured in the late-winter/early spring and during the summer. The model captures many of the features of the data but needs additional fine tuning to better reproduce the observed growth patterns. Chl-a concentrations are low at the beginning of the year, and the model generally reproduces the data. The spring bloom is more pronounced at station A4 where the model under-estimates the chl-a data. At the more easterly stations, the model over-estimates the low chl-a data and the model does approximate the lower late spring chl-a data that were measured in the April/May time-period. The timing of the summer bloom is fairly accurate. In general, the model under-estimates the higher chl-a data and over-estimates the lower chl-a data, and additional model calibration to the full time-period will be beneficial to further improving the spatial and temporal change observed in the chl-a data. Since the model calculates phytoplankton carbon and chl-a is calculated via a carbon to chl-a ratio, some of the chl-a differences between the model and data can ultimately be addressed with changes to the carbon to chl-a ratios of the individual algal groups.

The 2006 calibration for chl-a at CTDEEP stations is shown in Figure 19. Based on the chl-a data, a winter/spring bloom is not as evident as observed in 2005. High chl-a was measured during the summer, but high chl-a concentrations are mostly evident at station A4, the western most station. Temperature data indicate that 2006 had a warmer winter and cooler summer than 2005. The model under-estimates the higher chl-a data at station A4 during 2006 but compares more favorable at the other stations.

The 2005 calibration for chl-a at the DEP stations is shown in Figure 20. Some missing data at the beginning of the year make it difficult to assess whether there was a system wide spring bloom; however, data at station E10 (closest to CTDEEP station A4) suggest a winter/spring bloom did occur at some locations. Most stations show a summer bloom, but stations E4 and E2, do not show high chl-a during the summer despite adequate nutrient levels to fuel a bloom. The model spring bloom is difficult to assess because monitoring data are not available for this period and the summer bloom model results are mixed. At stations N1, N4, K5A, N8, and E10, the summer bloom is under-estimated by the model. At the remaining stations, the model calculated summer chl-a compare more favorably to the data.

The 2006 model data comparison for chl-a at DEP stations is presented in Figure 21. The 2006 data show a similar pattern to the 2005 chl-a concentrations, but a small winter/spring bloom is more evident in the February 2006 data, and the summer bloom does not have chl-a concentrations as high as observed in 2005. In general, the model compares more favorably to the 2006 chl-a data than the 2005 data.

5.1.2.2 Particulate Carbon and Nitrogen

Particulate carbon (PC) and nitrogen (PN) are other indicators of algal biomass. Only CTDEEP collects particulate carbon and nitrogen data and the assumption is that PC and PN are primarily organic.

Figure 22 presents the 2005 CTDEEP PC model and data comparison. There is some evidence of a spring bloom in the data, primarily at station A4 with smaller peaks at other stations. The summer blooms are more evident with higher surface PC concentrations occurring in the summer. Higher PC is measured in the western LIS as compared to in eastern LIS and the data indicate the phytoplankton have higher carbon to chl-a ratios in the summer than the winter. The model under-estimates the spring bloom at the western station A4 but tends to over-estimate the PC concentrations at the eastern stations. As the calibration continues, modeling will assess what factors limit growth in the eastern LIS that are not limiting to growth in the western LIS during the spring. Overall, the model generally reproduces the summer PC.

The 2005 PN model versus CTDEEP data comparison, in Figure 23, shows a similar pattern of increased summer concentrations as the PC data. The model compares more favorably to the western LIS spring PN data than the PC data but over-estimates the spring PN in eastern LIS. The summer PN model results compare favorably to the data, but the bottom PN is over-estimated.

The 2006 PC and PN model versus data comparisons are presented in Figures 24 and 25. A winter/spring bloom is less evident in the 2006 data, partially because there is less data in 2006. Both PC and PN show the same pattern with peak concentrations occurring in the June/July timeframe. The model comparison to the data is fairly good.

5.1.3 Nutrients

Inorganic nutrients fuel algal growth and their concentrations are influenced by the uptake and recycle from phytoplankton and other biogeochemical processes.

5.1.3.3 Nitrogen

Nitrogen tends to be the limiting nutrient for phytoplankton growth in LIS and the NY/NJ Harbor area. Therefore, it is the nutrient that has been the target for management decisions. Figures 26 through 29 present the TN model versus data comparisons for CTDEEP and DEP stations in 2005 and 2006.

At station A4 in 2005, TN concentrations are generally between 0.4 and 1.0 mgN/L, and at station C2, TN concentrations are generally below 0.5 mgN/L. East of station C2, there is a gradual decline in TN concentrations from west to east. The importance of the starting model initial conditions for nitrogen can be observed in the preliminary calibration results as there is a decreasing trend in the model output from the beginning of the year to the end. As the TN calibration continues with the full time-period, the initial conditions for CY05 will be further evaluated.

The 2006 distribution of TN concentrations is similar to 2005, but there is not much of a difference between peak TN concentrations from stations C2 to M3. The model over-estimates the TN in most places, in part, due to the initial conditions used for the preliminary model calibration.

TN concentrations at the DEP stations are higher than in LIS. East River TN concentrations range from approximately 0.5 mgN/L to greater than 3.0 mgN /L. Note the scale change between the CTDEEP and DEP model-data figures. TN concentrations in the Hudson River and Raritan Bay are generally greater than 1.0 mgN/L. The TN concentrations decrease closer to the NY Bight at station N16. In most cases, the model compares favorably to the data. There are occasional high TN concentrations that the model under-estimates and some concentrations at station N8 near the Verrazano Bridge are under-estimated.

TN concentrations in 2006 appear to be more temporally variable than during 2005 and the model reproduces the 2006 TN data. Since the model reproduces the DEP station E10 data but over-estimates the CTDEEP station A4 data, either a loss mechanism (e.g., nitrification, algal uptake) or the longitudinal mixing between the East River and western LIS may not be captured well in the preliminary calibration and will be further evaluated as the model calibration continues.

DIN is used by phytoplankton for growth and it is the sum of NH4 and NO23. The RCA model uses a nitrogen Michaelis half-saturation concentration of 0.01 mgN/L. At this concentration, the maximum phytoplankton growth rate is reduced by half due to nitrogen limitation. The 2005 and 2006 CTDEEP DIN data (Figures 30 and 31) show the potential for nitrogen limitation at the surface during periods from April through October, with the exception at station A4. The DIN data show a seasonal pattern with higher concentrations at the beginning and end of the year with a reduction during the spring and summer and a recovery in the fall. The model generally reproduces the surface data, but there are occasions during the summer when the model under-estimates the data. The model is currently over-estimating the bottom DIN during the summer. This appears primarily due to the model over-estimating the sediment NH4 flux, as will be discussed in Section 5.2. This will be addressed as the calibration of the full time-period continues. There are also some issues with the initial conditions, which will also be addressed as the model calibration continues.

At the DEP stations (Figures 32 and 33), the DIN concentrations do not indicate a nitrogen limitation with the exception of station N16 near Breezy Point. The model comparisons to the DIN at the DEP stations are good in both 2005 and 2006. The temporal pattern of the model at stations N16 and E10 could be improved in 2005 but otherwise the model matches the data well.

5.1.3.4 Phosphorus

Phosphorus, while needed for phytoplankton growth, is less critical in LIS because it currently is in excess and does not limit phytoplankton growth. TP also has a seasonal signal with a reduction in the spring, and increase through the late summer, and a decrease in the fall to winter (Figures 34 through 37). There is a concentration gradient from higher concentrations in the west of LIS to lower concentrations in the east. The general shape and magnitude of the model TP concentrations matches the data fairly well with the 2005 data, but the model over-estimates the TP concentrations. The model calibration to the 2006 TP data is similar to the 2005 calibration.

The DEP data show a similar temporal pattern from station to station and with the CTDEEP data. Note the scale change on the y-axis such that the maximum scale is twice that for the CTDEEP data. TP concentrations were generally above 0.1 mgP/L, except for station N16. The model reproduces some of the 2005 TP data but over-estimates the TP during the warmer months at stations N1, E4, N4, E2, and E7. This appears to be due to an over-estimate of the sediment PO4 flux.

2006 TP concentrations were generally lower than the 2005 concentrations at the DEP stations. The model generally reproduces the timing and magnitude of the 2006 TP data.

The model calculates lower sediment PO4 fluxes leading to a better comparison to the data.

The DIP or PO4 is the fraction of phosphorus used for phytoplankton growth. The majority of the TP at CTDEEP and DEP stations is DIP, so the temporal and spatial patterns observed in the TP data are observed in the DIP data. DIP concentrations were higher in 2005 than 2006 in LIS. The RCA model has been assigned a Michaelis half-saturation concentration of 0.001 mgP/L for DIP and the data do not approach these concentrations in LIS. Figures 38 through 41 present the model versus data comparisons for DIP. At the CTDEEP stations during 2005, the uptake of DIP during the spring is over-estimated by the model resulting in low DIP model output in the spring compared to the data. As the DO decreases and temperature increases during the summer, the sediment PO4 flux increases, and the model begins to reproduce the DIP data. In 2006, the model does not calculate as much algal uptake in the spring, so the model matches the CTDEEP data more favorably during this period. The model still exceeds the DIP data in the late summer.

As with TP, the DEP DIP data were higher than CTDEEP LIS data and 2005 peaks were higher than during 2006 (Figures 40 and 41). The model calibration to DIP is similar to the TP data with the calibration to the 2006 data being more favorable than 2005.

5.1.3.5 Silica

Silica is the third nutrient used by phytoplankton for growth, but it is only used by diatoms. Diatoms are the dominant phytoplankton group in LIS year-round but are less dominant during the warmer summer months. Figures 42 through 45 show model versus data comparisons for DSI. At the CTDEEP stations in 2005, a decline in DSI concentrations is observed in the late winter/spring followed by an increase during the summer. The easternmost stations show a short duration decrease in the late summer suggesting a small diatom bloom during that period. The model is assigned a Michaelis half saturation concentration of 0.02 mgSi/L. During 2005, there is no evidence of nutrient limitation due to silica. The model calibration to the data is generally favorable, but there are periods during the spring when the concentrations, especially at the surface, are under-estimated. The model does capture the higher surface DSI than the bottom DSI in eastern LIS, which may be due to the influence of the Connecticut River.

The 2006 CTDEEP DSI data are a little different than 2005. More DSI uptake was observed in the spring of 2006 and a late summer/early fall decrease in DSI is observed in most of LIS. DSI concentrations approached nutrient limiting conditions in 2006. The model does not reproduce the earlier silica uptake in 2006, and the general magnitude of the model results reproduce the data for most of the year.

The 2005 and 2006 model versus data comparison for the DEP data are presented in Figures 44 and 45. Some data are missing during the spring, so it is difficult to determine if a spring diatom bloom occurred. Stations closer to the LIS (E10 and E7) show more evidence of decreased DSI concentrations in the late spring than at other stations. DSI concentrations at the DEP stations were higher than in LIS and did not approach nutrient limiting conditions. The model over-estimates the DSI data at stations N1 and N4 by quite a bit. This is most likely due to the assigned DSI Loads in the Hudson River that were based on limited data. As the model calibration continues, the Hudson River loads will be re-evaluated.

5.1.4 DOC

The oxidation of DOC represents a loss of dissolved oxygen and is an important component in the DO balance. Figures 46 through 49 present model versus data comparisons for 2005 and 2006 at CTDEEP and DEP stations. The 2005 and 2006 CTDEEP DOC data are considered suspect because the concentrations are higher than observed in years both before and after this time-period. In most other years the DOC concentrations tend to range from 2 to 3 mgC/L. However, the data may still be able to provide some insight as to how the model is performing. It is expected that the majority of the DOC data are relatively refractory, and the data do not show much spatial variability as is observed in most other constituents.

The model DOC concentrations are a similar magnitude as the DOC data in eastern LIS and the model also reproduces the apparent peaks in the spring and summer. Further west, the model under-estimates the suspect DOC data. The model DOC output compares favorably to the DEP data in 2005 and 2006.

5.1.5 TOC

TOC was calculated as the sum of the CTDEEP DOC data and the PC data. DEP does not collect PC data and TOC is not presented for the HS stations. The majority of the TOC data is DOC. Figures 50 and 51 present the model comparison to the TOC data. Since the model generally reproduces the DOC data, the model generally reproduces the TOC data, especially in western LIS. The model under-estimates the TOC data in eastern LIS in the second half of 2005.

5.1.6 BOD

BOD5 and BOD30 measurements are available at a limited number of CTDEEP stations. Data are shown in Figures 52 and 53. Spatially and temporally, the BOD data somewhat follow the pattern of phytoplankton growth with higher concentrations in the west and peaks occurring during the spring and summer.

The model estimated BOD5 concentrations are based on calculated organic carbon and NH4 concentrations and the associated carbon oxidation and nitrification rates. At the CTDEEP stations in 2005 and 2006, the model generally reproduces the observed data but is biased high although most measurements are less than 2-3 mg/L.

5.1.7 Dissolved Oxygen

Dissolved oxygen is required by higher trophic aquatic organisms and is one of the few constituents in LIS that has numerical water quality criteria. The CTDEEP and DEP grab sampling data were compared against the model results.

Figures 54 and 55 present the 2005 and 2006 model versus data figures for the CTDEEP stations. The data show a temporal trend associated with seasonal oxygen saturation concentrations with higher DO concentrations in the winter and lower concentrations in the summer. The western most stations show oxygen deficits below saturation with the lowest concentrations in the bottom waters at station A4. Minimum summer DO concentrations increase in the eastward direction and by station J2 the DO concentrations are above 5.0 mg/L. The 2005 and 2006 data show similar patterns. The model reproduces the temporal and spatial patterns observed in the 2005 data but does over-estimate some of the lower DO data at station A4. The surface data are generally reproduced well by the model but the model tends to under-estimate the data near the end of the year. The model also compares favorably to the 2006 data with the model over-estimating the low DO data at station A4 and under-estimating DO near the end of the year, similar to 2005.

The DEP stations, shown in Figures 56 and 57, show spatial variability in the DO data from the East River and into the Hudson River. The "N" stations in the Hudson River have minimum DO concentrations near or slightly below 5.0 mg/L, Raritan Bay stations have periods where the DO approaches 3.0 mg/L, and the East River "E" stations generally have the lowest DO concentrations, especially at station E10, which is closest to CTDEEP station A4. The model reproduces both the 2005 and 2006 DEP data reasonably well at most stations. The bottom DO at station N1 is under-estimated during 2005 and the bottom DO is over-estimated at stations K5A and E10. In general, the overall model DO calibration is quite good.

5.2 Sediment

Model results were compared to sediment data collected by Mazur (2020) in 2016 and 2017 and some LISICOS sediment oxygen demand from Balcom et al. (2007) for 2005/2006. Sediment fluxes change from year to year based on organic matter deposition, temperature, and overlying water concentrations. Sediment porewater and solid phase concentrations tend to change more slowly due to the relative magnitude of the model burial rate (<1 cm/yr) to the depth of the active sediment layer (~10 cm). While the water column can reach a new equilibrium in less than a year due to changes in loadings, the sediment with a sedimentation rate of 0.25 cm/yr would take 40 years to reach a new equilibrium due to changes in sedimentation. Due to the relative rate of change in organic matter deposition, temperature, and overlying water concentrations in LIS, it is possible to compare sediment model results to data collected more than a decade later to provide some perspective on the magnitude and timing of the fluxes and sediment concentrations being calculated by the model with the understanding that there is year-to-year variability and potential changes over time due to varying algal levels.

Figure 58 presents the 2005 model output versus the sediment POC concentrations. Station locations can be found in Figure 3. The model includes three classes of reactivity. The most reactive is G1 and the least reactive is G3. Most of the organic matter that settles to the sediment is G1, but because G3 is relatively inert, the highest percentage of the sediment is in the G3 reactivity class. This means the data should be compared to the model output representing G3 sediment POC. The data indicate that in the western sound the sediment is approximately 2-4% carbon and at the central LIS station it is approximately 0.5% carbon. These concentrations are in agreement with or slightly higher than the concentrations compiled by Poppe et al. (2000). The model under-estimates the POC concentrations except in central LIS (CLIS station). The 2006 model versus data comparisons for POC are shown in Figure 59 and the model also compares favorably to the data. The model results for the G2 and G3 carbon for 2005 and 2006 are essentially the same due to the slow reactivity of these two sediment POC classes. The G1 carbon varies slightly as it is dependent on the rate at which POC settles to the sediment and water column POC concentrations.

The 2005 sediment PON concentrations are presented in Figure 60. The PON data are about an order of magnitude lower than the POC data. The model generally reproduces the PON data but over-estimates the CLIS PON concentrations. The 2006 results shown in Figure 61 differ in a similar way as the 2005 and 2006 POC comparisons with the G1 nitrogen showing the only real differences between the two years.

Figure 62 compares the 2005 sediment NH4 flux model results to the data. The flux data are higher in the western portion of the LIS and the model reproduces this pattern, but over-estimates the sediment NH4 flux at the WLIS, EXR, and ARTG stations. Improvements to the sediment NH4 flux model calibration will occur as the model calibration continues. Near zero sediment NH4 fluxes were measured in eastern LIS and the model calculates a small flux during the summer. The model calculates similar temporal and spatial patterns in the sediment NH4 fluxes in 2006 (Figure 63) and again over-estimates the data at the western LIS stations. As the model calibration continues and the sediment NH4 fluxes are improved, the water column NH4 model calibration will also improve.

Small sediment NO3 fluxes were measured at the LIS stations (Figure 64) with larger fluxes measured at the western LIS stations. The model calculates variable fluxes into and out of the sediment and in 2005 the model fluxes are a similar magnitude as the data. The model calculates similar sediment NO3 fluxes during 2006 as calculated in 2005 (Figure 65).

Figures 66 and 67 present model versus data comparisons for sediment denitrification. The calculated denitrification rates are small and the model over-estimates the data and will be further evaluated as the model calibration continues.

The model versus data comparisons of sediment PO4 fluxes for 2005 are presented in Figure 68. The data show higher fluxes in western LIS and during the summer. The model generally reproduces the fluxes during 2005 but shows the high sensitivity of the calculated fluxes to the model calculated low DO levels. The model calculated high sediment PO4 flux at the ELIS station is unexpected and will be evaluated further as the model calibration continues. Figure 69 presents the model and data for 2006 and shows that the model has less variability in the fluxes during 2006, presumably due to different temperatures and DO levels as compared to 2005. The high sediment PO4 flux is again calculated at the ELIS stations and will be further evaluated.

The model versus data SOD comparisons for 2005 and 2006 are compared in Figures 70 and 71. Both model and data show higher fluxes in western LIS and during the summer. The model generally reproduces the magnitude of the SOD data (particularly the 2005/06 data), with the exception of ARTG where the model over-estimates the data.

5.3 Algal Production

Figures 72 through 75 provide examples of the model calculated gross primary production (TGPP), total respiration (TRESP), and net community production (NCP) compared to the Goebel et al. 2002 and 2003 data. The Goebel paper only provides ranges per month at the stations, so the data do not represent specific stations. It would be expected that western stations, such as A4, would have rates on the higher end of the range and that these rates decrease further east. Therefore, as we compare the model to these rates, model results at A4 are compared to the high end of the data, and at a station like F2 are compared to the mid-range of the data. Also, note that the model is being compared to data from a different year, so year-to-year variations exist between the model output and data.

Figures 72 and 73 present the model comparison to TGPP, TRESP, and NCP data at station A4 with the 2005 and 2006 model results. The measurements and model results are for the photic zone. The photic zone is defined as the depth above where 1% of the surface light remains. Note that the model is comparing net primary production (NPP) to NCP. The model produces similar TGPP results for the two years and the model generally reproduces the peak TGPP measurements but tends to over-estimate the data during early and late summer. The model also compares favorably to the peak TRESP. NCP is also generally reproduced by the model, but with higher estimates to either side of the peak summer production. The model calculated TGPP, TRESP, and NCP all begin to increase earlier in 2006 as compared to 2005.

Figures 74 and 75 present another example of model and data comparisons of TGPP, TRESP, and NCP at station F2 for 2005 and 2006. The model results fall reasonably within the rates estimated by Goebel. Additional model to data comparisons with the Goebel data set are presented in Attachment 5.

Figures 76 and 77 compare Collins et al. (2013) data and model calculated TGPP, TRESP, and NCP near station I2. The Collins et al. dataset provided daily measurements and although the data are from 2010 the model compares favorably to the data. The data showed significant day to day variability, so the data are presented as monthly ranges (red symbols) and 5-day rolling averages (green lines). The model results are presented as daily averages (grey circles). The 2005 and 2006 model output generally reproduce the TGPP, TRESP, and NCP data and even shows a pattern similar to the 5-day rolling average data.

5.4 Skill Assessment Metrics

Model calibration is often accomplished through a subjective trial-and-error adjustment of model coefficients because many interrelated factors can influence model output. The experience and judgment of the modeler is a major factor in calibrating a model both accurately and efficiently. Although this method balances model comparison to data with the modeler's understanding of the physical, chemical, and biological characteristics of the system, it does not provide a quantitative measure of the "goodness of fit".

There is a large body of literature about coastal and estuarine modeling skill assessment (Blumberg et al., 1999; Fitzpatrick, 2009; Jolliff et al., 2009; Zhang et al., 2010; Ganju et al., 2016; and Ji, 2017). Typical measurements include relative error (RE), root mean square error (RMSE), correlation coefficient (r) and coefficient of determination (r^2). All statistical approaches have their limitations. Unfortunately, few references provide guidance as to the acceptable level of error for a satisfactory level of calibration for water quality models. Arhonditsis and Brett (2004) is one of the few examples of attempts to assess reasonable levels of calibration in water quality models. This reference as well as previous experience were used to guide the skill assessment statistical metric targets.

It should be noted that the correlation coefficient is dependent on the range of model and data with narrow ranges resulting in lower correlation coefficients as compared to model and data with wider ranges (https://www.bmj.com/rapid-response/2011/11/03/correlation-restricted-ranges-data-revisited, https://www.statisticshowto.com/restricted-range/).

Ultimately, the goal of model calibration is "not to curve fit model to data, but to describe the behavior of the data with a modeling framework of the principal mechanisms relevant to the problem" (Thomann, 1982). This ultimate goal requires a "weight of evidence" approach that balances both qualitative and quantitative skill assessment results with the model calibration guidance and acceptance targets provided by independent peer review.

The skill assessment metrics presented in the LIS Modeling Quality Assurance Project Plan (QAPP) (HDR, 2022) were used to perform a quantitative assessment of the model's ability to reproduce the available water quality data. The metrics included relative error, root mean square error and the correlation coefficient. Table 3 presents the targets for each metric. Meeting each metric is often dependent on the relative magnitude of the concentrations within a waterbody. When the magnitude varies widely within a waterbody the controlling metric can change. Using TN as an example, one area may have a TN concentration of 1.0 mgN/L and another with a concentration of 0.05 mgN/L. In the first case, it would be easier for the model to be within the 40% relative error target than being below 0.2 mgN/L RMSE. In the second case it would be easier for the model to meet the RSME than the RE because the concentrations are small and a small difference in the model calculation could still produce a large relative error. This is where the modeler must use judgment and weigh the qualitative and quantitative factors to assess the calibration.

- Relative error (RE): $= 100 \times \frac{|\bar{M}-\bar{O}|}{\bar{O}}$
- Root mean square error (RMSE): $= \sqrt{\sum_{i=1}^{n}\left(\frac{(M_i-O_i)^2}{n}\right)}$
- Correlation coefficient (r): $= \frac{\sum(O_i-\bar{O}) \times (M_i-\bar{M})}{\sqrt{\sum(O_i-\bar{O})^2} \times \sqrt{\sum(M_i-\bar{M})^2}}$

where: M_i – model output point

O_i – observed data point

$\bar{M}$ – model mean

$\bar{O}$ – observed data mean

n – number of observations

After a review of the model comparisons to the metrics in Table 3, the chosen targets could be reassessed as the model calibration continues.

Table 3. Skill Assessment Metric Targets[a]

Parameter	Relative Error[b]	RMSE[b]	Correlation Coefficient[b]
DO	< 10%	< 1.0 mg/L	> 0.8
BOD5	< 40%	< 0.75 mg/L	> 0.7
TOC	< 40%	< 1.0 mg/L	> 0.7
TN	< 40%	< 0.2 mgN/L	> 0.7
DIN	< 40%	< 0.1 mgN/L	> 0.7
NH4	< 40%	< 0.1 mgN/L	> 0.7
NO23	< 40%	< 0.1 mgN/L	> 0.7
TP	< 40%	< 0.05 mgP/L	> 0.7
PO4	< 40%	< 0.02 mgP/L	> 0.7
Chl-a	< 40%	< 15 µg/L	> 0.7

Note:

[a] These skill assessment metric targets will be updated in upcoming QAPP revisions.

[b] DEP tentatively agreed to these targets; however, final agreement is reserved until further MEG, DEP and EPA review and discussion.

5.5 Model Output Comparison to Metrics

Comparison of model to data using statistics can be challenging especially with limited data and constituents that can vary greatly both temporally over a day and spatially. The current approach compares hourly model output to the time of data collection, and for data without sampling times model output from noon was used. The skill assessment statistics were split between the unstratified season (Winter, November-April) and the stratified season (Summer, May-October) as requested by the MEG. Additionally, the stations were grouped into regions in LIS and the East River as shown in Figure 78. Currently, the statistics are first generated by station and then medians are developed for the different groups. This approach may be revisited as the model calibration continues either by taking model results over an averaging period (e.g., daily, weekly, or monthly), or grouping the model results and data at the stations in each group first and then calculating the statistics.

Grouped skill assessment statistics of the model are presented in Table 4 (Winter 2005), Table 5 (Summer 2005), Table 6 (Winter 2006), and Table 7 (Summer 2006) for the CTDEEP and DEP HS stations. In the future, we may group years together and calculate only summer and winter season statistics. Changing to seasonal skill assessment created some difficulties in meeting the RE targets for some constituents. For example, average NO23 concentrations can vary by more than an order of magnitude seasonally, so a small deviation from the data can result in a high RE.

Overall, further improvements to the model calibration for CY05 and CY06 are needed for certain parameters based on the skill assessment metrics and will be re-evaluated as the model calibration progresses to the remaining calibration years (CY07-14). This observation is also supported by the qualitative model-data comparisons presented in Sections 5.1 through 5.3. The revisions to the ROMS hydrodynamic model have resulted in meeting most all skill assessment targets in both years although additional revisions will be made to reduce the model calculated salinity levels as currently the model is biased high.

Many of the skill assessment metrics do not meet the targets (Table 3) set during QAPP development although for some parameters and groupings the targets are met for one or two metrics. For example, the correlation coefficient target is frequently met for DO and the RE target is frequently met for TOC. As the model calibration continues using additional data from CY07-14, the metrics will improve and revisions to the skill assessment calculation method may be warranted. These potential revisions may include grouping all station data for the regions before calculating the metrics by year or for all calibration years; and/or developing averaging intervals before calculating the metrics. In addition, the targets will be re-assessed as the model calibration continues.

Table 4. Grouped Skill Assessment Statistics for CTDEEP & DEP Stations – Winter 2005

Parameter	Region	Surface Layer			Bottom Layer		
		Median Relative Error	RMS Error	Correlation Coefficient	Median Relative Error	RMS Error	Correlation Coefficient
CHLA	West Narrows	67%	2.91	0.63	-	-	-
	East Narrows	68%	7.06	0.23	41%	2.98	0.31
	West Basin	121%	5.42	0.10	44%	2.70	0.17
	Central Basin	86%	3.31	0.77	48%	1.95	0.43
	East Basin	196%	3.12	0.40	131%	2.51	0.36
DIN	West Narrows	34%	0.25	0.18	-	-	-
	East Narrows	47%	0.14	0.65	38%	0.11	0.45
	West Basin	59%	0.09	0.65	48%	0.11	0.25
	Central Basin	48%	0.08	0.66	48%	0.10	0.50
	East Basin	27%	0.08	0.77	77%	0.10	0.82
NH4	West Narrows	33%	0.19	0.78	-	-	-
	East Narrows	86%	0.05	0.42	45%	0.05	0.29
	West Basin	84%	0.02	0.50	517%	0.05	0.33
	Central Basin	145%	0.02	0.45	365%	0.03	0.60
	East Basin	176%	0.01	0.70	164%	0.02	0.52
NO23	West Narrows	39%	0.16	0.37	-	-	-
	East Narrows	55%	0.11	0.67	48%	0.11	0.49
	West Basin	72%	0.10	0.51	67%	0.10	0.25
	Central Basin	74%	0.10	0.44	71%	0.10	0.35
	East Basin	54%	0.09	0.50	53%	0.09	0.52
DO	West Narrows	19%	1.49	0.98	24%	1.61	0.95
	East Narrows	22%	2.32	0.88	20%	1.90	0.94
	West Basin	15%	1.79	0.87	20%	1.80	0.95
	Central Basin	12%	1.30	0.90	15%	1.51	0.96
	East Basin	6%	1.05	0.95	9%	1.04	0.91
BOD5	West Narrows	-	-	-	-	-	-
	East Narrows	167%	1.41	0.69	186%	1.09	0.81
	West Basin	124%	1.15	0.62	108%	0.74	0.43
	Central Basin	61%	0.90	0.52	59%	0.70	0.53
	East Basin	22%	0.62	0.18	21%	0.48	0.50
PO4	West Narrows	61%	0.10	0.81	-	-	-
	East Narrows	23%	0.02	0.88	26%	0.03	0.89
	West Basin	29%	0.03	0.83	31%	0.03	0.87
	Central Basin	35%	0.03	0.87	42%	0.03	0.84
	East Basin	67%	0.02	0.31	79%	0.02	0.49

Parameter	Region	Surface Layer			Bottom Layer		
		Median Relative Error	RMS Error	Correlation Coefficient	Median Relative Error	RMS Error	Correlation Coefficient
SAL	West Narrows	3%	0.97	0.98	5%	1.28	0.93
	East Narrows	2%	0.73	0.81	1%	0.42	0.94
	West Basin	2%	0.65	0.83	1%	0.60	0.88
	Central Basin	1%	0.48	0.89	1%	0.47	0.91
	East Basin	1%	0.30	0.96	1%	0.49	0.92
TEMP	West Narrows	3%	0.59	1.00	3%	0.83	1.00
	East Narrows	11%	0.87	1.00	22%	1.57	0.99
	West Basin	8%	0.94	1.00	16%	1.41	1.00
	Central Basin	10%	1.03	1.00	16%	1.18	1.00
	East Basin	5%	0.91	1.00	9%	1.02	1.00
TN	West Narrows	22%	0.38	0.35	-	-	-
	East Narrows	25%	0.20	0.44	30%	0.18	0.24
	West Basin	46%	0.18	0.48	63%	0.20	0.61
	Central Basin	73%	0.21	0.21	82%	0.23	0.07
	East Basin	95%	0.24	0.19	111%	0.27	0.01
TP	West Narrows	59%	0.11	0.85	-	-	-
	East Narrows	48%	0.05	0.93	44%	0.05	0.94
	West Basin	46%	0.04	0.89	50%	0.04	0.89
	Central Basin	42%	0.04	0.90	49%	0.04	0.68
	East Basin	95%	0.04	0.51	75%	0.04	0.19
TOC	West Narrows	-	-	-	-	-	-
	East Narrows	13%	0.66	0.91	11%	1.35	0.34
	West Basin	10%	0.70	0.83	16%	0.81	0.50
	Central Basin	16%	0.85	0.53	20%	1.30	0.28
	East Basin	27%	1.89	0.26	18%	1.57	0.29

Table 5. Grouped Skill Assessment Statistics for CTDEEP & DEP Stations – Summer 2005

		Surface Layer			Bottom Layer		
Parameter	Region	Median Relative Error	RMS Error	Correlation Coefficient	Median Relative Error	RMS Error	Correlation Coefficient
CHLA	West Narrows	66%	7.38	0.38	-	-	-
	East Narrows	52%	8.16	0.23	48%	2.29	0.23
	West Basin	41%	3.93	0.52	88%	2.42	0.24
	Central Basin	68%	3.58	0.51	66%	2.14	0.15
	East Basin	44%	2.94	0.48	115%	2.72	0.27
DIN	West Narrows	80%	0.44	0.58	-	-	-
	East Narrows	88%	0.06	0.80	139%	0.15	0.70
	West Basin	68%	0.02	0.19	127%	0.11	0.51
	Central Basin	82%	0.02	-	172%	0.12	0.51
	East Basin	80%	0.03	0.11	94%	0.04	0.23
NH4	West Narrows	53%	0.20	0.76	-	-	-
	East Narrows	72%	0.02	0.61	127%	0.07	0.22
	West Basin	53%	0.01	0.09	118%	0.06	0.07
	Central Basin	71%	0.01	0.25	149%	0.06	0.67
	East Basin	40%	0.01	0.20	59%	0.02	0.41
NO23	West Narrows	114%	0.24	0.65	-	-	-
	East Narrows	100%	0.05	0.94	351%	0.08	0.53
	West Basin	95%	0.01	0.19	393%	0.07	0.41
	Central Basin	100%	0.01	-	285%	0.06	0.47
	East Basin	101%	0.03	0.49	389%	0.03	0.41
DO	West Narrows	18%	1.32	0.91	16%	1.05	0.94
	East Narrows	16%	2.51	0.32	13%	0.75	0.97
	West Basin	6%	1.07	0.62	8%	0.87	0.96
	Central Basin	11%	1.16	0.79	8%	0.70	0.97
	East Basin	13%	1.24	0.96	10%	0.78	0.99
BOD5	West Narrows	-	-	-	-	-	-
	East Narrows	76%	1.29	0.55	65%	0.90	0.32
	West Basin	94%	1.23	0.59	55%	0.73	0.40
	Central Basin	136%	1.28	0.27	155%	1.07	0.36
	East Basin	116%	0.96	0.41	103%	0.87	0.45
PO4	West Narrows	72%	0.09	0.88	-	-	-
	East Narrows	60%	0.01	0.88	31%	0.01	0.98
	West Basin	79%	0.01	0.81	48%	0.02	0.90
	Central Basin	82%	0.01	0.80	74%	0.02	0.70
	East Basin	52%	0.01	0.84	48%	0.01	0.82

Parameter	Region	Surface Layer			Bottom Layer		
		Median Relative Error	RMS Error	Correlation Coefficient	Median Relative Error	RMS Error	Correlation Coefficient
SAL	West Narrows	1%	0.48	0.90	2%	0.61	0.91
	East Narrows	1%	0.54	0.90	3%	0.82	0.96
	West Basin	1%	0.47	0.96	3%	0.83	0.97
	Central Basin	2%	0.75	0.80	3%	0.95	0.91
	East Basin	3%	1.34	0.75	5%	1.42	0.78
TEMP	West Narrows	3%	0.79	0.98	4%	0.97	0.97
	East Narrows	4%	1.05	0.99	5%	1.07	0.99
	West Basin	5%	1.13	0.99	4%	0.84	1.00
	Central Basin	5%	1.03	0.99	6%	0.94	1.00
	East Basin	8%	1.33	0.99	13%	2.10	0.99
TN	West Narrows	41%	0.60	0.14	-	-	-
	East Narrows	92%	0.34	0.31	110%	0.33	0.10
	West Basin	93%	0.23	0.32	155%	0.33	0.12
	Central Basin	89%	0.20	0.43	133%	0.29	0.24
	East Basin	85%	0.20	0.13	147%	0.24	0.04
TP	West Narrows	58%	0.10	0.70	-	-	-
	East Narrows	67%	0.05	0.75	70%	0.04	0.92
	West Basin	47%	0.02	0.89	75%	0.04	0.86
	Central Basin	68%	0.02	0.50	97%	0.04	0.67
	East Basin	94%	0.03	0.88	108%	0.03	0.38
TOC	West Narrows	-	-	-	-	-	-
	East Narrows	17%	1.17	0.30	22%	1.59	0.41
	West Basin	14%	1.00	0.16	26%	1.42	0.58
	Central Basin	15%	0.94	0.55	21%	1.36	0.52
	East Basin	21%	1.05	0.81	22%	0.96	0.71

Table 6. Grouped Skill Assessment Statistics for CTDEEP & DEP Stations – Winter 2006

Parameter	Region	Surface Layer			Bottom Layer		
		Median Relative Error	RMS Error	Correlation Coefficient	Median Relative Error	RMS Error	Correlation Coefficient
CHLA	West Narrows	108%	4.20	0.57	-	-	-
	East Narrows	37%	4.39	0.75	28%	1.26	0.91
	West Basin	50%	5.03	0.10	40%	2.08	0.27
	Central Basin	48%	2.22	0.90	23%	1.19	0.77
	East Basin	41%	1.76	-	21%	1.08	-
DIN	West Narrows	25%	0.26	0.41	-	-	-
	East Narrows	27%	0.10	0.91	30%	0.08	0.35
	West Basin	53%	0.05	0.95	31%	0.05	0.81
	Central Basin	72%	0.04	0.89	44%	0.06	0.35
	East Basin	26%	0.01	-	67%	0.04	-
NH4	West Narrows	16%	0.12	0.23	-	-	-
	East Narrows	75%	0.03	0.69	431%	0.06	0.70
	West Basin	102%	0.02	0.60	390%	0.05	0.03
	Central Basin	164%	0.02	0.51	435%	0.04	0.78
	East Basin	120%	0.01	-	207%	0.01	-
NO23	West Narrows	24%	0.18	0.50	-	-	-
	East Narrows	66%	0.07	0.94	57%	0.07	0.51
	West Basin	82%	0.06	0.92	59%	0.05	0.91
	Central Basin	89%	0.06	0.89	74%	0.06	0.60
	East Basin	42%	0.01	-	59%	0.05	-
DO	West Narrows	13%	1.13	0.98	17%	1.54	0.97
	East Narrows	15%	1.50	-	18%	1.93	-
	West Basin	12%	1.61	0.82	19%	1.71	0.94
	Central Basin	10%	0.99	-	12%	1.33	-
	East Basin	17%	1.00	-	8%	0.59	-
BOD5	West Narrows	-	-	-	-	-	-
	East Narrows	229%	1.65	0.49	256%	1.26	0.10
	West Basin	210%	1.28	0.17	119%	0.98	0.34
	Central Basin	52%	0.76	0.45	47%	0.69	0.53
	East Basin	75%	0.50	-	81%	0.47	-
PO4	West Narrows	65%	0.09	0.93	-	-	-
	East Narrows	53%	0.03	0.97	50%	0.03	0.95
	West Basin	57%	0.03	0.98	65%	0.04	0.96
	Central Basin	58%	0.03	0.97	81%	0.05	0.90
	East Basin	195%	0.06	-	170%	0.04	-

Parameter	Region	Surface Layer			Bottom Layer		
		Median Relative Error	RMS Error	Correlation Coefficient	Median Relative Error	RMS Error	Correlation Coefficient
SAL	West Narrows	6%	1.44	0.97	7%	1.91	0.84
	East Narrows	4%	1.08	0.89	6%	1.57	0.78
	West Basin	5%	1.20	0.92	6%	1.52	0.94
	Central Basin	5%	1.29	0.83	6%	1.76	0.73
	East Basin	6%	1.61	-	7%	2.19	-
TEMP	West Narrows	7%	1.00	1.00	7%	1.15	1.00
	East Narrows	9%	0.84	1.00	17%	1.35	1.00
	West Basin	10%	1.05	1.00	14%	1.49	1.00
	Central Basin	4%	0.91	1.00	14%	1.48	1.00
	East Basin	3%	0.44	-	5%	0.77	-
TN	West Narrows	31%	0.50	0.47	-	-	-
	East Narrows	71%	0.25	0.91	65%	0.24	0.84
	West Basin	67%	0.22	0.79	72%	0.22	0.75
	Central Basin	68%	0.19	0.21	82%	0.20	0.41
	East Basin	86%	0.14	-	62%	0.11	-
TP	West Narrows	62%	0.10	0.77	-	-	-
	East Narrows	79%	0.06	0.94	82%	0.06	0.97
	West Basin	87%	0.05	0.94	84%	0.06	0.97
	Central Basin	84%	0.05	0.71	114%	0.07	0.94
	East Basin	154%	0.07	-	132%	0.06	-
TOC	West Narrows	-	-	-	-	-	-
	East Narrows	22%	1.56	0.44	16%	0.74	0.75
	West Basin	21%	1.29	0.50	17%	1.15	0.57
	Central Basin	14%	0.68	0.75	23%	0.95	0.45
	East Basin	24%	0.87	-	22%	0.74	-

Table 7. Grouped Skill Assessment Statistics for CTDEEP & DEP Stations – Summer 2006

Parameter	Region	Surface Layer			Bottom Layer		
		Median Relative Error	RMS Error	Correlation Coefficient	Median Relative Error	RMS Error	Correlation Coefficient
CHLA	West Narrows	57%	3.75	0.55	-	-	-
	East Narrows	144%	9.78	0.41	40%	4.68	0.46
	West Basin	36%	4.99	0.16	35%	1.17	0.51
	Central Basin	29%	3.28	0.35	50%	1.98	0.22
	East Basin	94%	3.85	0.87	47%	1.90	0.49
DIN	West Narrows	73%	0.39	0.47	-	-	-
	East Narrows	84%	0.05	0.35	224%	0.20	0.48
	West Basin	75%	0.05	0.45	192%	0.12	0.57
	Central Basin	76%	0.03	0.33	161%	0.11	0.28
	East Basin	93%	0.08	0.27	99%	0.07	0.48
NH4	West Narrows	39%	0.20	0.16	-	-	-
	East Narrows	86%	0.04	0.43	122%	0.08	0.42
	West Basin	62%	0.04	0.36	181%	0.06	0.45
	Central Basin	79%	0.04	0.34	137%	0.06	0.39
	East Basin	18%	0.06	0.20	90%	0.07	0.55
NO23	West Narrows	126%	0.23	0.59	-	-	-
	East Narrows	100%	0.02	0.33	830%	0.12	0.53
	West Basin	98%	0.02	0.34	428%	0.08	0.61
	Central Basin	95%	0.01	0.50	276%	0.05	0.29
	East Basin	137%	0.04	0.25	87%	0.02	0.73
DO	West Narrows	15%	1.47	0.90	18%	1.47	0.91
	East Narrows	25%	2.29	0.64	18%	1.72	0.75
	West Basin	9%	1.57	0.73	18%	1.38	0.89
	Central Basin	9%	1.17	0.66	12%	1.10	0.91
	East Basin	13%	1.32	0.97	8%	0.64	0.98
BOD5	West Narrows	-	-	-	-	-	-
	East Narrows	87%	1.72	0.58	242%	1.18	0.55
	West Basin	230%	1.65	0.32	329%	1.11	0.30
	Central Basin	148%	1.33	0.30	134%	0.89	0.33
	East Basin	194%	1.07	0.39	126%	0.83	0.35
PO4	West Narrows	77%	0.08	0.85	-	-	-
	East Narrows	50%	0.01	0.96	49%	0.03	0.93
	West Basin	79%	0.01	0.96	92%	0.02	0.92
	Central Basin	87%	0.01	0.86	30%	0.02	0.92
	East Basin	81%	0.02	0.80	44%	0.01	0.80

Parameter	Region	Surface Layer			Bottom Layer		
		Median Relative Error	RMS Error	Correlation Coefficient	Median Relative Error	RMS Error	Correlation Coefficient
SAL	West Narrows	4%	1.19	0.83	6%	1.73	0.83
	East Narrows	6%	1.48	0.59	7%	1.76	0.49
	West Basin	5%	1.32	0.60	7%	1.73	0.39
	Central Basin	5%	1.37	0.45	6%	1.67	0.08
	East Basin	3%	1.32	0.68	7%	1.78	0.23
TEMP	West Narrows	5%	1.14	0.97	4%	0.89	0.98
	East Narrows	6%	1.32	0.98	2%	0.91	0.98
	West Basin	7%	1.85	0.98	3%	0.65	0.99
	Central Basin	5%	1.25	0.99	4%	0.67	1.00
	East Basin	5%	0.87	0.99	4%	0.82	0.99
TN	West Narrows	30%	0.48	0.14	-	-	-
	East Narrows	61%	0.26	0.53	103%	0.31	0.46
	West Basin	49%	0.18	0.47	113%	0.29	0.24
	Central Basin	55%	0.17	0.51	100%	0.26	0.37
	East Basin	54%	0.19	0.57	81%	0.19	0.53
TP	West Narrows	71%	0.09	0.82	-	-	-
	East Narrows	95%	0.06	0.87	56%	0.04	0.91
	West Basin	63%	0.03	0.86	119%	0.05	0.85
	Central Basin	67%	0.02	0.90	113%	0.04	0.90
	East Basin	118%	0.04	0.80	81%	0.03	0.29
TOC	West Narrows	-	-	-	-	-	-
	East Narrows	17%	1.13	0.33	18%	1.02	0.79
	West Basin	14%	0.85	0.23	15%	1.19	0.29
	Central Basin	15%	1.05	0.26	20%	1.07	0.40
	East Basin	24%	1.40	0.20	26%	1.66	0.23

6 Discussion and Next Steps

A preliminary model calibration of the LIS RCA water quality model was completed using data for CY05 and CY06 with over 125 model runs completed to investigate various model coefficients both in groups and alone. The preliminary model calibration also involved use of improved ROMS hydrodynamic model calibration inputs. CY05 and CY06 are two years of the 10-year model calibration period of CY05-14. The goals of the preliminary calibration were to develop the process of creating model inputs and assessing the resulting model calibration using these approaches. Additionally, the preliminary calibration process provided an understanding of how and where the preliminary calibration may be improved during the calibration to the full time-period and for further review and discussion with the MEG.

Model versus data comparisons were completed qualitatively using visual graphical comparisons and quantitatively using skill assessment statistical metrics. Preliminary calibration included model-data comparisons to chl-a, organic and inorganic nitrogen, phosphorus, and silica, organic carbon, and DO data from CTDEEP and DEP sources. The monitoring stations from these data sources included areas throughout LIS, Hudson River, and NY/NJ Harbor. A summary of the preliminary calibration efforts is presented below along with the next steps planned to further improve the model calibration to the full time-period (CY05-14).

6.1 Model Calibration

The accurate calibration of the model to phytoplankton (algal) biomass, nutrients, and DO is a delicate balance between the factors affecting algal light, nutrient, and temperature limitations, nutrient and carbon biogeochemical processes, as well as algal growth and respiration, predation and settling rates coupled with proper representation of hydrodynamic mixing and transport processes. In addition, these factors must be controlled to produce the correct magnitude of algal primary production. The magnitude of many of these factors is not precisely known, so they are typically estimated through a trial-and-error process bounded by literature or prior modeling experience.

The preliminary calibration process has not only involved adjusting the coefficients that control these processes, but it also involved modifications to the RCA model code. Modifications included the addition of a TSS state-variable, changes to the light extinction formulation, adjustment of the phytoplankton predation formulation, and use of a variable carbon to chl-a ratio.

It should be noted that the model currently uses model rate coefficients that are spatially constant and adjusted by water temperature. For example, the algal growth rates or carbon oxidation rates assigned in the model are the same in all locations (e.g., LIS, coastal ocean, East River, Hudson River). Further evaluation will be completed during calibration to the full time-period to assess whether assigning spatial variability to certain model coefficients can help improve model calibration.

Qualitatively, the model compares favorably to the majority of the parameters, but the time-series figures show there are areas that can be improved. The chl-a calibration generally looks good but there are regional differences, such as the existence of a spring bloom, that the model is not able to differentiate. Some of the nutrient issues are related to initial conditions and high sediment nutrient fluxes, which will be addressed. The increased algal predation has improved the comparison to DOC and TOC, but it is not clear yet whether the good comparisons will carry into the other calibration years when the measured DOC is lower. The model calibration to DO looks good, but the BOD5 is over-estimated, so improving the comparison to BOD5 may impact the DO calibration. Based on the time-series figures, the current calibration is considered good for expanding the calibration time-period to determine if the current model inputs can also reproduce the additional data during the other calibration years.

Quantitatively, the model needs improvement related to meeting skill assessment targets. This includes completing some additional analysis to determine the best method to calculate the statistics for fairly comparing the model to data. Typically, models do not compare well to data on an hour-to-hour basis, as the metrics are calculated now, due to tidal variability, algal photosynthesis and respiration variation over the day, and small-scale mixing processes not represented due to model segment size. As the model calibration continues, additional methods will be evaluated for calculating the statistics.

6.2 Next Steps

There are still areas where the preliminary model calibration needs to be improved. These other areas for improvement tend to be site-specific and need to be addressed locally and not universally in the model. For example, the model calibration of algal growth, DO, and nutrient levels in western LIS and the East River (West Narrows) could still be improved to better capture the DO decrease observed in July/August and the winter/spring algal blooms. Additionally, model calibration improvements of DO, phosphorus, and silica levels in the Hudson River could also be improved.

Some of these model calibration improvements can be addressed through re-evaluation of external loadings (e.g., silica in the Hudson River) or adjustments to sediment flux model coefficients to control sediment nutrient fluxes with the overlying water column. In some cases, the timing of algal blooms in LIS could be improved in the model that will require re-evaluation of algal temperature optimums and algal growth rates. The timing of the winter/spring algal blooms in the model will also impact nutrient levels, which could also use improvement in the model calibration.

The model is sensitive to the RCA water quality model initial conditions used for a model run, especially for nutrients (e.g., if the initial nutrient concentrations are too high, the model can over-predict the spring algal bloom). This effect is most noticeable for CY05 model runs as CY04 model runs have not been completed for beginning CY05 model runs. As the additional model calibration years are used, this initial condition effect will become less of an issue although the model needs to reproduce data at the end of a year as the results are used to begin a model run for the next year.

To date, the preliminary model calibration has focused on CY05 and CY06 datasets. The river and WRRF flows, and associated loads, were similar between these two years and further model calibration to the remaining calibration years (CY07-14) will provide a wider range of environmental conditions for calibrating the LIS models. These additional model calibration years will provide the opportunity to develop a consistent set of model calibration coefficients that best represents the observed data for CY05-14.

HDR has identified the following areas for RCA model calibration improvement as model calibration continues using data from the full time-period (CY05-14).

- DO calibration in certain areas (i.e., western LIS/East River, Hudson River) – The model preliminary calibration reproduces the low bottom layer DO levels in most areas of LIS but further improvements are needed to better represent the low DO levels in western LIS and in the East River during July and August. The main area of focus will be on algal growth in these locations. In addition, DO levels in the Hudson River tend to be lower than the observed data upstream and higher than the data downstream during summer months. The areas of focus will include algal growth, carbon oxidation rates and atmospheric reaeration.
- Inorganic nutrient calibration (NH4, NO23, PO4) – The model tends to over-predict inorganic nutrient levels in LIS during the summer months in the bottom layer. A similar observation occurs in the East River and in the Hudson River (although mainly for PO4). The model over-prediction of inorganic nutrient levels will focus on the sediment flux model coefficients, net settling rates, and the effect they have on calculated sediment nutrient fluxes. In the Hudson River area, external loads will also be re-evaluated.
- Winter/spring algal bloom calibration – In western LIS, the timing of the winter/spring algal bloom has a direct impact on the calculated DO levels during these time-periods. At some locations in western LIS, the winter/spring bloom is under-predicted and could be better reproduced. The areas of focus will include the winter/spring algal group growth rates and algal growth temperature optimums. In addition, re-evaluation of the variable carbon to chl-a formulation will be reviewed as at some LIS locations the PC model calibration is good but the chl-a calibration could be improved.
- TN and TP calibration – Primarily in LIS, the model over-predicts the TN and TP data. The over-prediction appears to be primarily related to an over-prediction of total dissolved nitrogen (TDN) and total dissolved phosphorus (TDP). The area of focus will include algal recycle fractions and improved model calibration to inorganic nutrients. In addition, external model loads will be re-evaluated.
- ROMS hydrodynamic model calibration – Further improvements in the hydrodynamic model calibration to salinity and water elevations are planned. The focus areas for the calibration improvements are adjustments to the assigned salinity boundary condition inputs and water elevation tidal range. In addition, the model calibration of water temperatures in eastern LIS could be improved and will also focus on adjustments to boundary condition inputs (particularly in the bottom layers).

7 References

Arhonditsis, G.B. and M.T. Brett. 2004. Evaluation of the current state of mechanistic aquatic biogeochemical modeling. Marine Ecology Progress Series, 271:13–26.

Balcom, P.H., J.M. Godfrey, D. C. Bennett1, G. A. Grenier, C. G., Cooper, D. R. Cohen, D. A. Arbige, and W. F. Fitzgerald. 2007. Deploying Benthic Chambers to Measure Sediment Oxygen Demand in Long Island Sound. Proceedings of the American Academy of Underwater Sciences 26th Symposium. Dauphin Island, AL: AAUS; 2007.

Barclay, J.R. and J.R. Mullaney, 2021. Simulation of groundwater budgets and travel times for watersheds on the north shore of Long Island Sound, with implications for nitrogen-transport studies. U.S. Geological Survey Scientific Investigations Report 2021–5116, 84 p., https://doi.org/ 10.3133/ sir20215116.

Blumberg, A., L.A. Khan and J. St. John, 1999. Three-dimensional hydrodynamic model of New York Harbor Region. ASCE Jour. Hydraulic Engineering.

Calbet, A., 2008. The trophic roles of microzooplankton in marine systems. ICES Journal of Marine Science, 65: 325-331.

Cerco, C.F. and M.R. Noel, 2004. Process-based primary production modeling in Chesapeake Bay. Mar. Ecol. Prog. Ser. 282: 45-58, 2004.

Collins, J.R, P.A. Raymond, W.F. Bohlen, and M.M. Howard-Strobel. 2013. Estimates of New and Total Productivity in Central Long Island Sound from In Situ Measurements of Nitrate and Dissolved Oxygen. Estuaries and Coasts (2013) 36:74–97.

Cooperative Institute for Research in Environmental Sciences (CIRES) at the University of Colorado, Boulder. 2014: Continuously Updated Digital Elevation Model (CUDEM) - 1/9 Arc-Second Resolution Bathymetric-Topographic Tiles. [indicate subset used]. NOAA National Centers for Environmental Information. https://doi.org/10.25921/ds9v-ky35.

Defne, Z., F. Spitz, V. DePaul, and T.A. Wool, 2017. Toward a Comprehensive Water-Quality Modeling of Barnegat Bay: Development of ROMS to WASP Coupler, Journal of Coastal Research 78(sp1), 34-45, https://doi.org/10.2112/SI78-004.1

Fitzpatrick, J.J., 2009. Assessing skill of estuarine and coastal eutrophication models for water quality managers. Jour. Marine Systems, 76: 195-211.

Foreman, M.G.G. 1977. Manual for Tidal Height Analysis and Prediction. Pacific Marine Science Report 77-10. Revised July 1996. Institute of Marine Sciences, Patricia Bay, Victoria, BC.

Ganju, N.K., M.J. Brush, B. Rashleigh, A.L. Aretxabaleta, P. del Barrio, J.S. Grear, L.A. Harris, S.J. Lake, G. McCardell, J. O'Donnell, D.K. Ralston, R.P. Signell, J.M. Testa and J.M.P. Vaudrey, 2016. Progress and Challenges in Coupled Hydrodynamic-Ecological Estuarine Modeling, Estuaries & Coasts, 39: 311-332.

Goebel, N. L., J.N. Kremer, and C. A. Edwards. 2006. Primary Production in Long Island Sound. Estuaries and Coasts, Vol. 29, No. 2, p. 232–245, April 2006.

HDR, 2021. Hydrodynamic & Water Quality Model Selection and Setup. For the New York City Department of Environmental Protection DEP LIS-HWQMS Project. Contract: BEPA-LIS-HWQMS; PIN: 82619BEPALIS.

HDR, 2022. Quality Assurance Project Plan (v 2.0), Long Island Sound Hydrodynamic and Water Quality Modeling. 8/22/2022.

HDR, 2023. LIS ROMS Hydrodynamic Model Inputs and RCA Water Quality Model Load Development Approach Memo. 1/6/2023.

HydroQual, 2005. Application of RAINMAN to New York/New Jersey Harbor. Interim Submittal. Submitted to the NY/NJ Harbor Estuary Program, Pathogen Work Group. September 2005.

Ji, Zhen-Gang, 2017. Hydrodynamics and Water Quality Modeling Rivers, Lakes, and Estuaries. John Wiley & Sons, Hoboken, NJ, 581 pp.

Jolliff, J.K., J.C. Kindle, I. Shulman, B. Penta, M.A.M. Friedrichs, R. Helber, and R.A. Arnone, 2009. Summary diagrams for coupled hydrodynamic-ecosystem model skill assessment. Journal of Marine Systems, 76: 64–82.

Levin, Julia, John Wilkin, Naomi Fleming, and Javier Zavala-Garay, 2018. Mean circulation of the Mid-Atlantic Bight from a climatological data assimilative model, Ocean Modelling, Volume 128, Pages 1-14, ISSN 1463-5003, https://doi.org/10.1016/j.ocemod.2018.05.003.

Lopez, A., D. Carey, J. Carlton, R. Cerrato, H. Dam, R. DiGiovanni, C. Elphick, M. Frisk, C. Gobler, L. Hice, P. Howell, A. Jordaan, S. Lin, D. Lonsdale, M. McEnroe, K. McKown, G. McManus, R. Orson, B. Peterson, C. Pickerell, R. Rozsa, A. Siuda, E. Thomas, G. Taylor, S. Shumway, S. Talmage, M. Van Patten, J. Vaudrey, G. Wikfors, C. Yarish, and R. Zajac, 2013. Biology and Ecology of Long Island Sound. Springer Series on Environmental Management.

Mazur, C.I., A.N Al-Haj, N.E. Ray, I. Sanchez-Viruet, and R.W. Fulweiler, 2021. Low denitrification rates and variable benthic nutrient fluxes characterize Long Island Sound sediments. Biogeochemistry. https://doi.org/10.1007/s10533-021-00795-7.

Misut, P.E., N.A. Casamassina and D.A. Walter, 2021. Delineation of areas contributing groundwater and travel times to receiving waters in Kings, Queens, Nassau, and Suffolk Counties, New York: U.S. Geological Survey Scientific Investigations Report 2021–5047, 61 p., https://doi.org/ 10.3133/ sir20215047.

Moore, A.M., H.G. Arango, G. Broquet, B.S. Powell, A.T. Weaver, and J. Zavala-Garay. 2011. The Regional Ocean Modeling System (ROMS) 4-Dimensional Variational Data Assimilation Systems: I - System overview and formulation. Progress in Oceanography. Vol. 91, Issue 1, October 2011.

O'Donnel, J., G. McCardell, and H. Dam. 2010. Simulation of Long Island Sound with the System-Wide Eutrophication Model (SWEM): Inter-annual Variability and Sensitivity. EPA Grant Number: LI-97127101.

Poppe, L.J.; Knebel, H.J.; Mlodzinska, Z.J.; Hastings, M.E., and Seekins, B.A, 2000. Distribution Of Surficial Sediment in Long Island Sound and Adjacent Waters: Texture and Total Organic Carbon. Journal of Coastal Research, 16(3), 567-574. West Palm Beach (Florida), ISSN 0749-0208.

Schmidt, S.R. and M.M. Whitney, 2018. A Model Study on the Summertime Distribution of River Waters in Long Island Sound. Estuaries and Coasts 41:1002-1020.

Testa, J. M., Y. Li, Y.-J. Lee, M. Li, D. C. Brady, D. M. Di Toro, W. M. Kemp, and J. J. Fitzpatrick, 2014, Quantifying the effects of nutrient loading on dissolved O2 cycling and hypoxia in Chesapeake Bay using a coupled hydrodynamic-biogeochemical model, J. Mar. Syst., 139, 139–152, doi:10.1016/j.jmarsys.2014.05.018.

Thomann, R.V., 1982. Verification of Water Quality Models. Journal of Environmental Engineering, Vol. 108, No. EE5. October 1982.

Xu, J., R.R. Hood, and S. Chao, 2005. A Simple Empirical Optical Model for Simulating Light Attenuation Variability in a Partially Mixed Estuary. Estuaries. Vol. 28, No. 4, p. 572-580 August 2005.

Zhang, A. KW Hess & F. Aikman, 2010. User-based skill assessment techniques for operational hydrodynamic forecast systems. Vol 3, No. 2, Journal of Operational Oceanography.

8 Figures

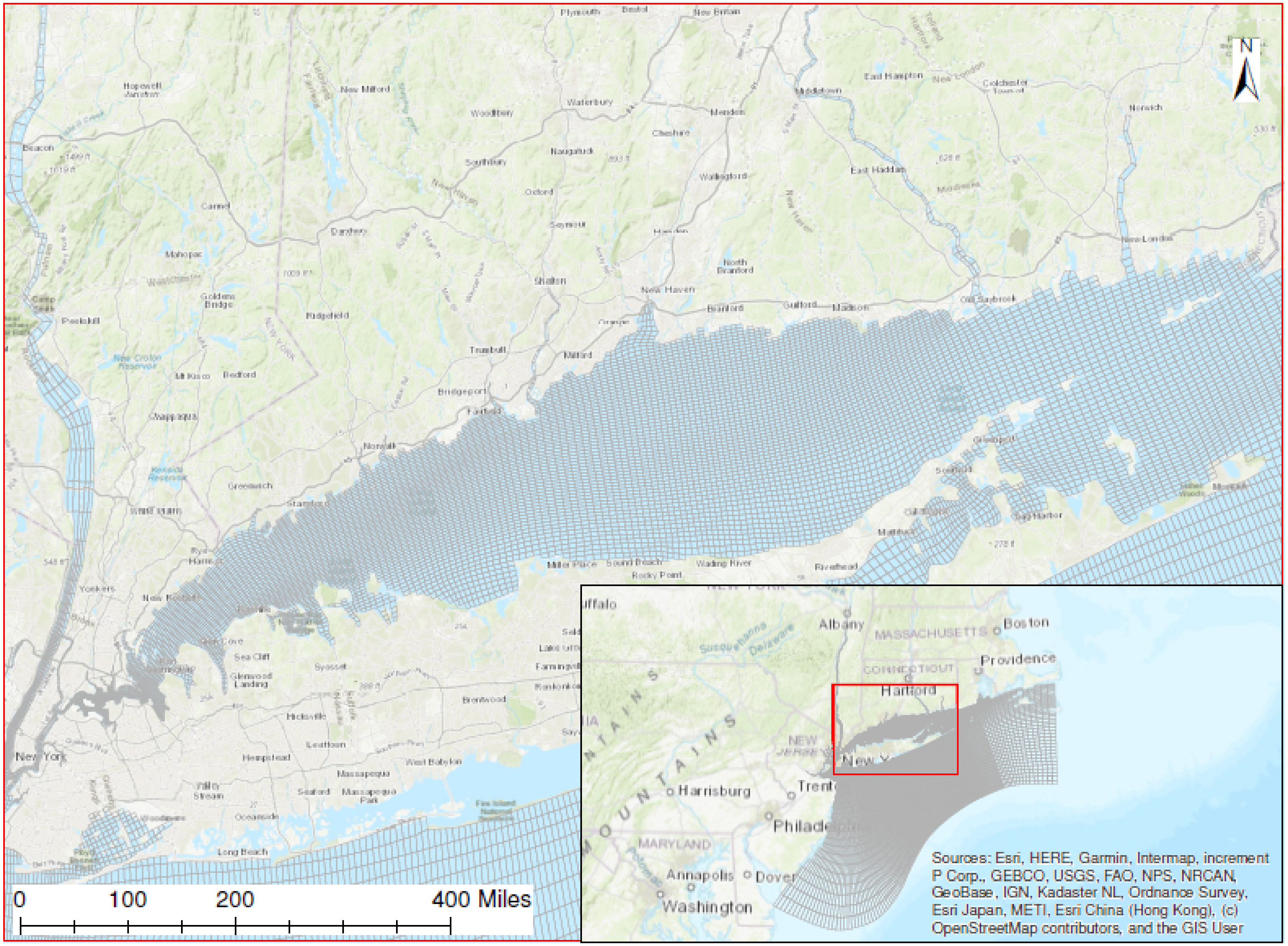

Figure 1. Project Area and Model Grid

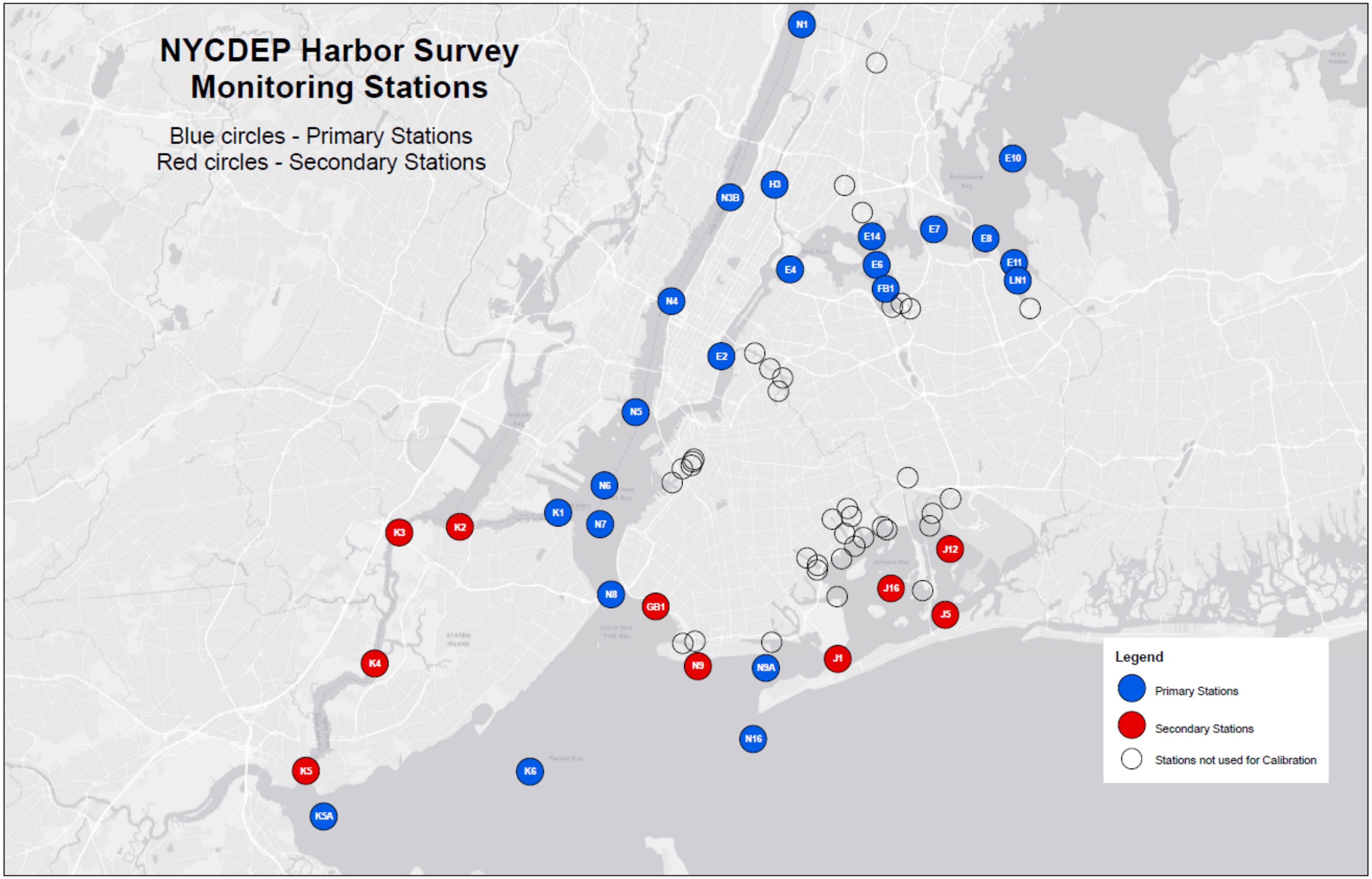

Figure 2. DEP Harbor Survey Water Quality Monitoring Stations

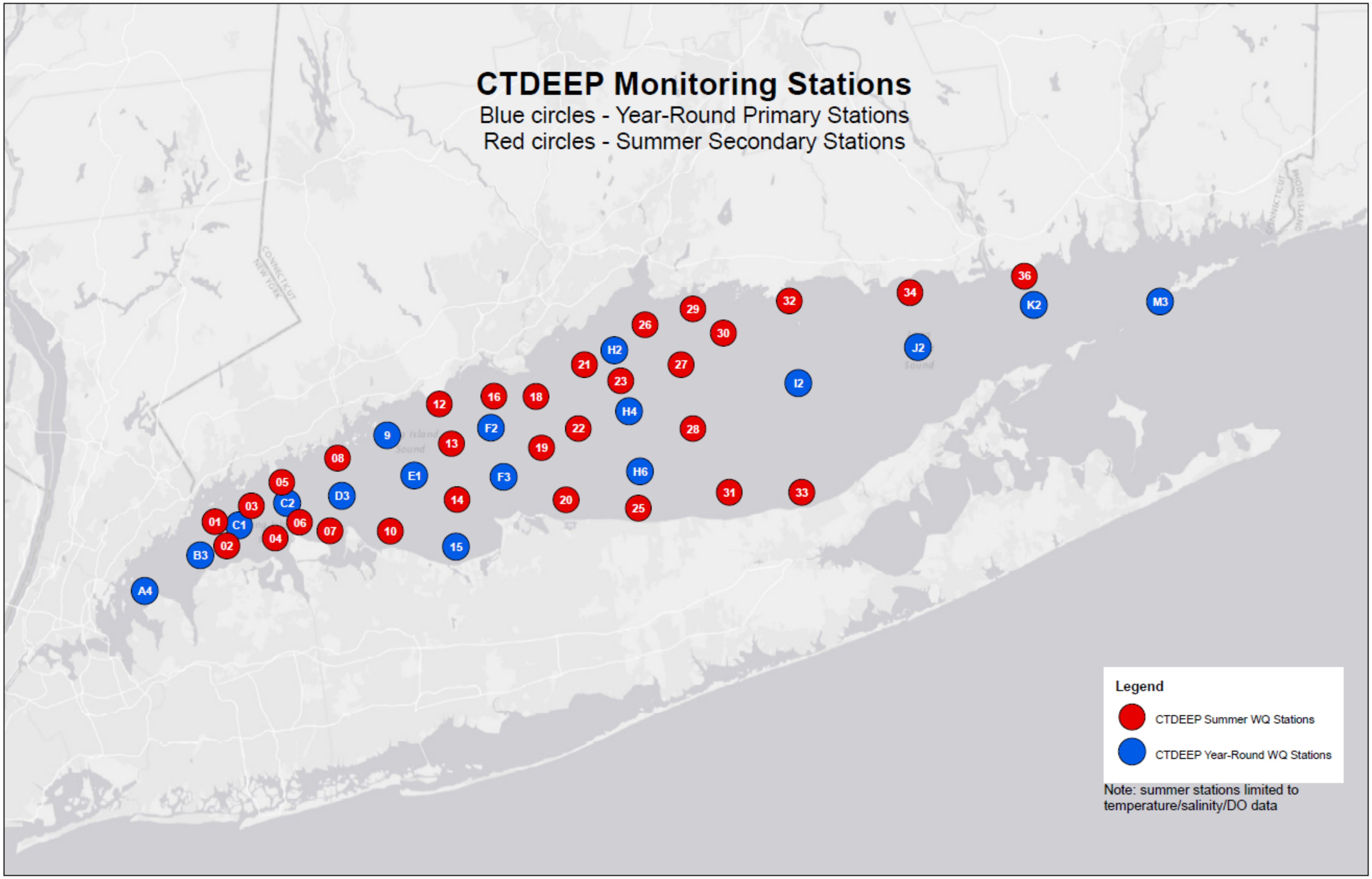

Figure 3. CTDEEP Water Quality Monitoring Stations

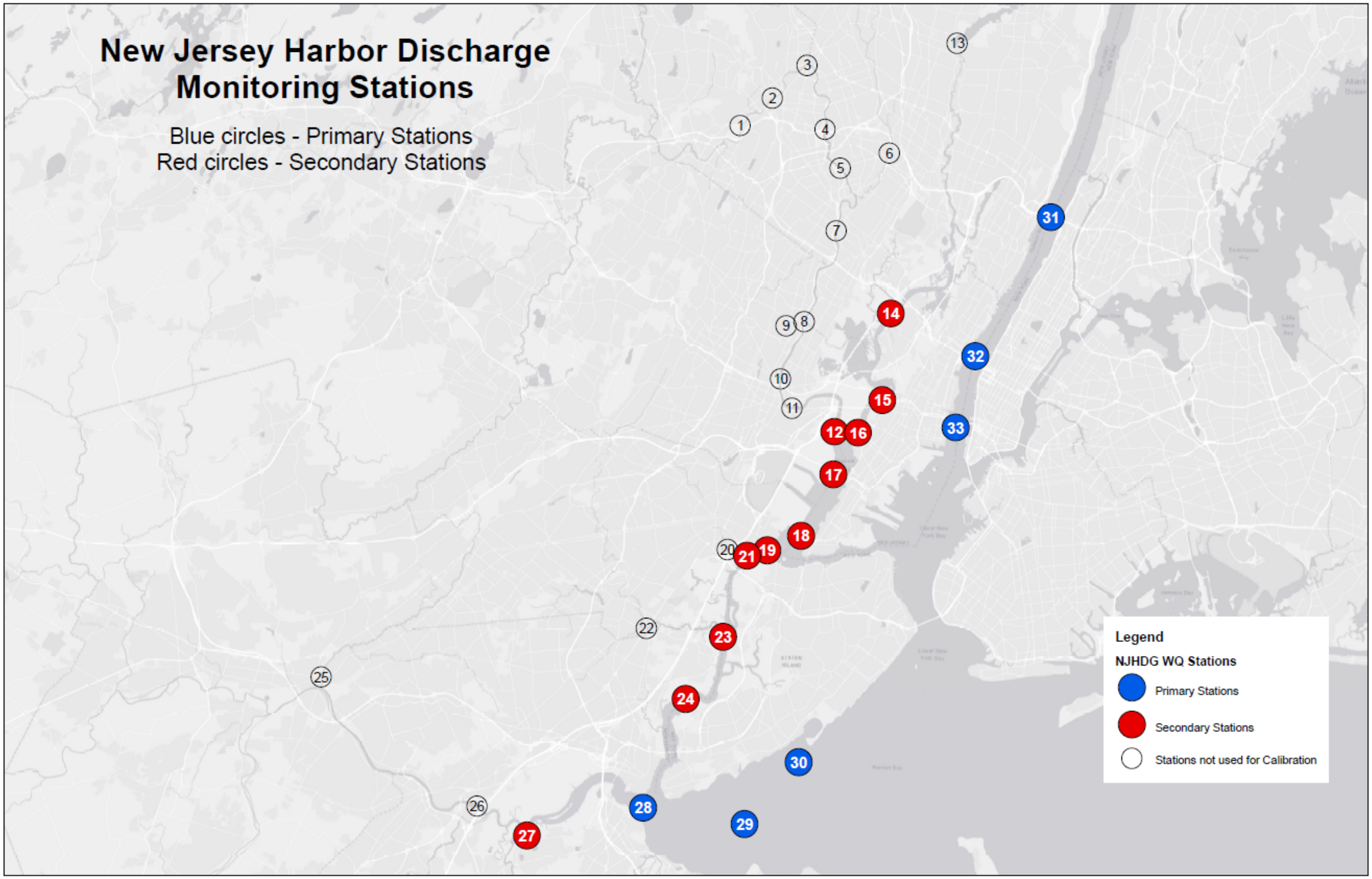

Figure 4. NJHDG Water Quality Monitoring Stations

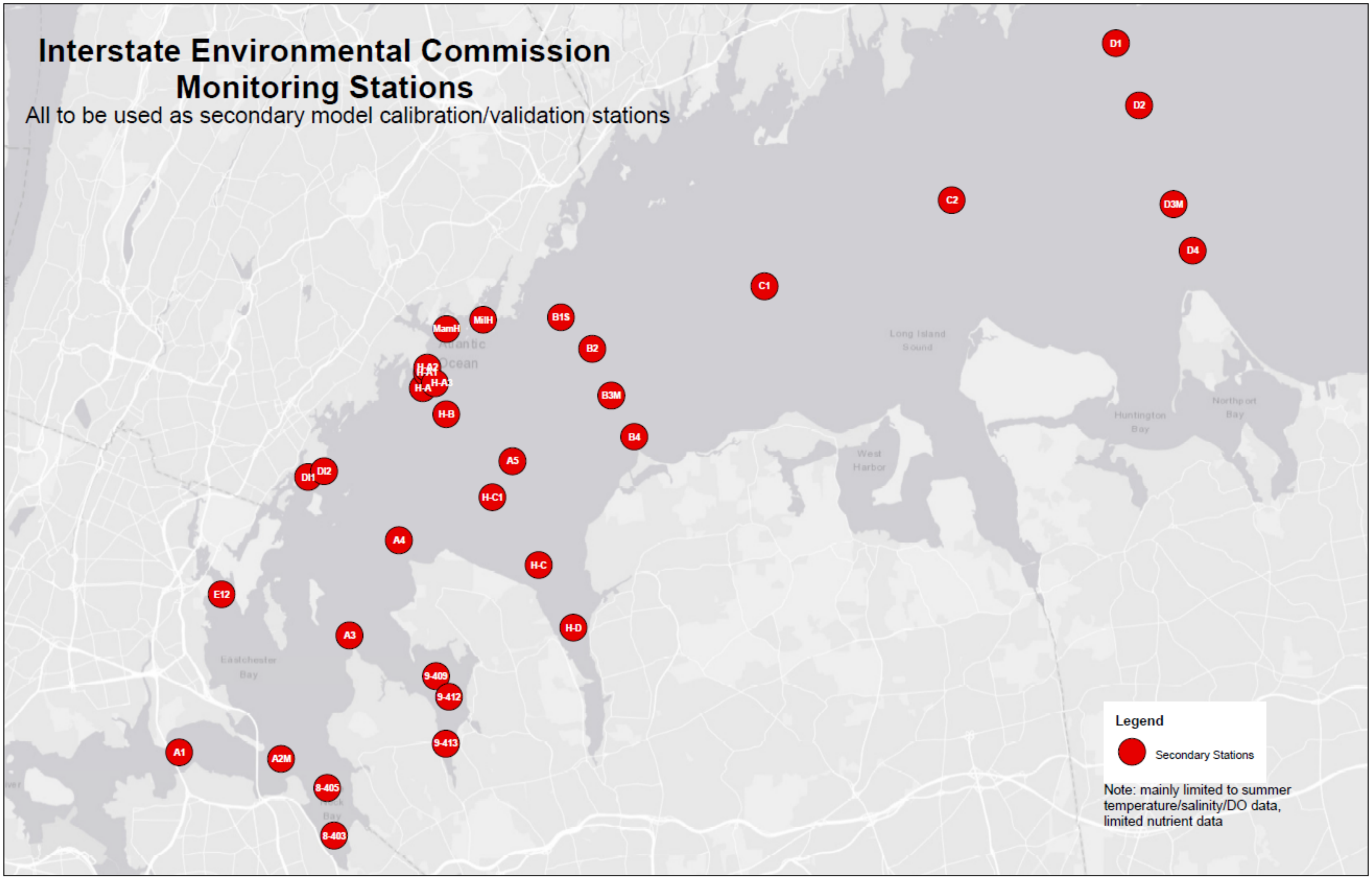

Figure 5. IEC Water Quality Monitoring Stations

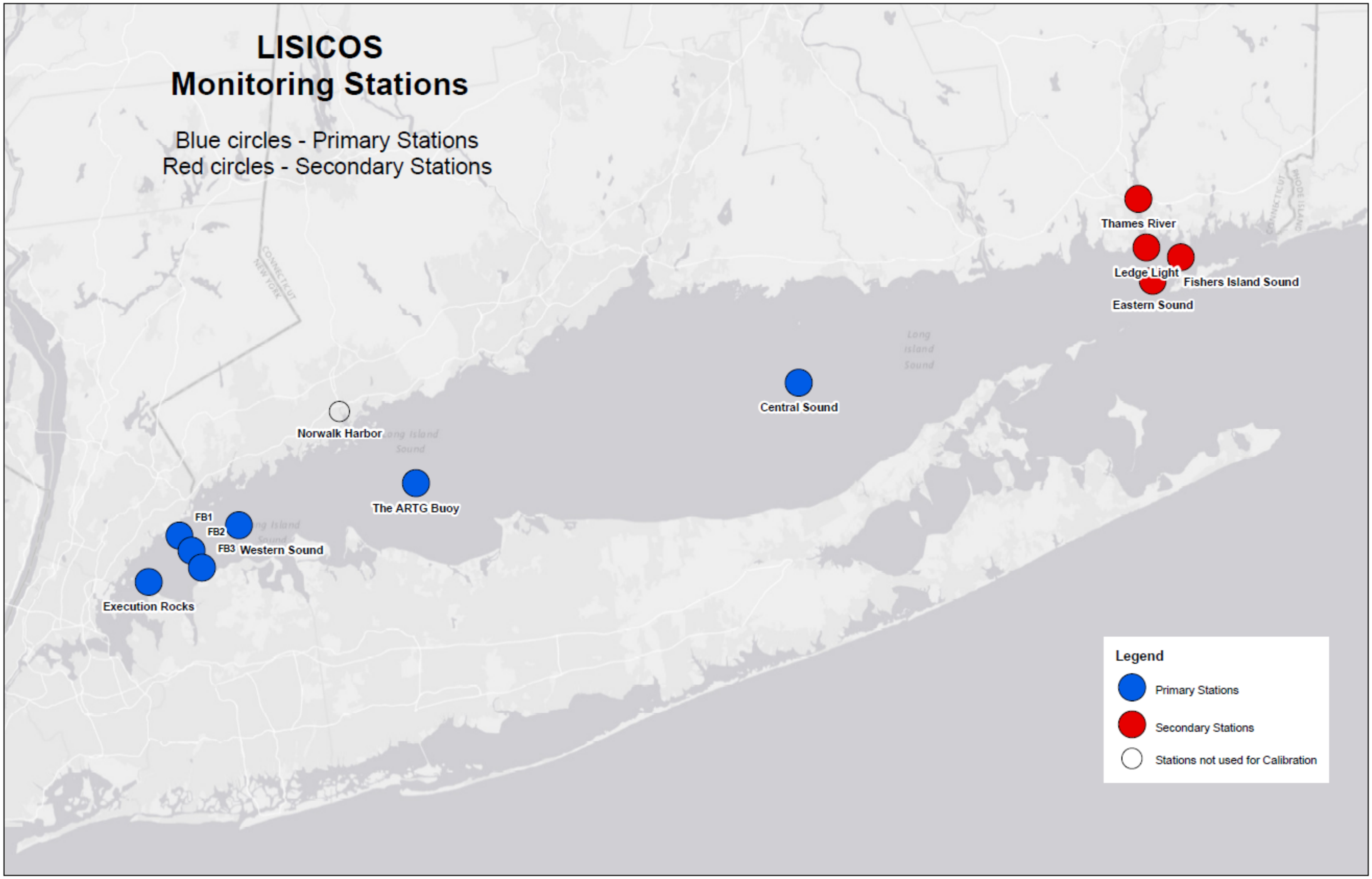

Figure 6. LISICOS Monitoring Stations

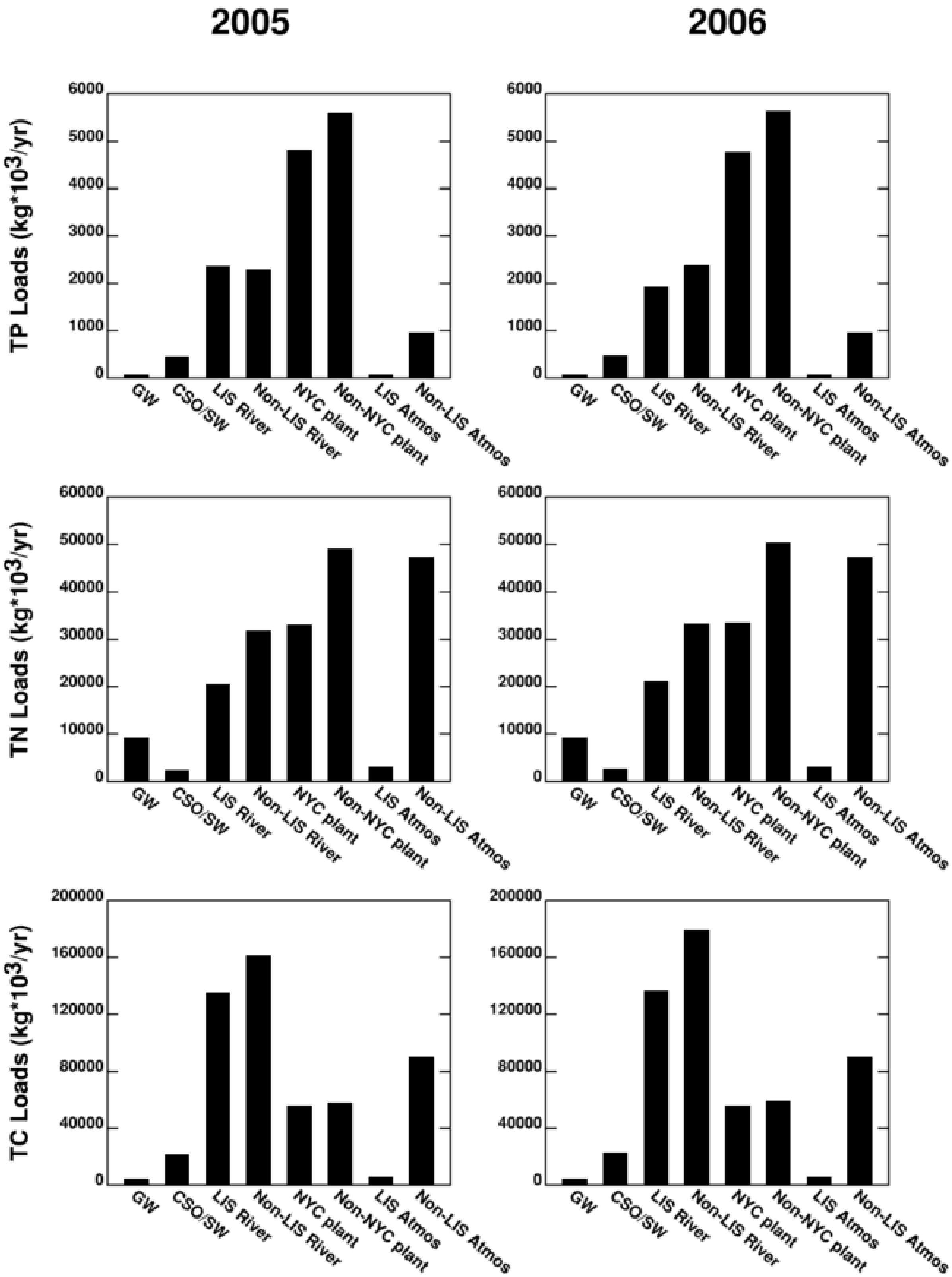

Figure 7. Model Loading Summary

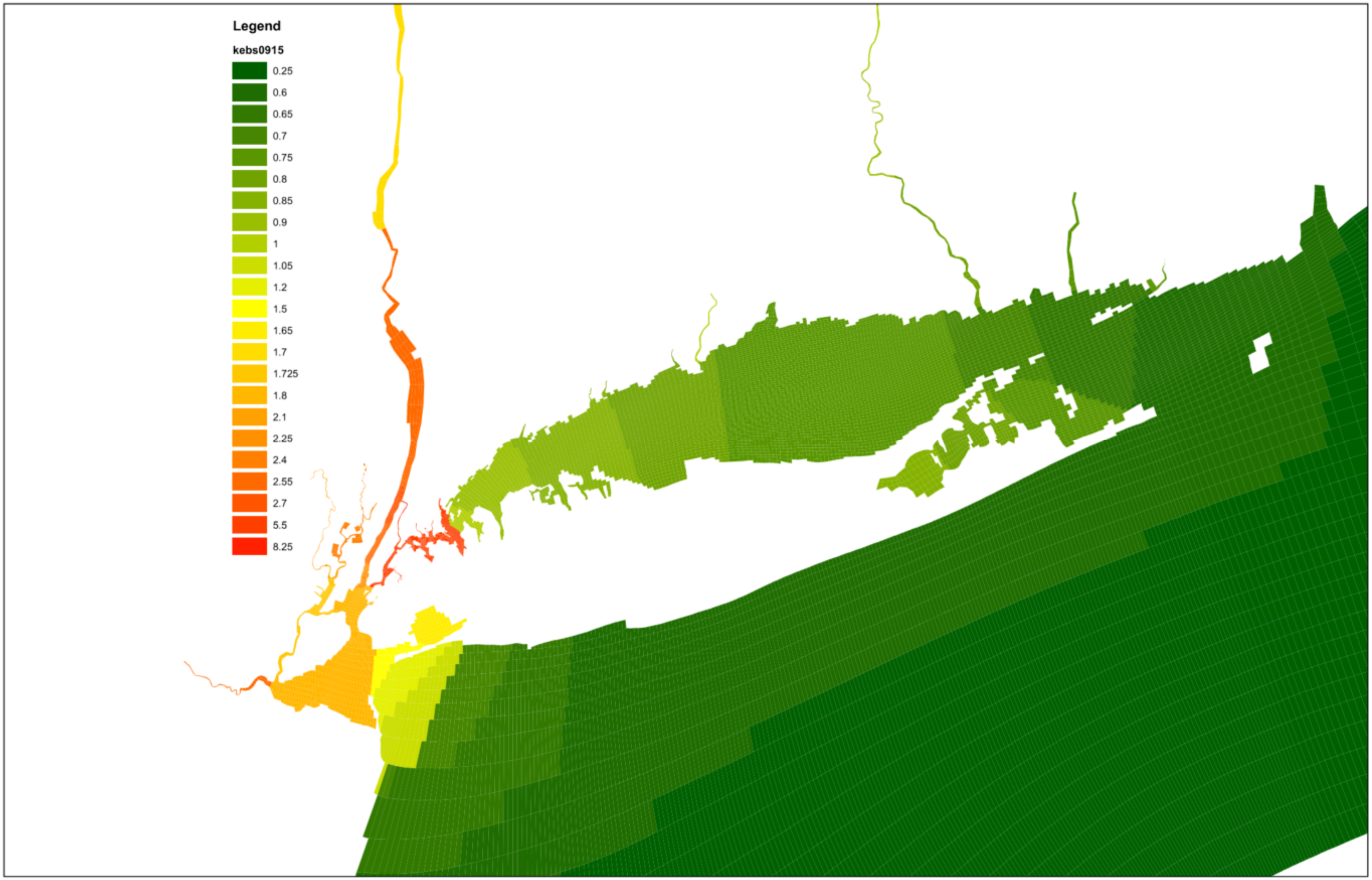

Figure 8. Spatially Variable Assigned Kebase

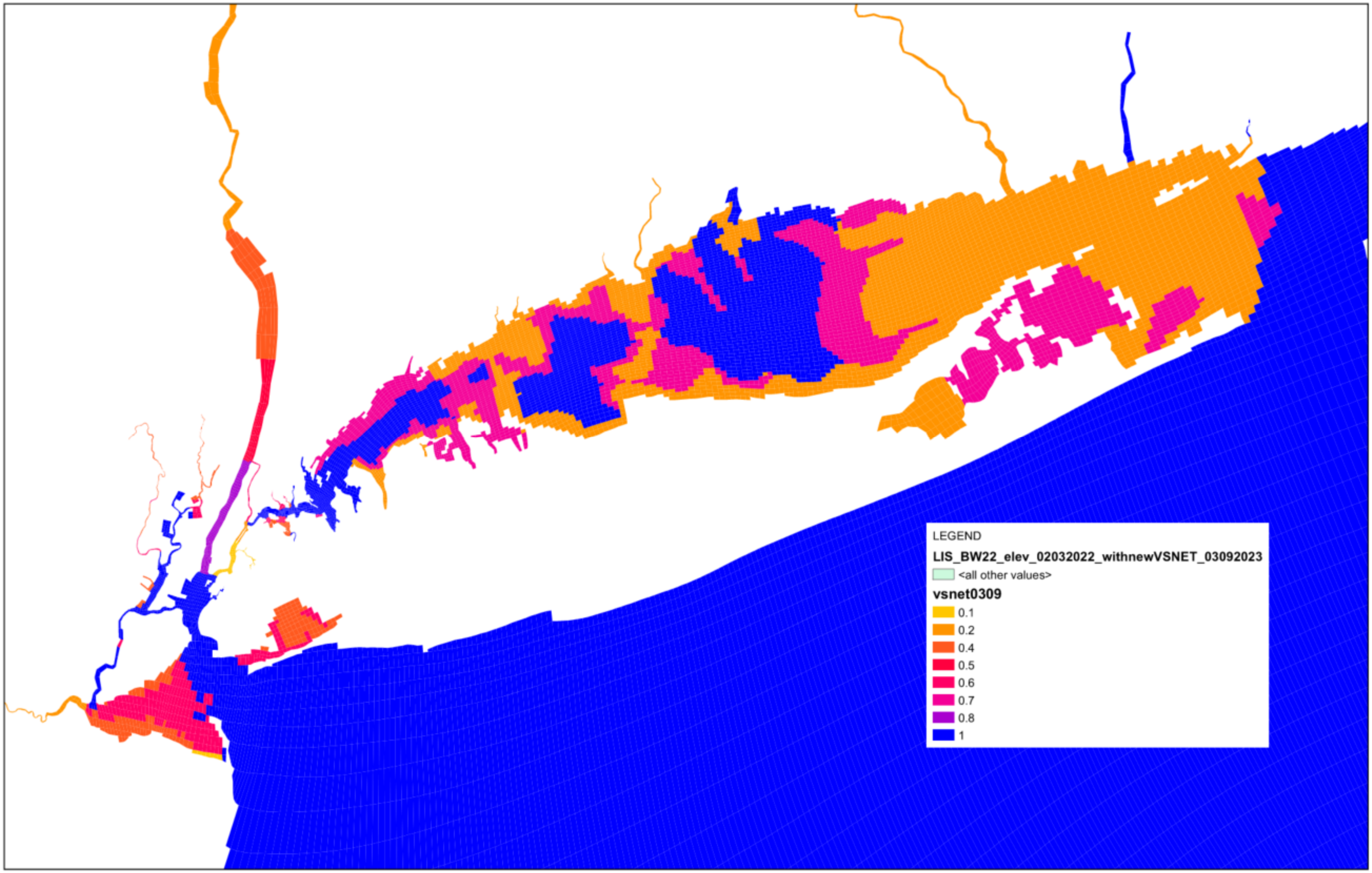

Figure 9. Assigned Vsnet

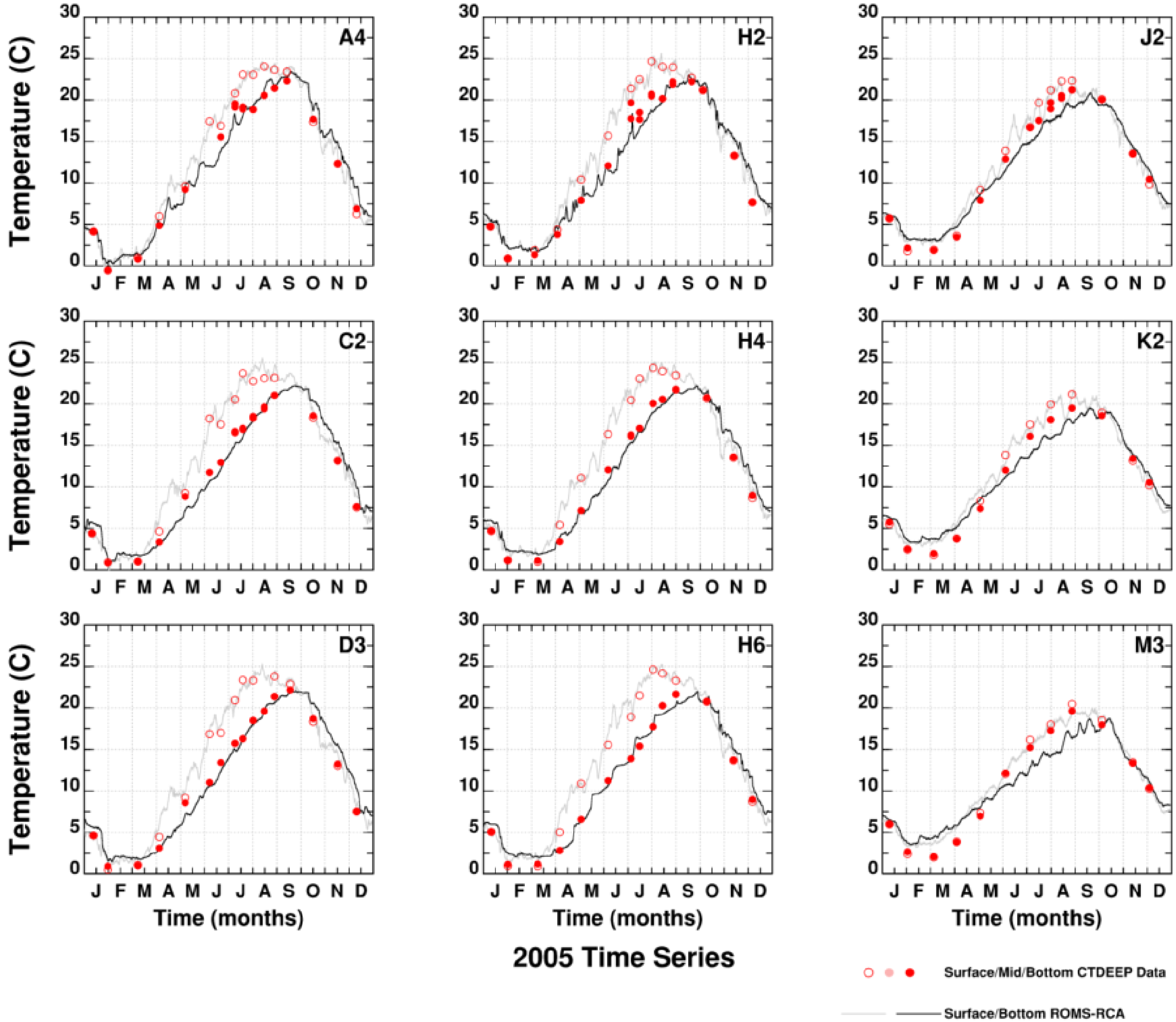

Figure 10. Model versus Data Comparisons for Temperature at Nine Representative CTDEEP Stations in 2005

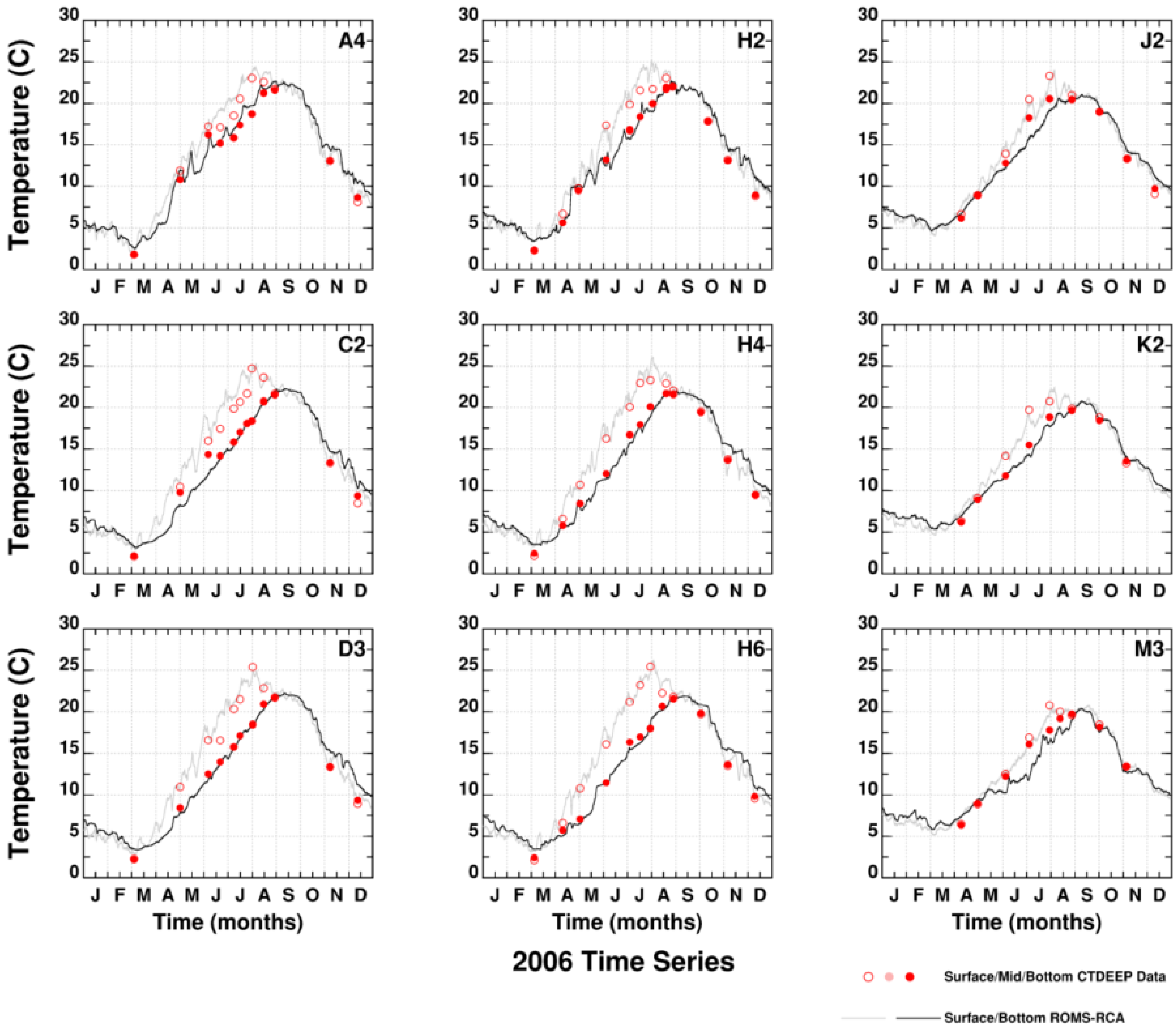

Figure 11. Model versus Data Comparisons for Temperature at Nine Representative CTDEEP Stations in 2006

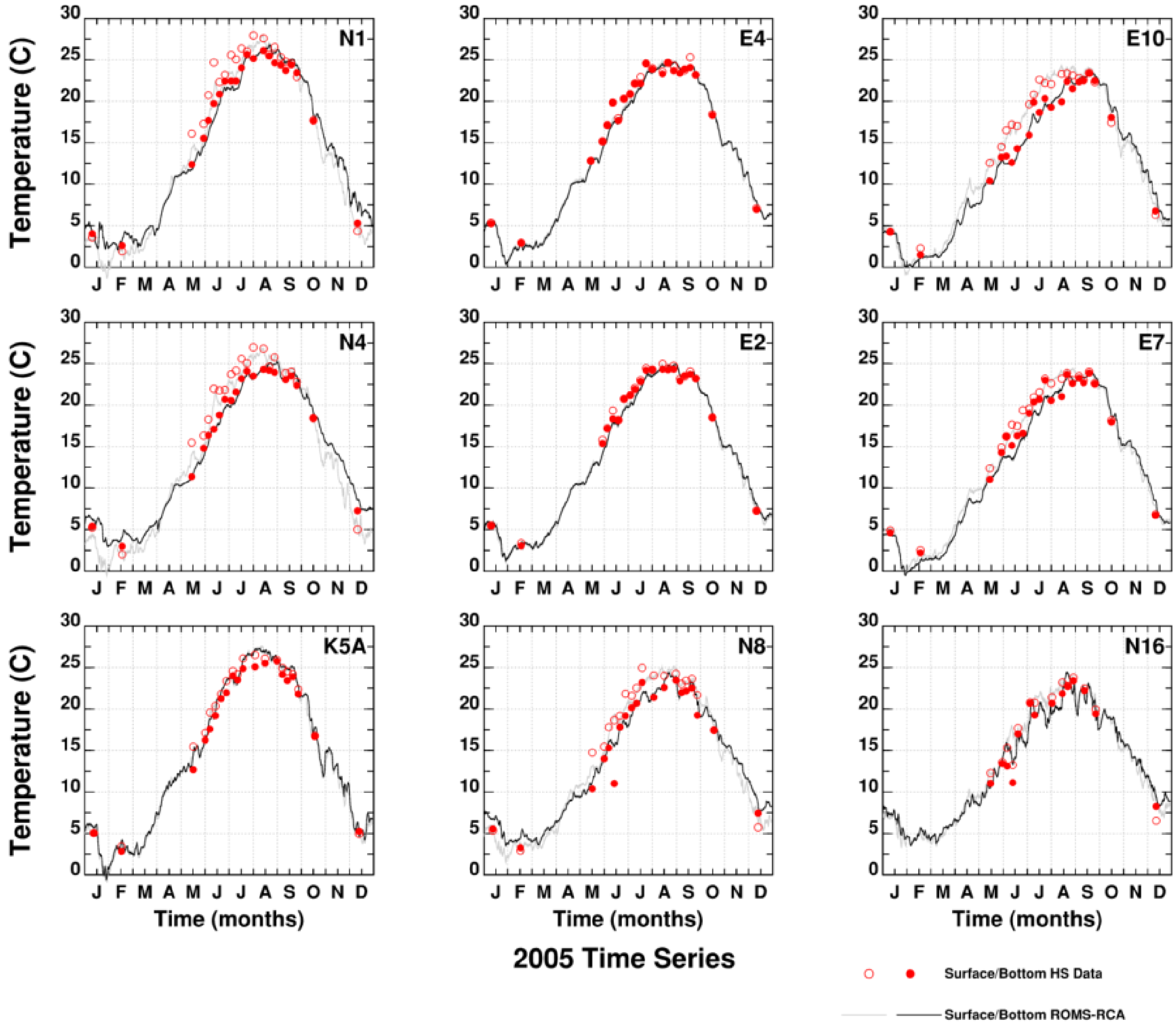

Figure 12. Model versus Data Comparisons for Temperature at Nine Representative DEP HS Stations in 2005

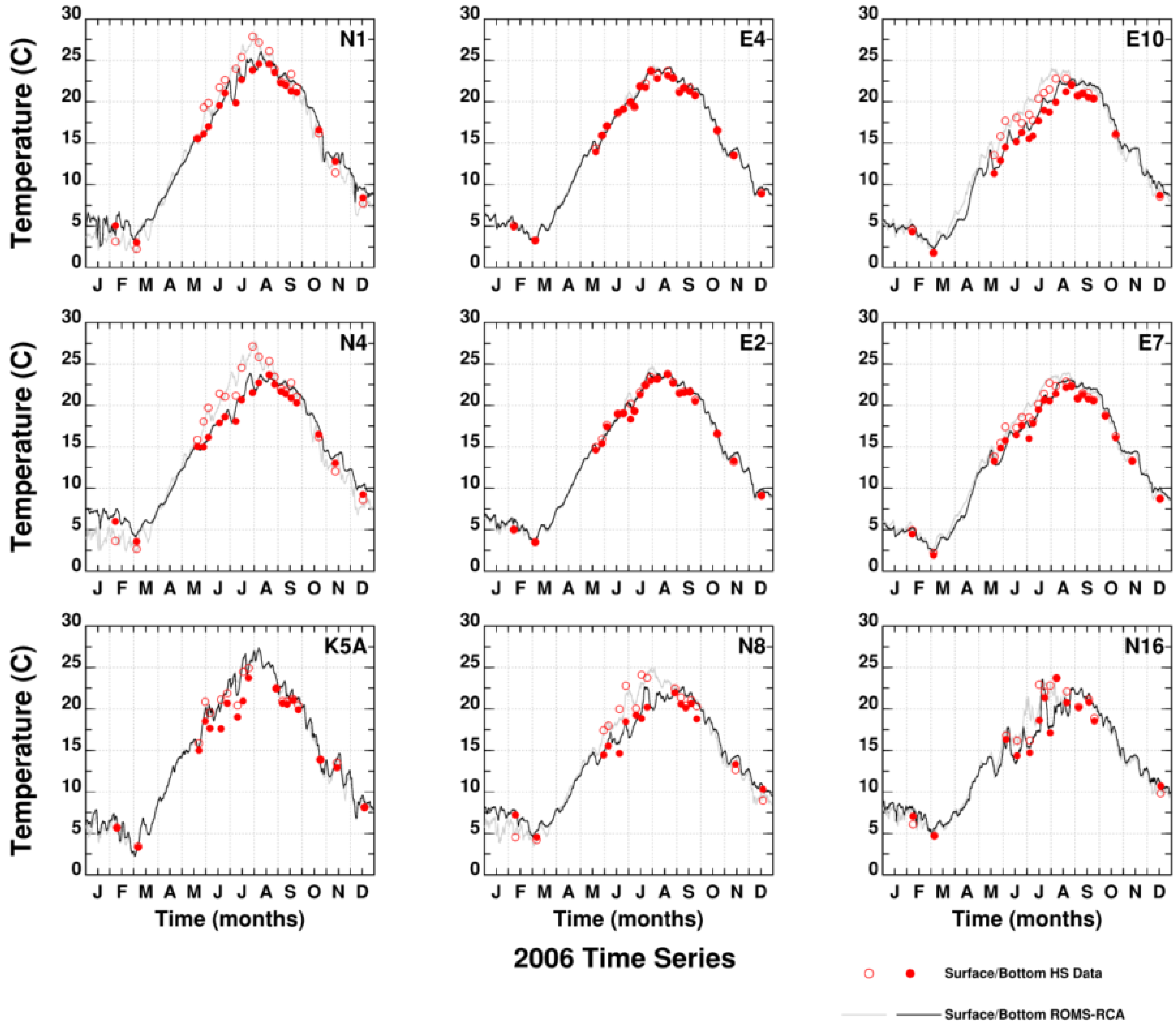

Figure 13. Model versus Data Comparisons for Temperature at Nine Representative DEP HS Stations in 2006

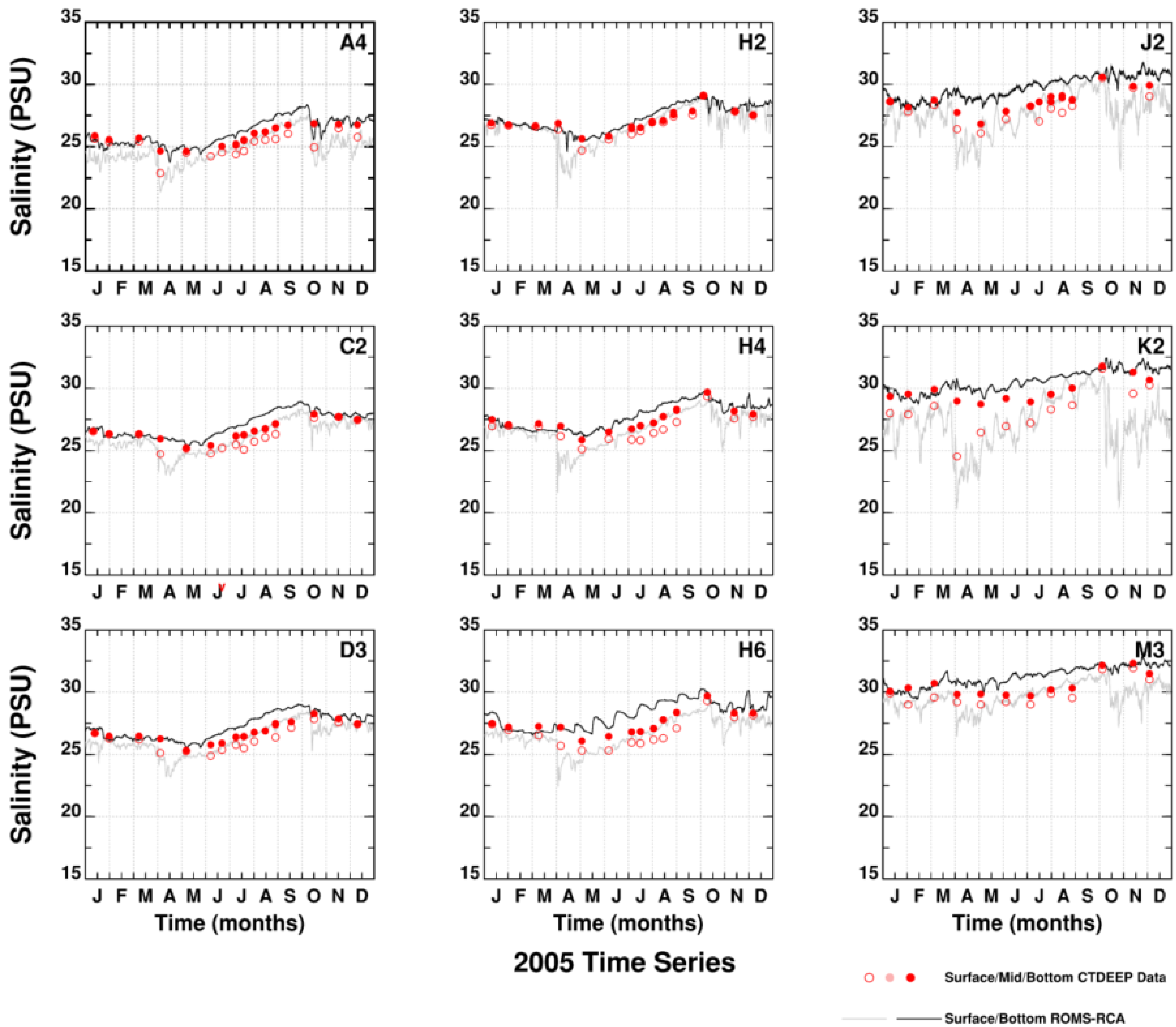

Figure 14. Model versus Data Comparisons for Salinity at Nine Representative CTDEEP Stations in 2005

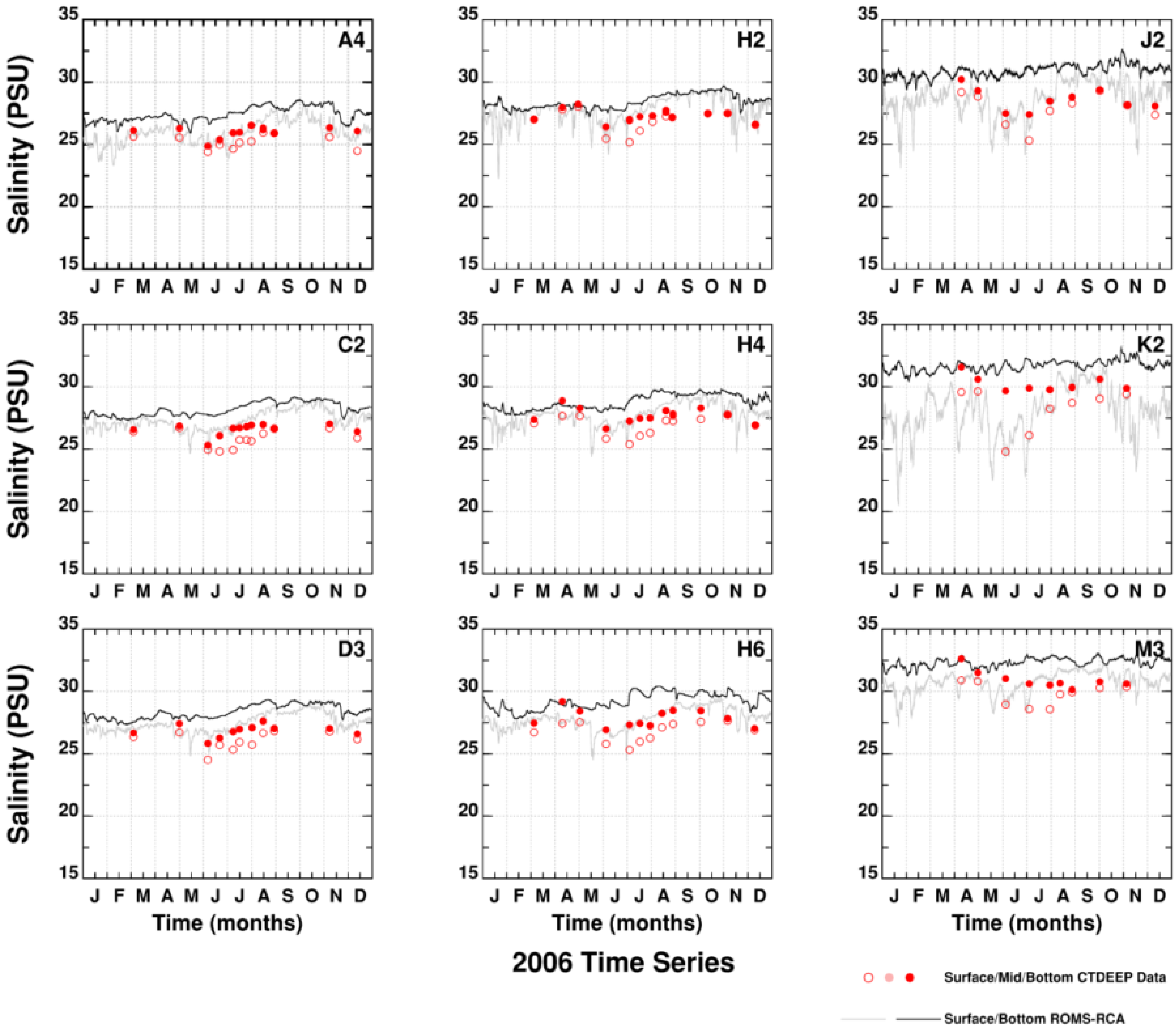

Figure 15. Model versus Data Comparisons for Salinity at Nine Representative CTDEEP Stations in 2006

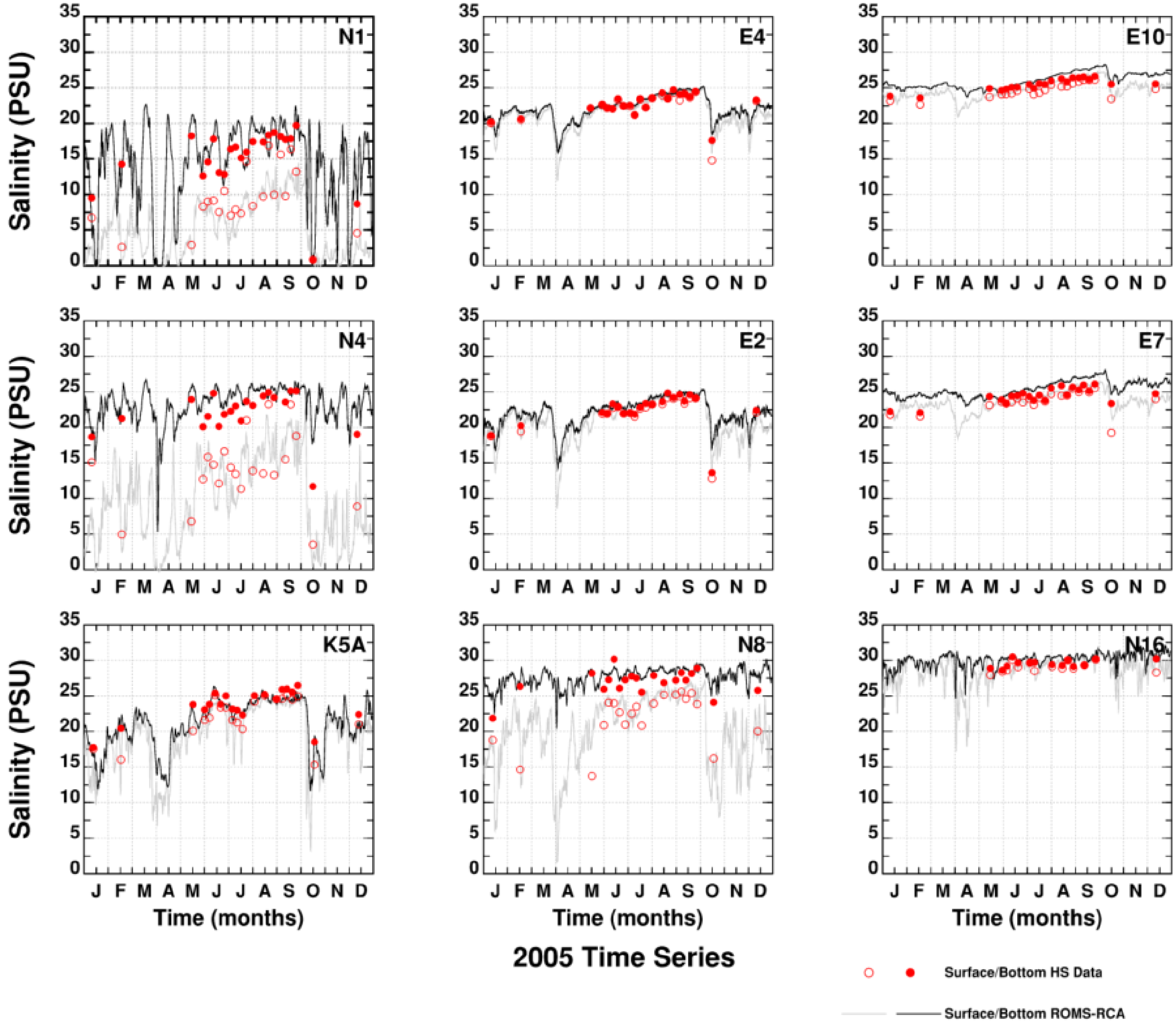

Figure 16. Model versus Data Comparisons for Salinity at Nine Representative DEP HS Stations in 2005

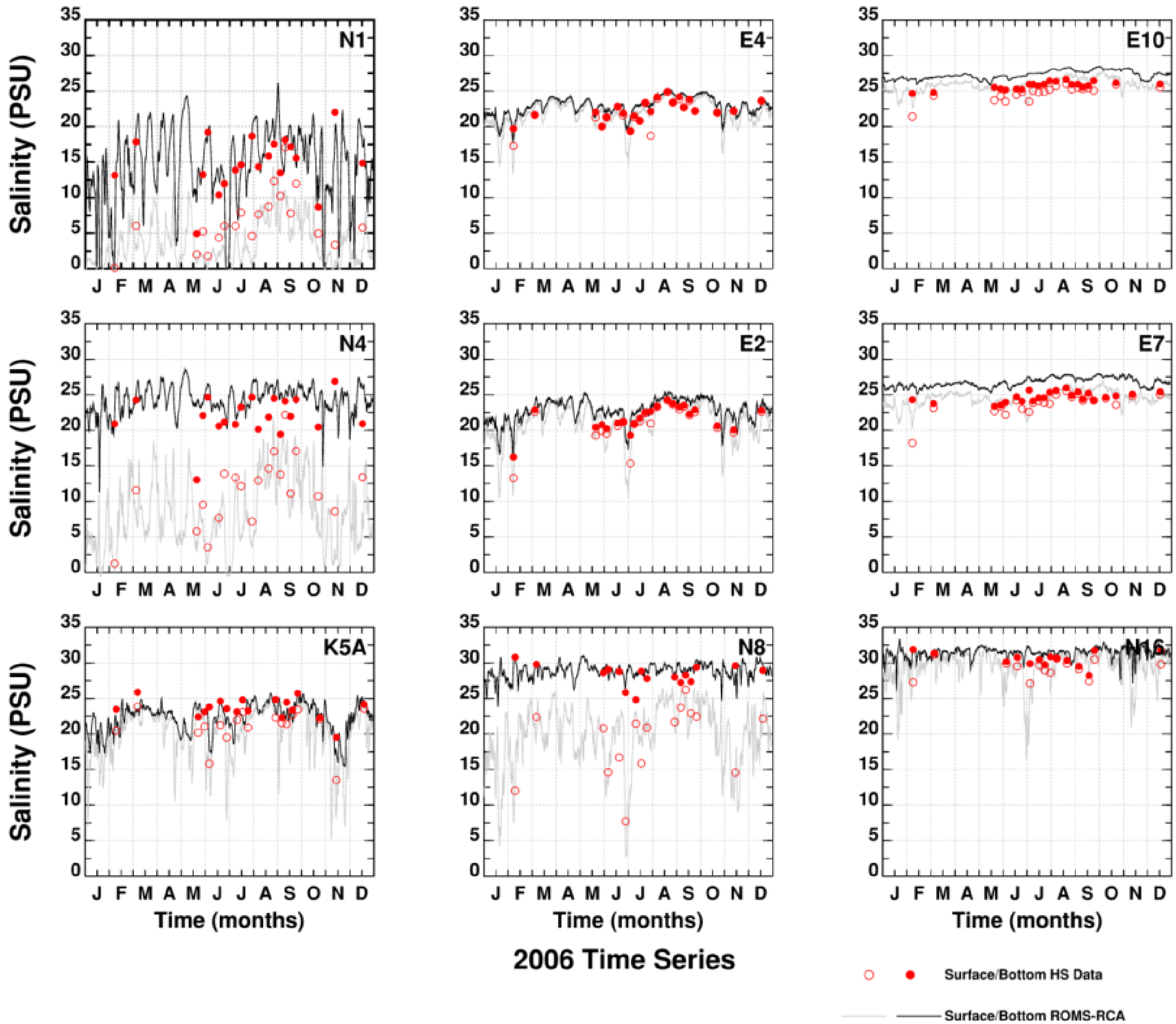

Figure 17. Model versus Data Comparisons for Salinity at Nine Representative DEP HS Stations in 2006

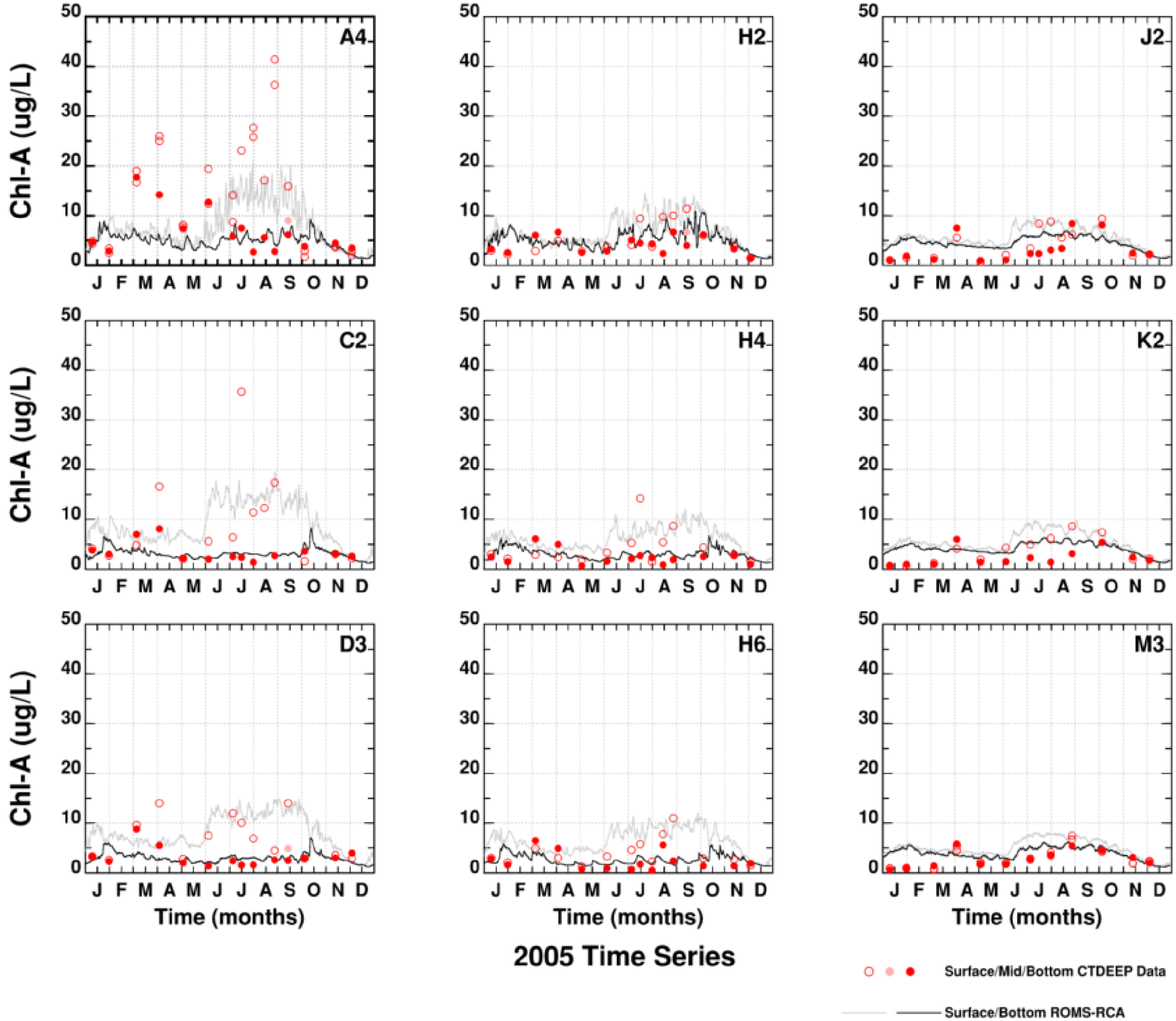

Figure 18. Model versus Data Comparisons for Chlorophyll-a at Nine Representative CTDEEP Stations in 2005

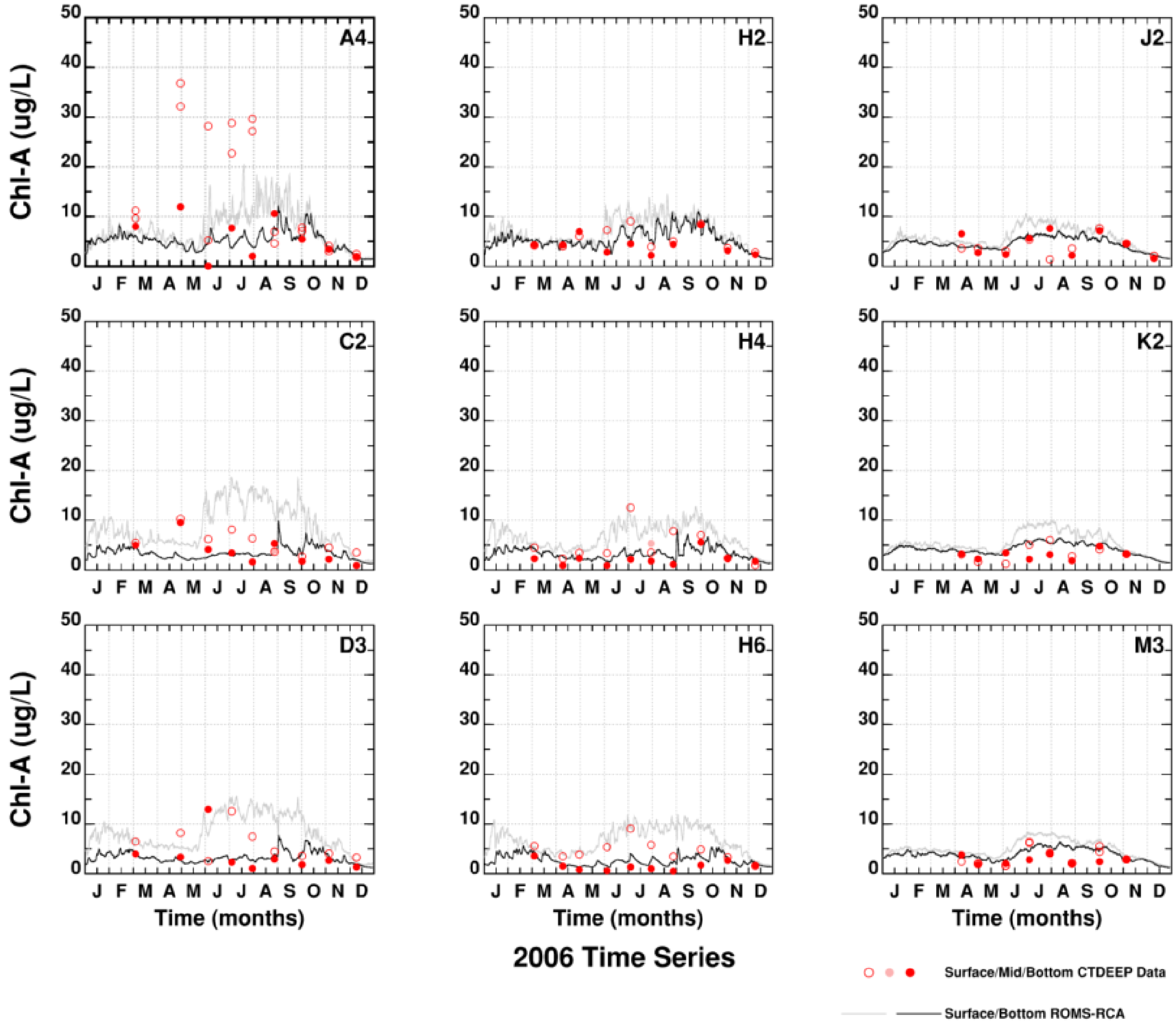

Figure 19. Model versus Data Comparisons for Chlorophyll-a at Nine Representative CTDEEP Stations in 2006

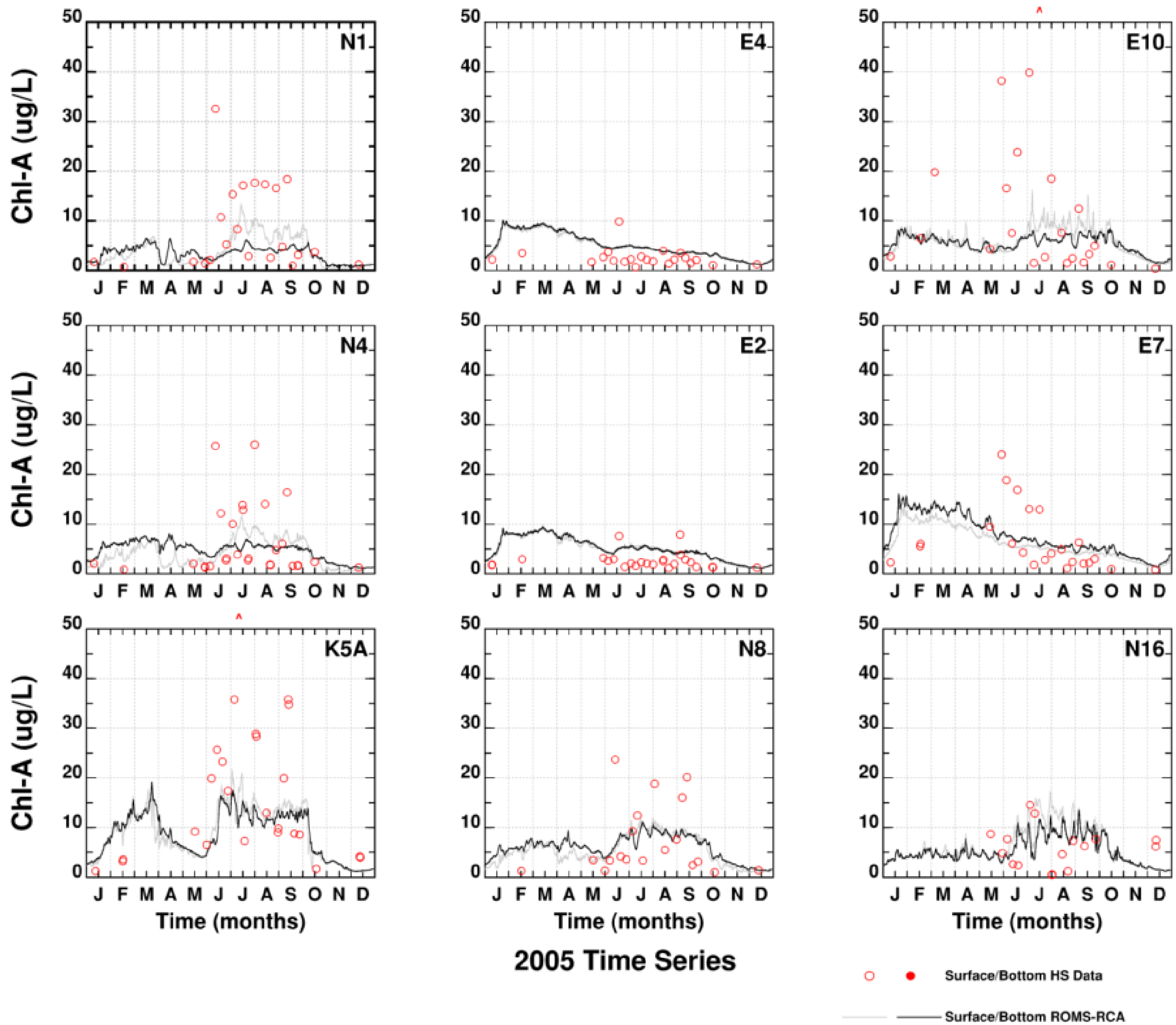

Figure 20. Model versus Data Comparisons for Chlorophyll-a at Nine Representative DEP HS Stations in 2005

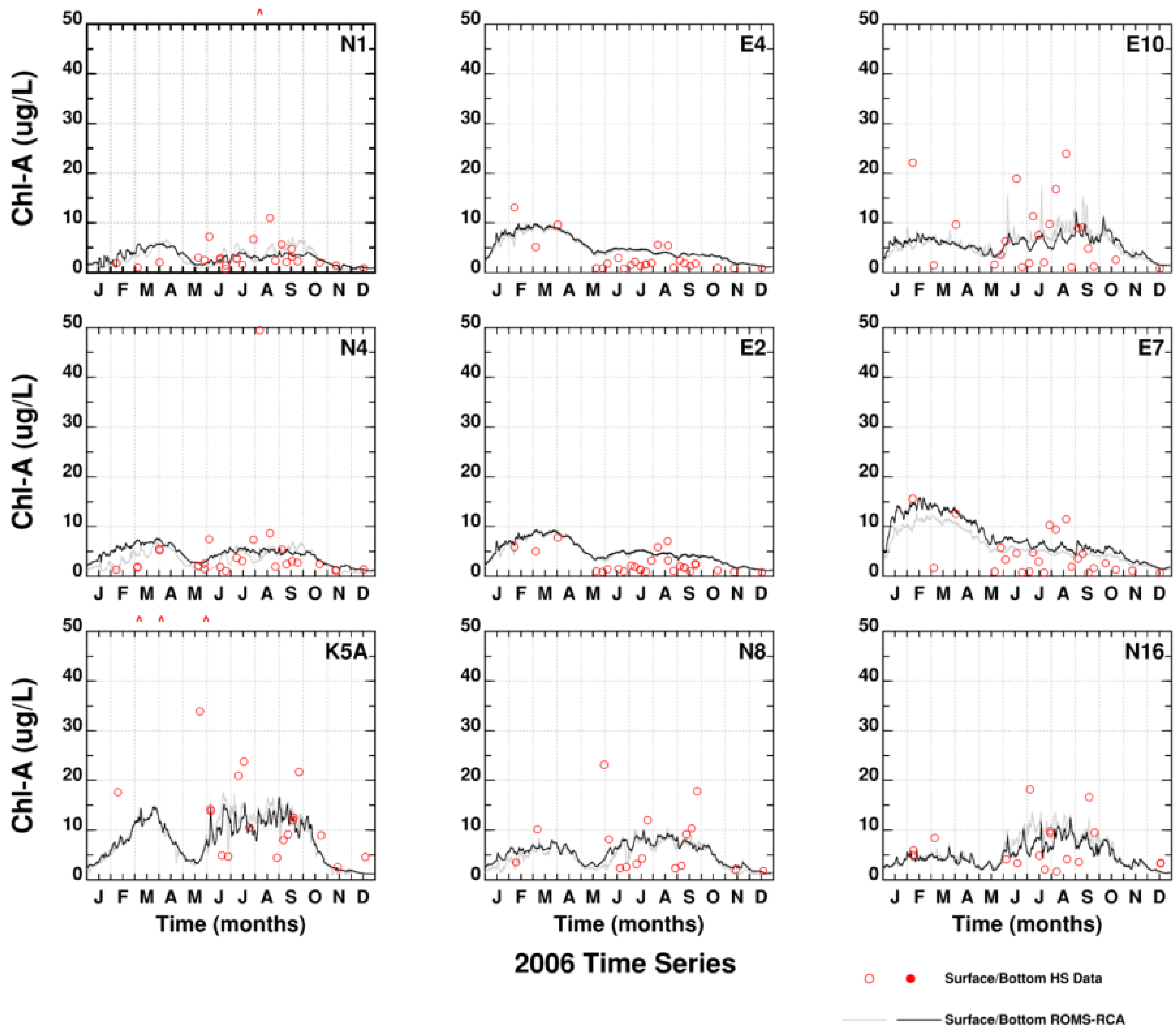

Figure 21. Model versus Data Comparisons for Chlorophyll-a at Nine Representative DEP HS Stations in 2006

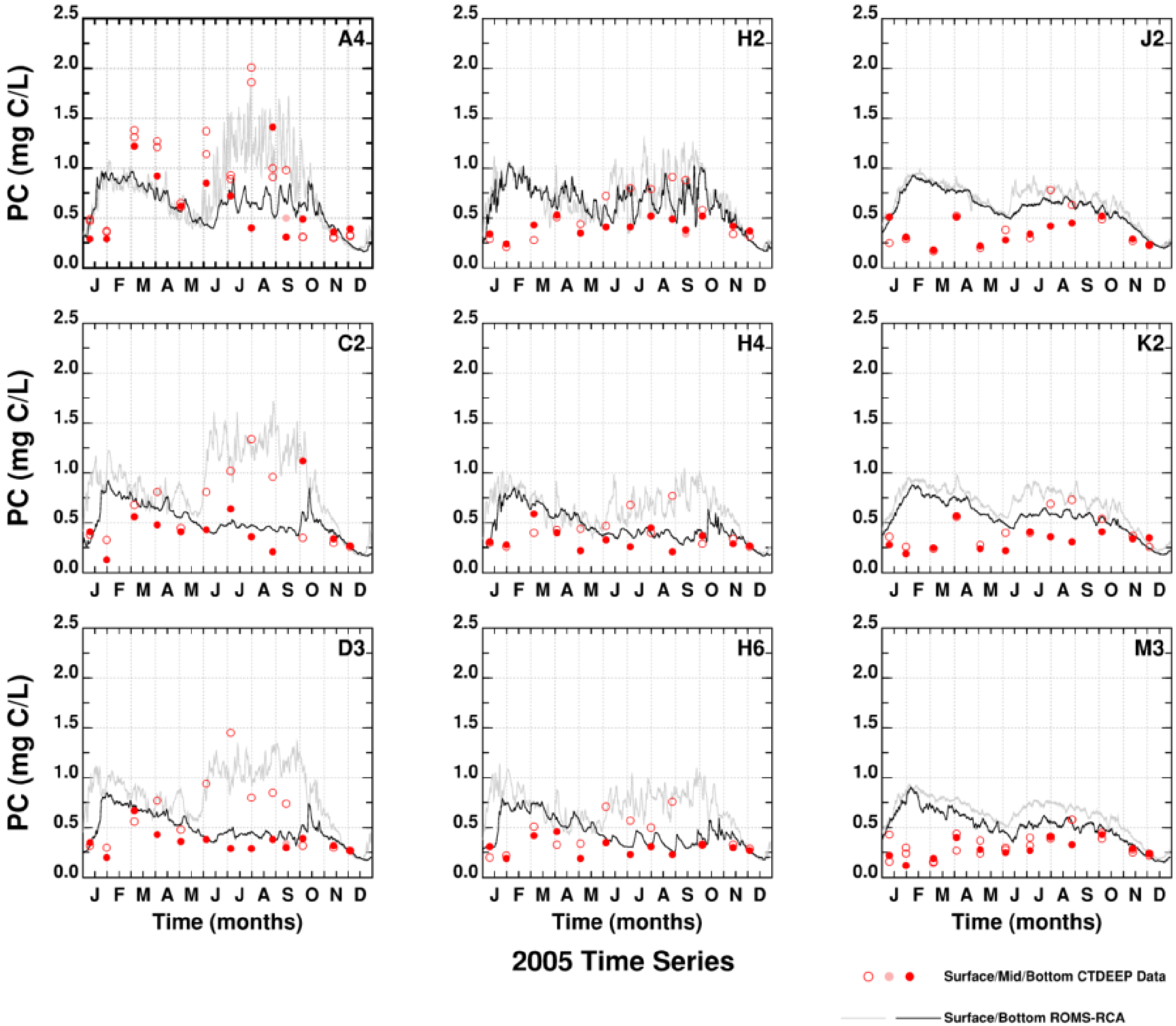

Figure 22. Model versus Data Comparisons for Particulate Carbon at Nine Representative CTDEEP Stations in 2005

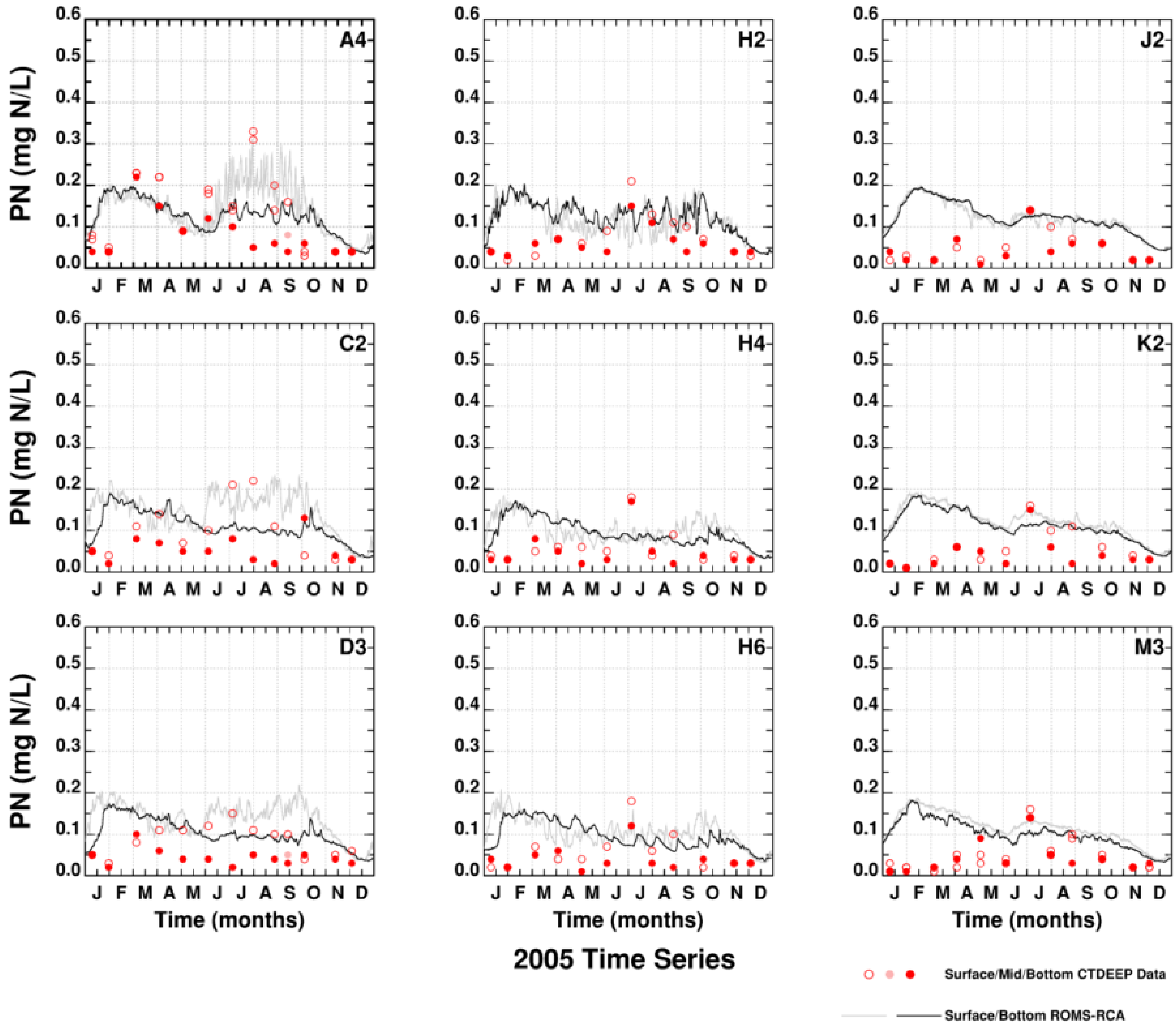

Figure 23. Model versus Data Comparisons for Particulate Nitrogen at Nine Representative CTDEEP Stations in 2005

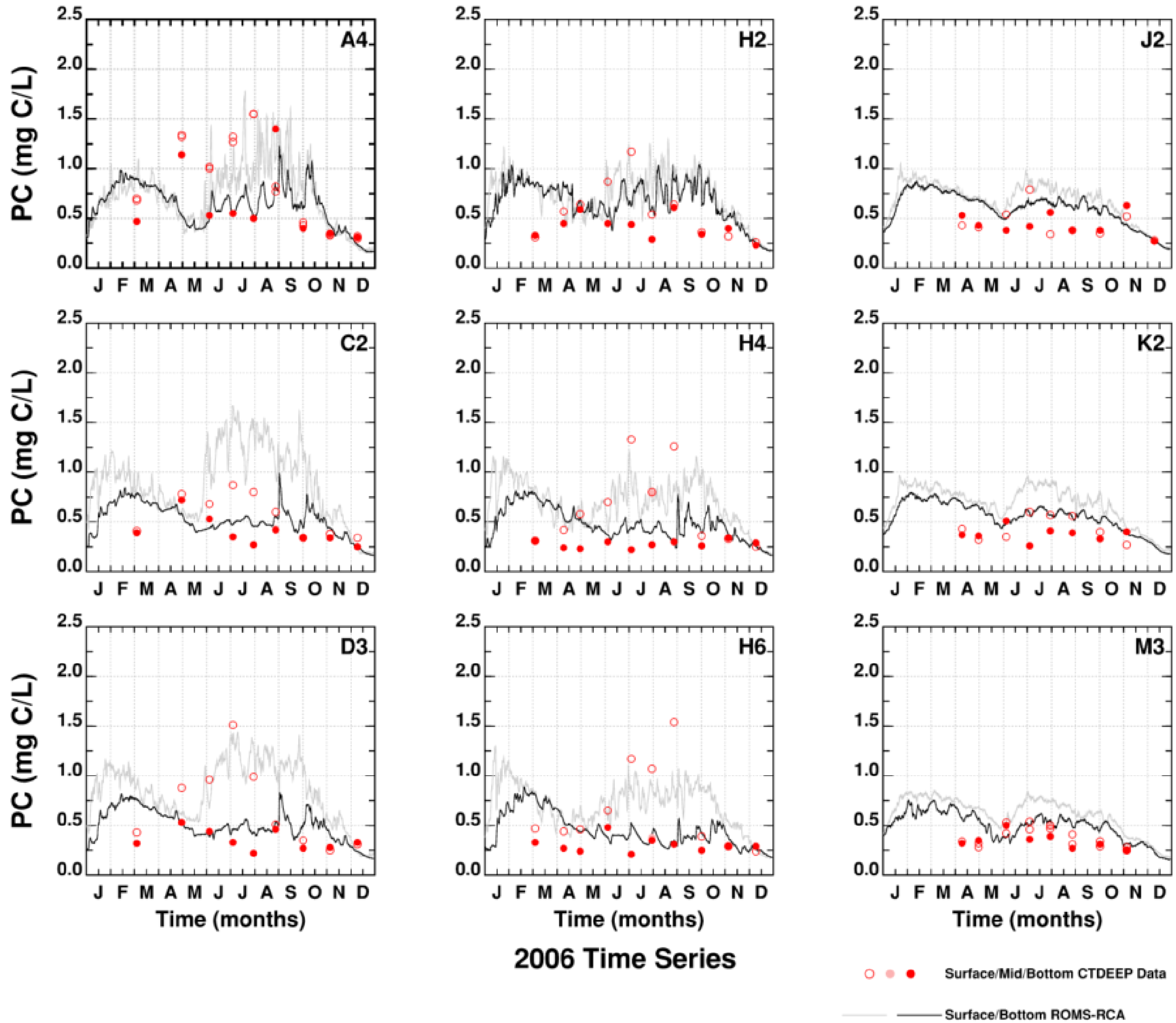

Figure 24. Model versus Data Comparisons for Particulate Carbon at Nine Representative CTDEEP Stations in 2006

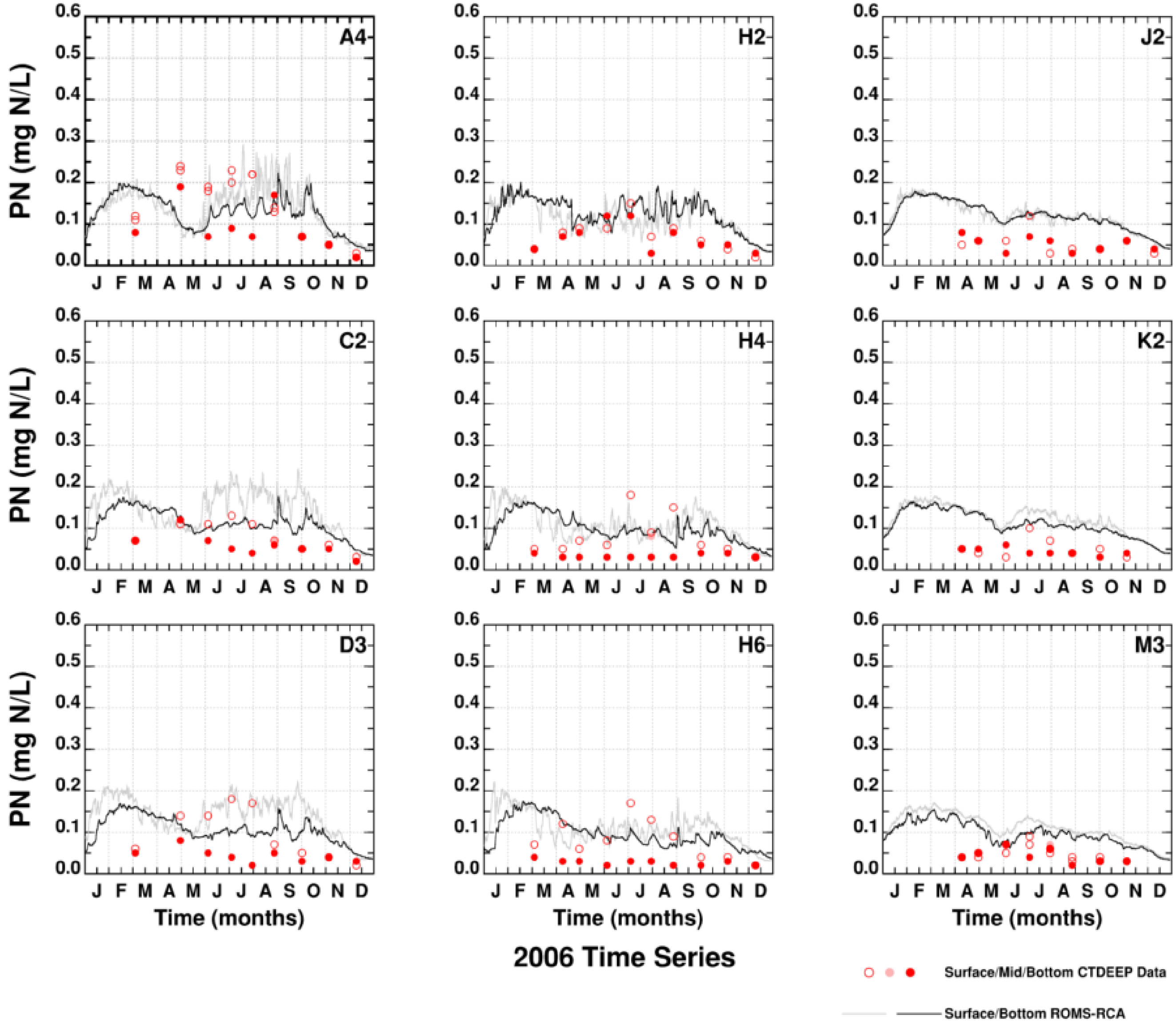

Figure 25. Model versus Data Comparisons for Particulate Nitrogen at Nine Representative CTDEEP Stations in 2006

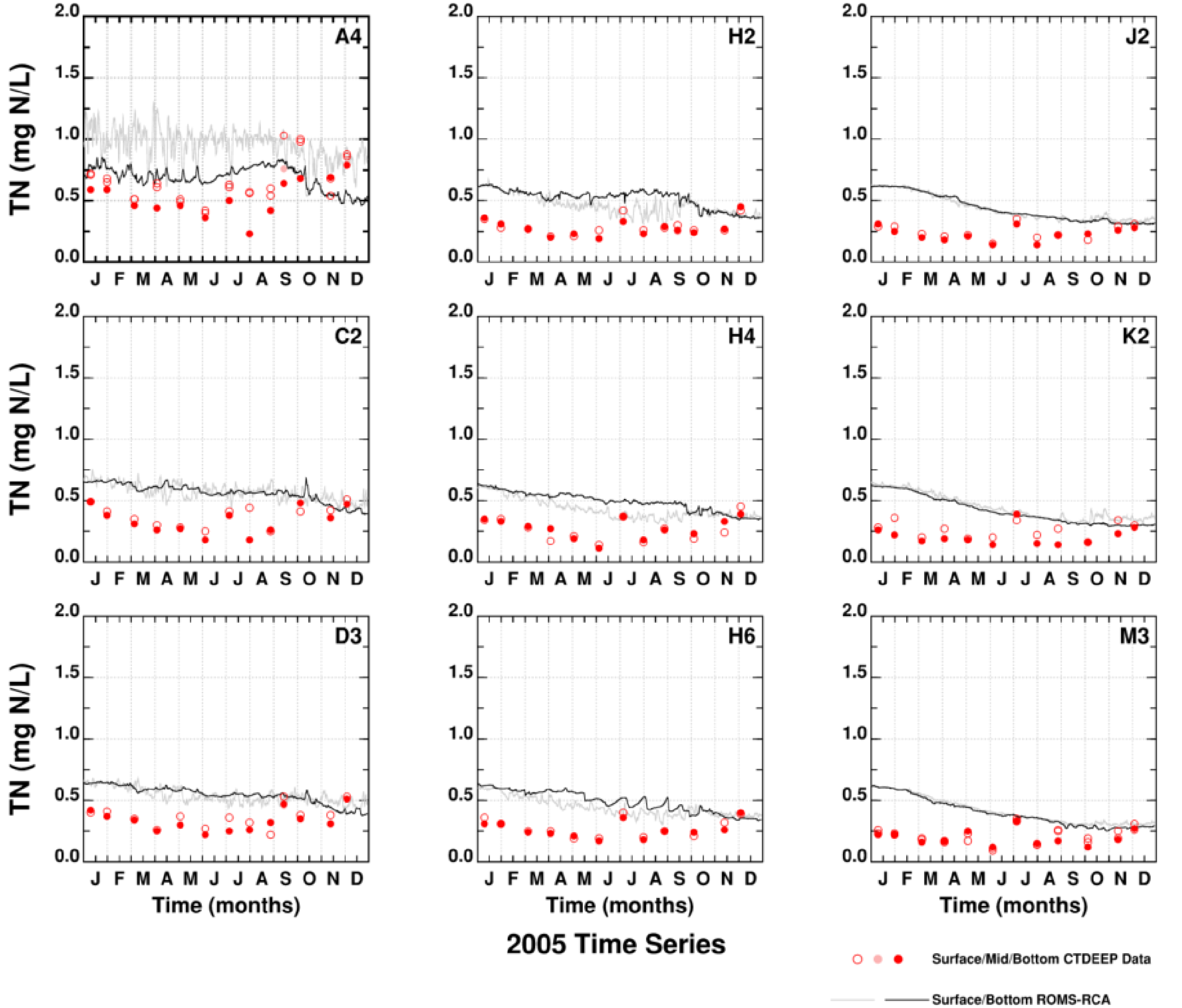

Figure 26. Model versus Data Comparisons for Total Nitrogen at Nine Representative CTDEEP Stations in 2005

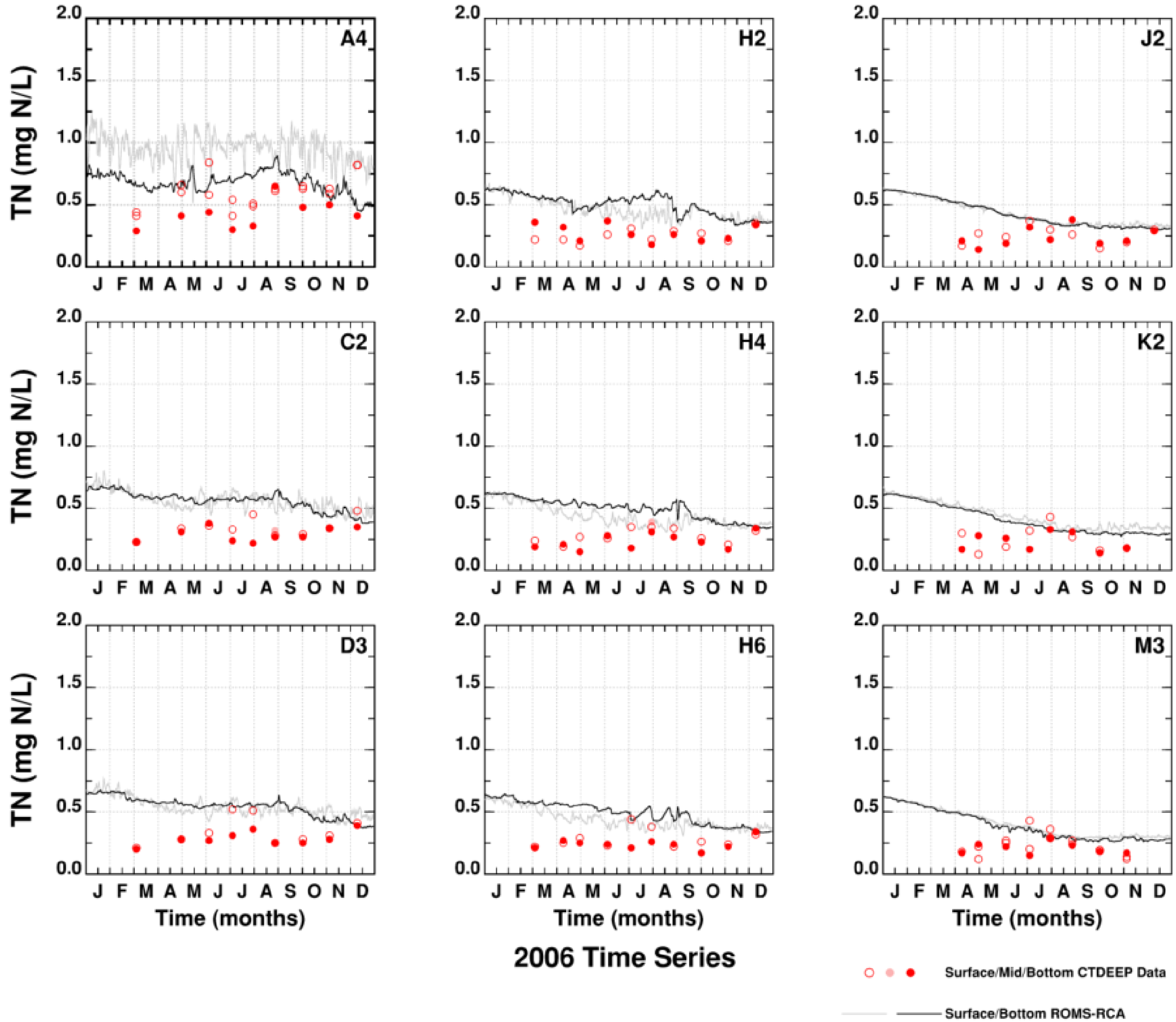

Figure 27. Model versus Data Comparisons for Total Nitrogen at Nine Representative CTDEEP Stations in 2006

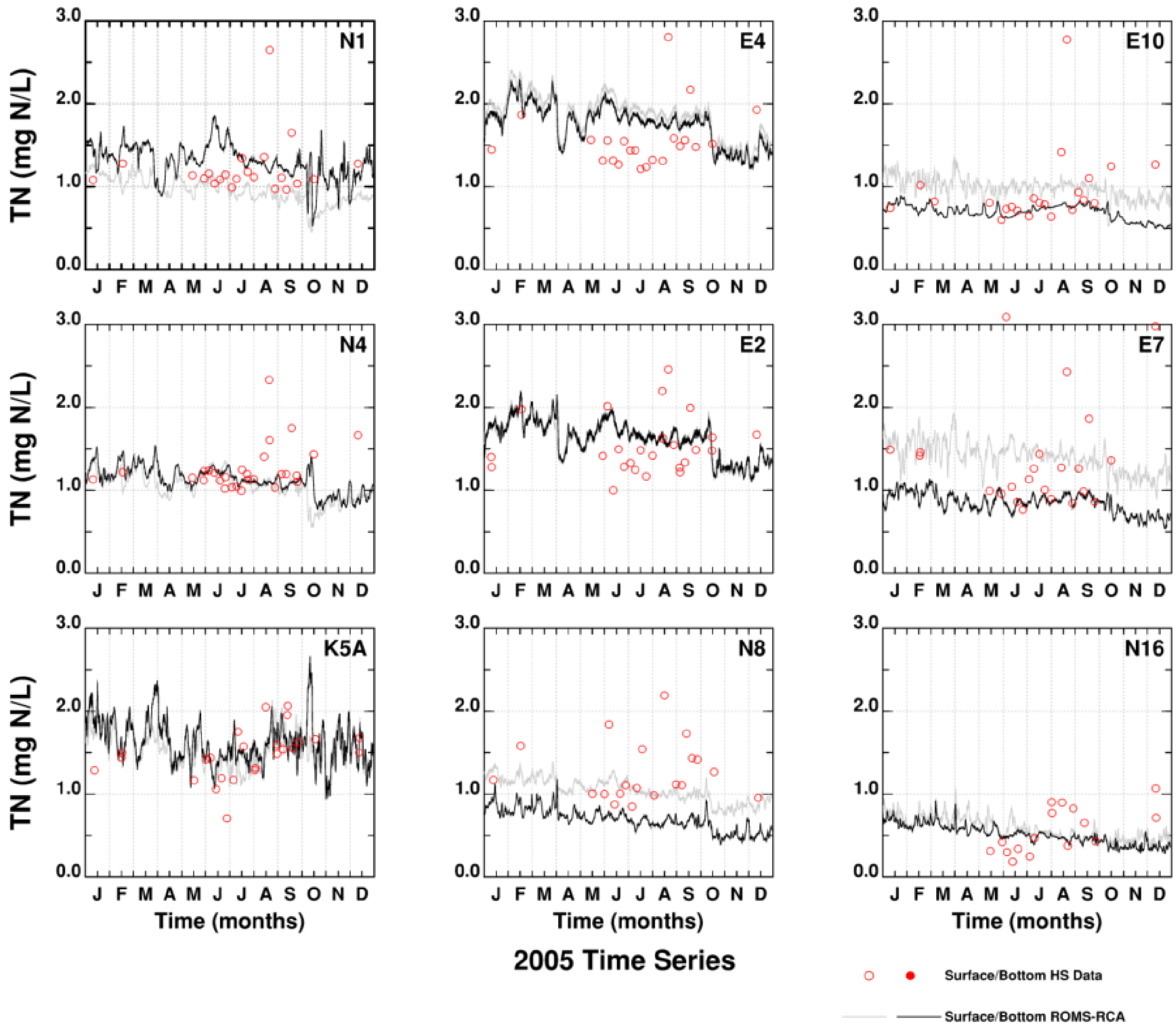

Figure 28. Model versus Data Comparisons for Total Nitrogen at Nine Representative DEP HS Stations in 2005

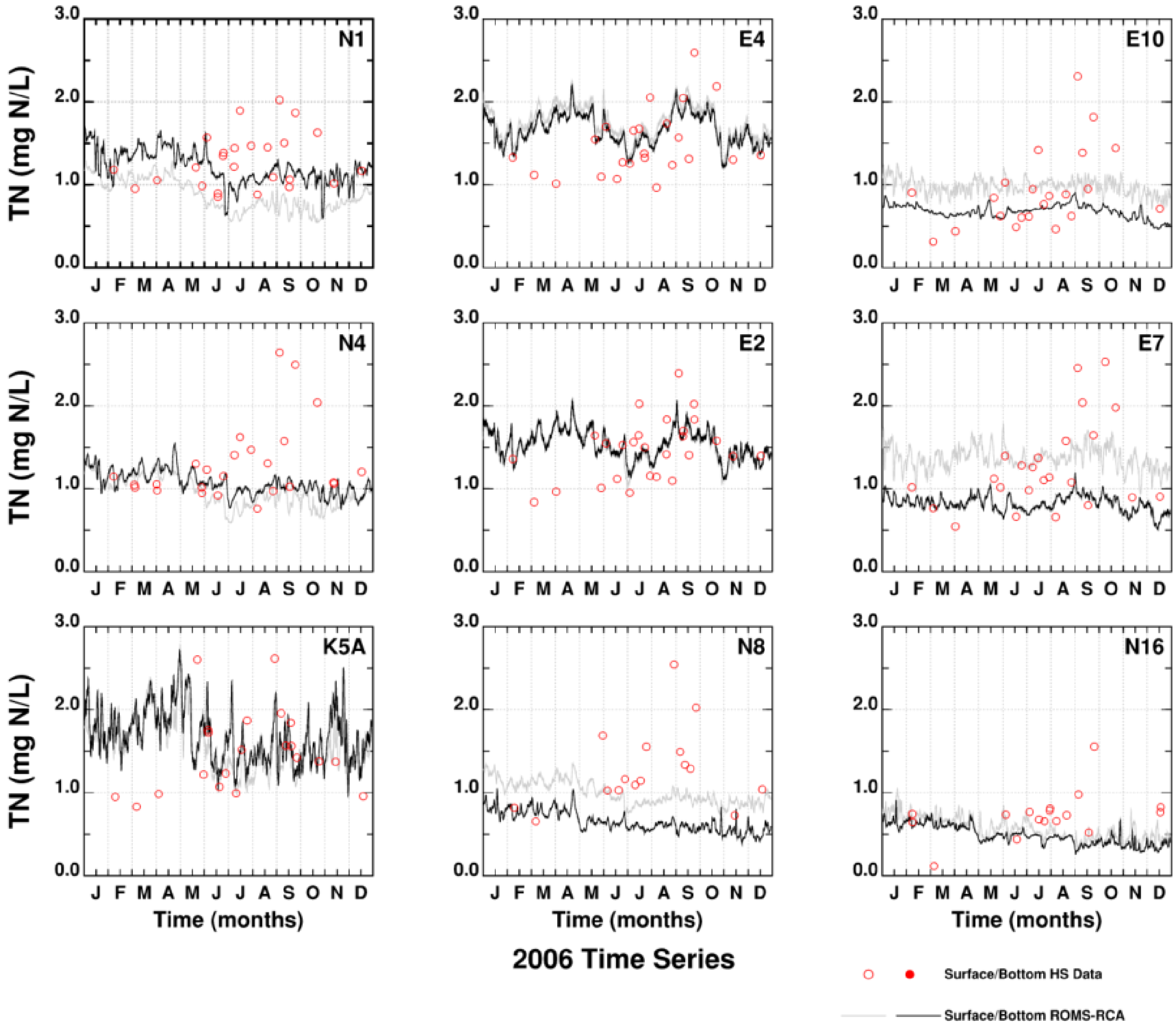

Figure 29. Model versus Data Comparisons for Total Nitrogen at Nine Representative DEP HS Stations in 2006

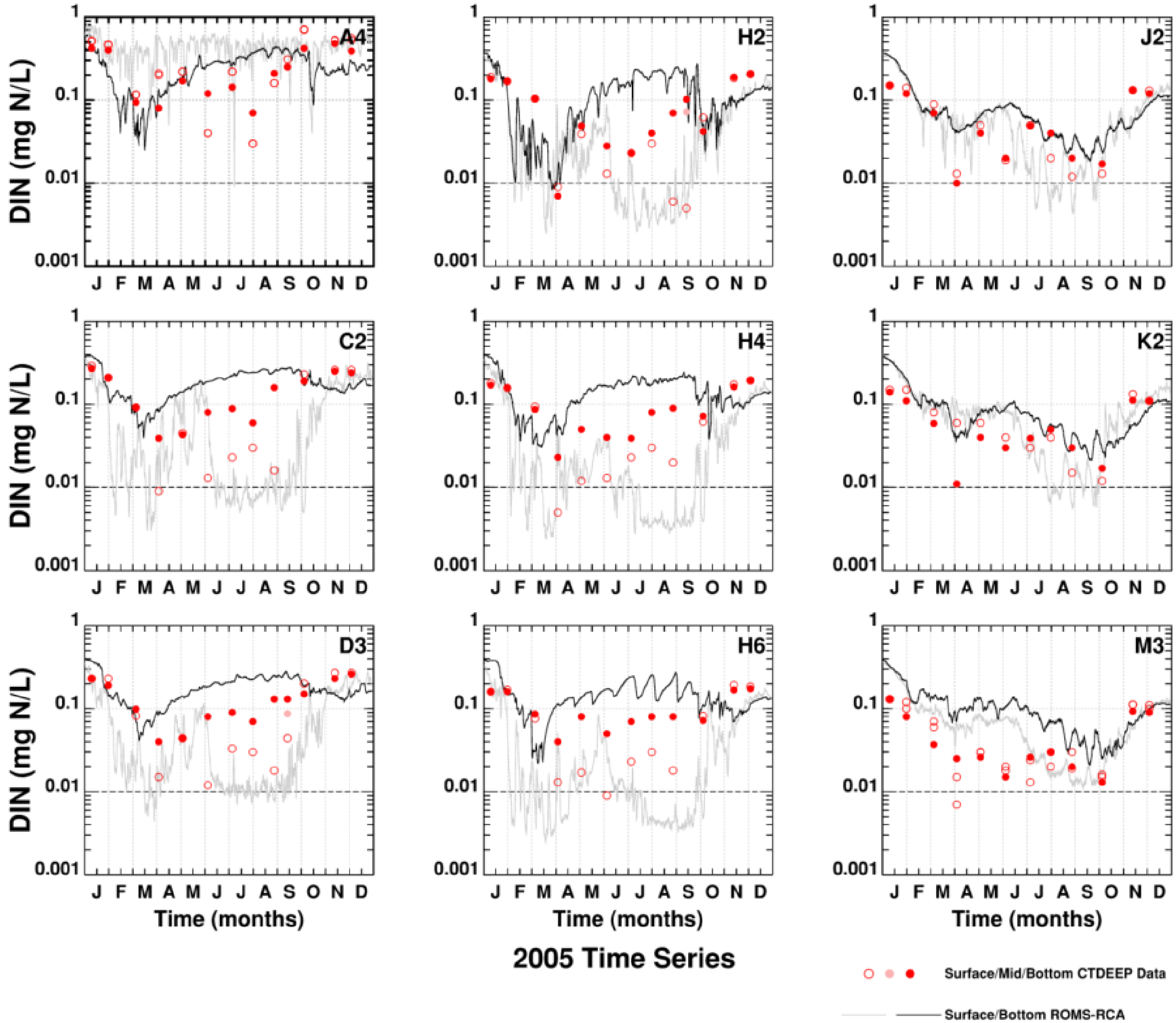

Figure 30. Model versus Data Comparisons for Dissolved Inorganic Nitrogen at Nine Representative CTDEEP Stations in 2005

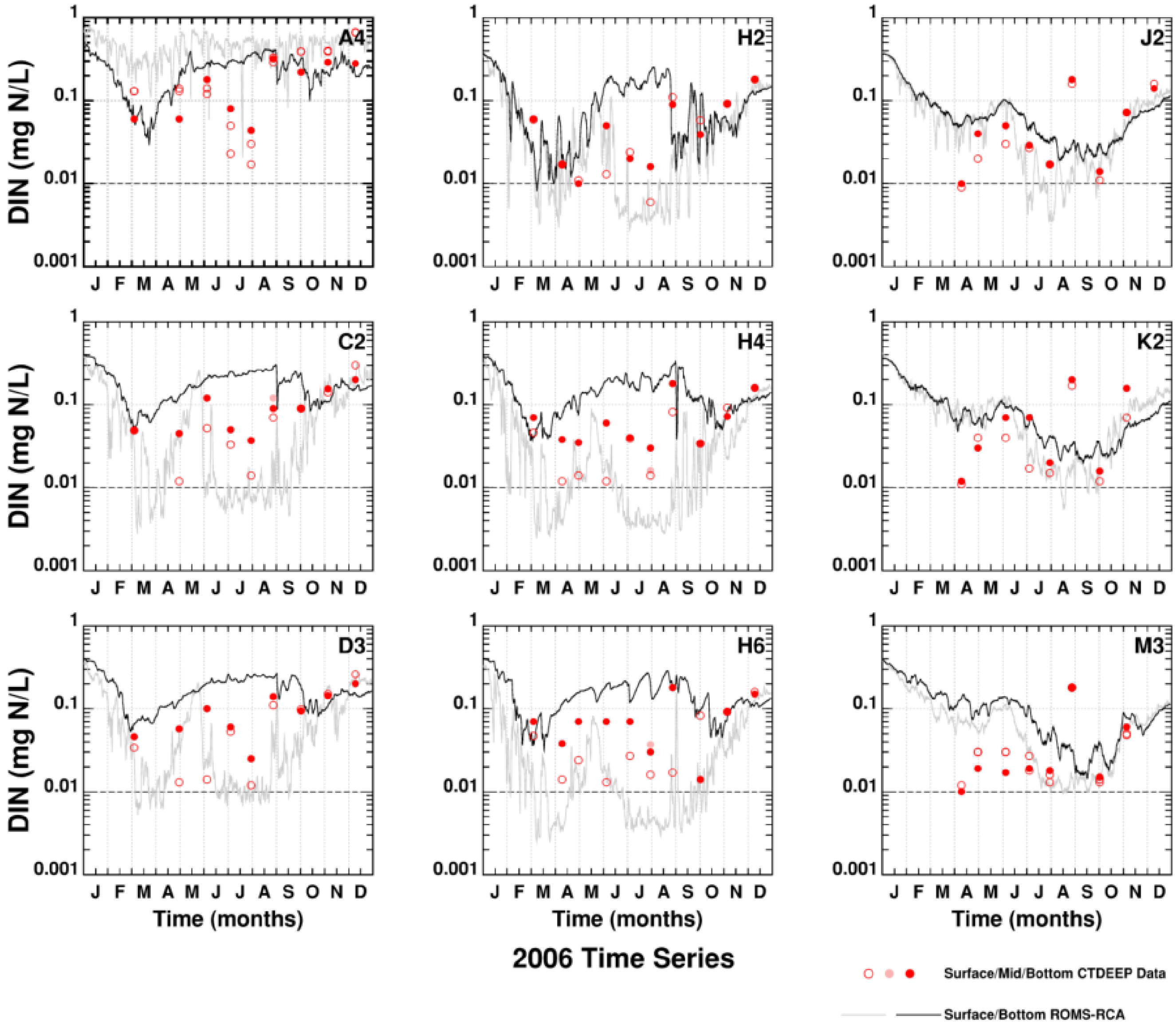

Figure 31. Model versus Data Comparisons for Dissolved Inorganic Nitrogen at Nine Representative CTDEEP Stations in 2006

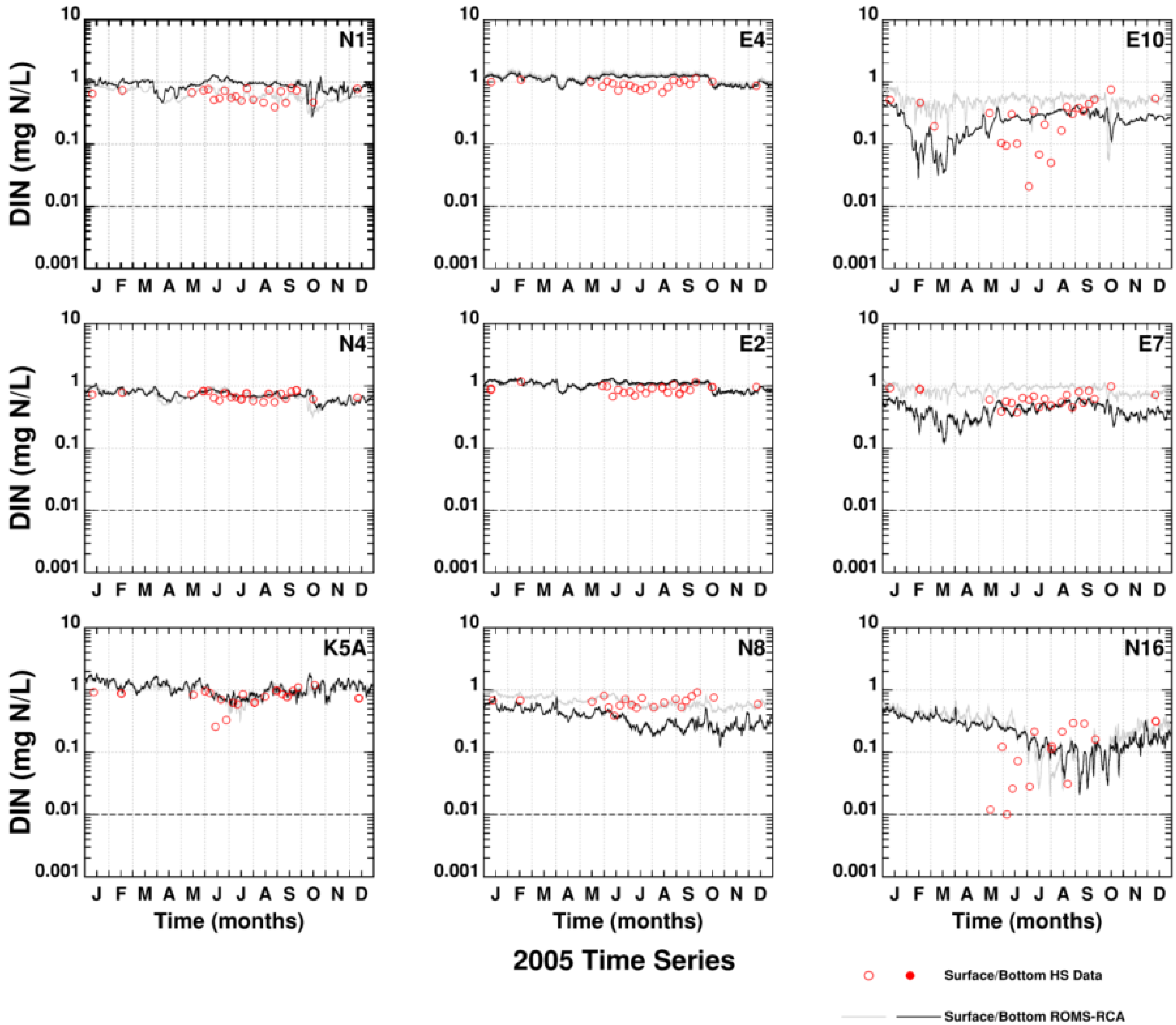

Figure 32. Model versus Data Comparisons for Dissolved Inorganic Nitrogen at Nine Representative DEP HS Stations in 2005

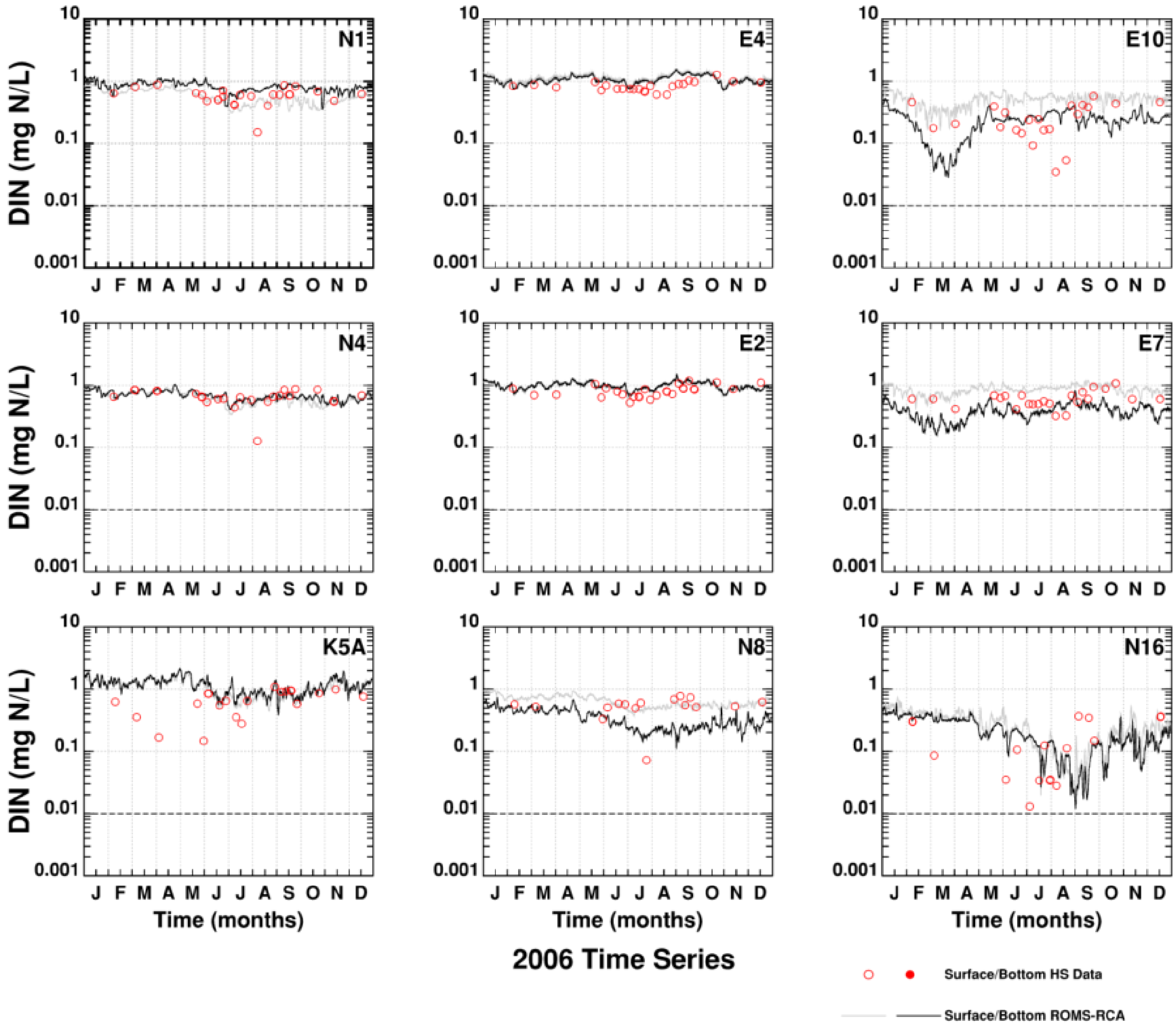

Figure 33. Model versus Data Comparisons for Dissolved Inorganic Nitrogen at Nine Representative DEP HS Stations in 2006

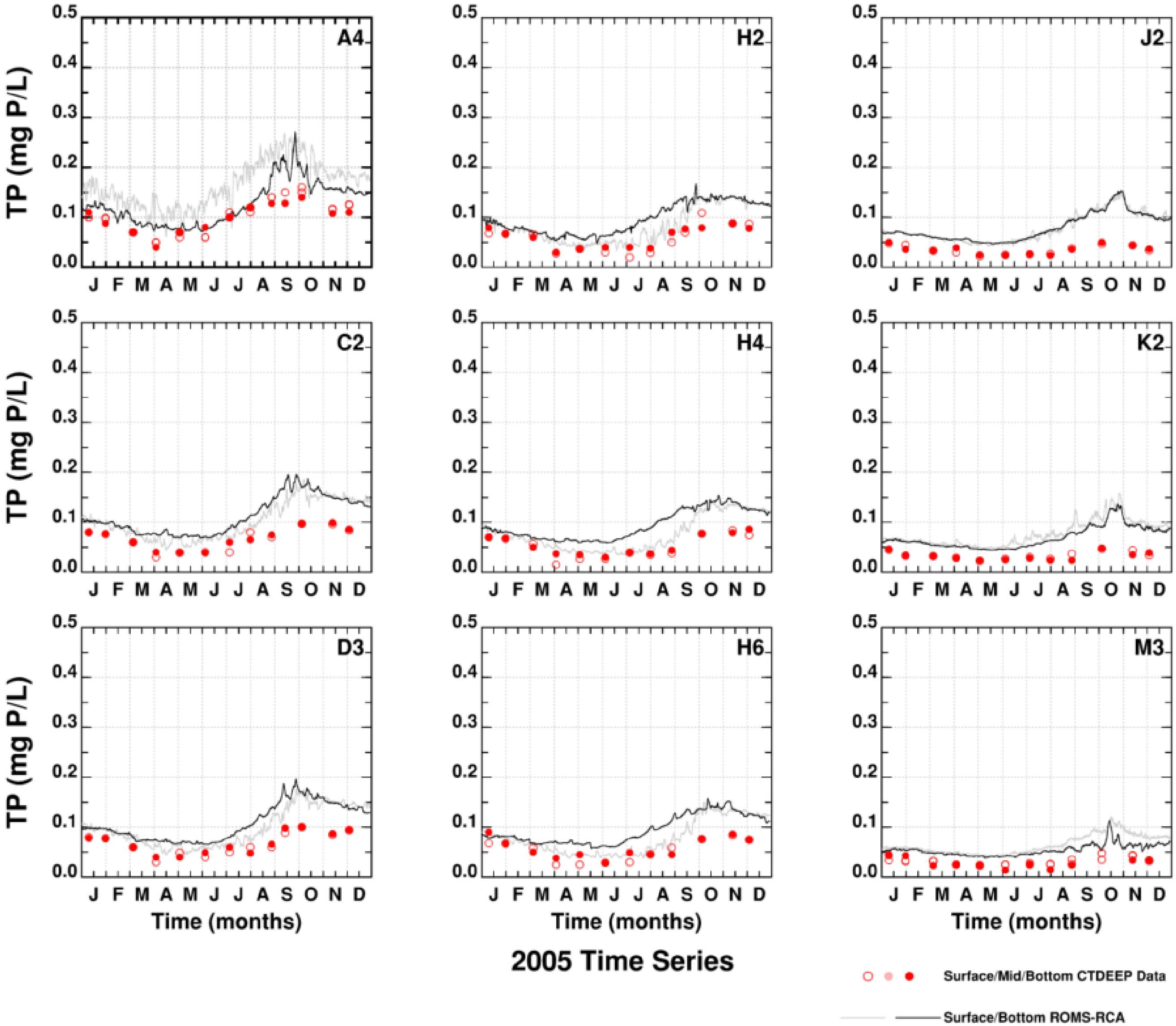

Figure 34. Model versus Data Comparisons for Total Phosphorus at Nine Representative CTDEEP Stations in 2005

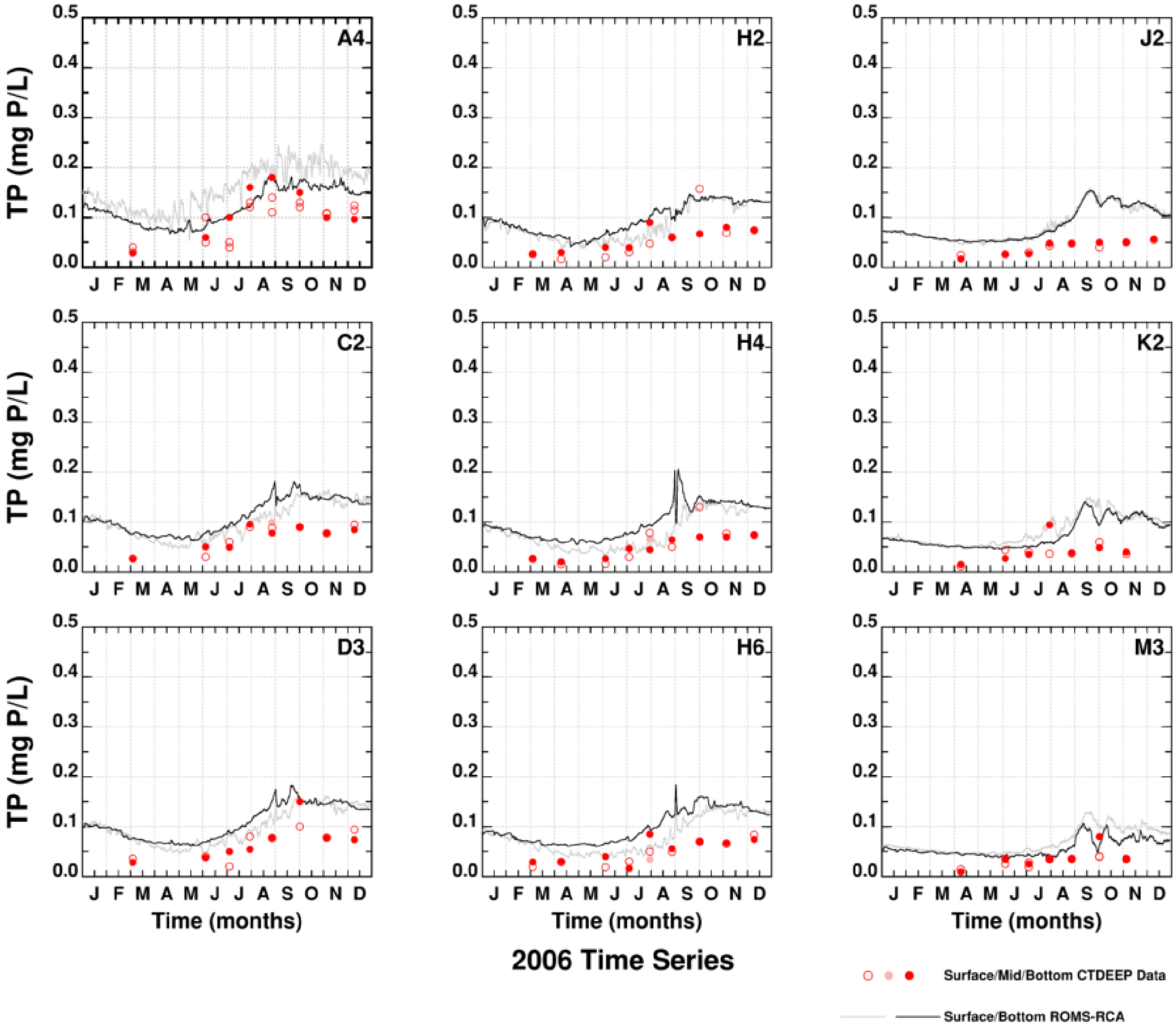

Figure 35. Model versus Data Comparisons for Total Phosphorus at Nine Representative CTDEEP Stations in 2006

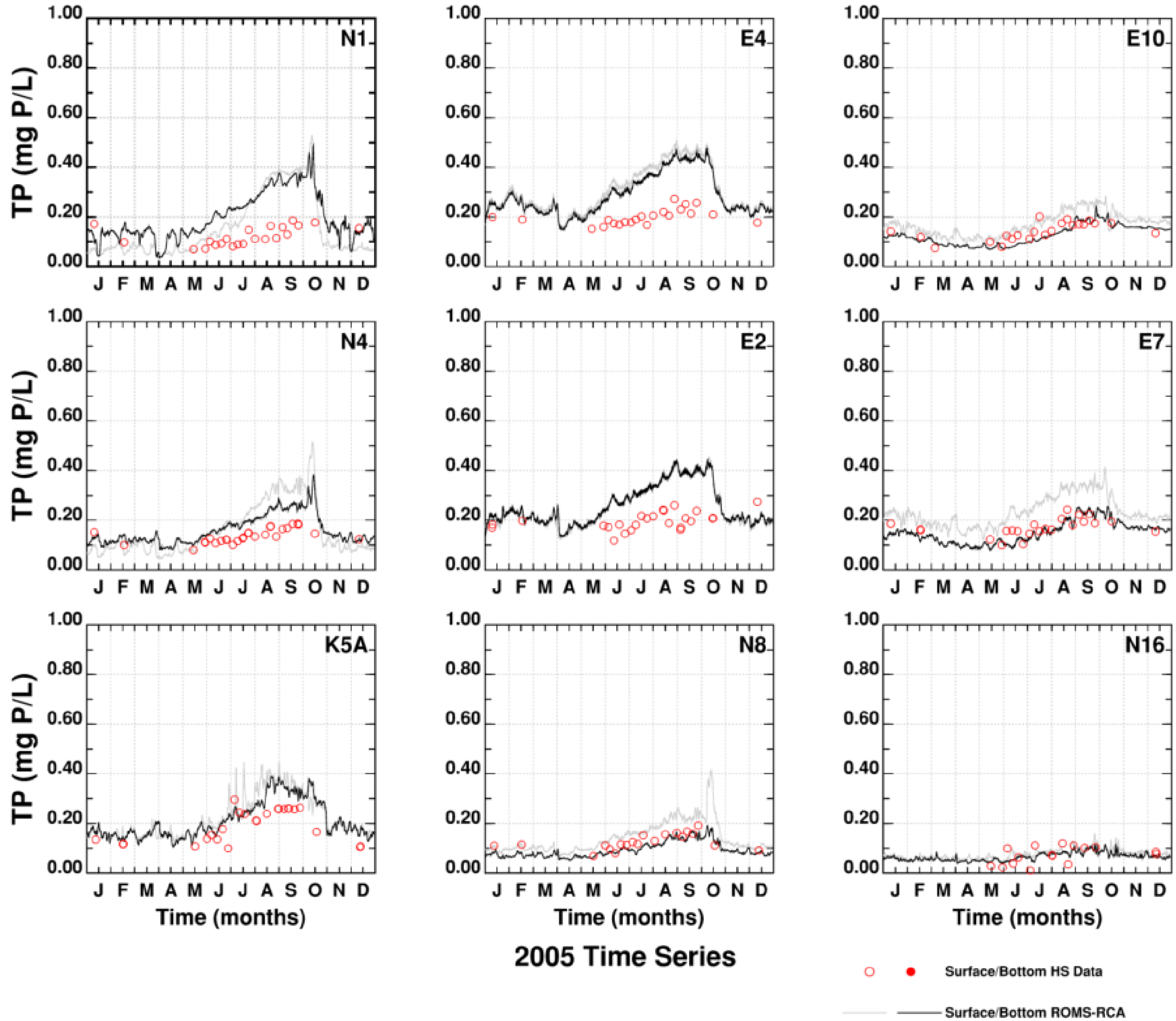

Figure 36. Model versus Data Comparisons for Total Phosphorus at Nine Representative DEP HS Stations in 2005

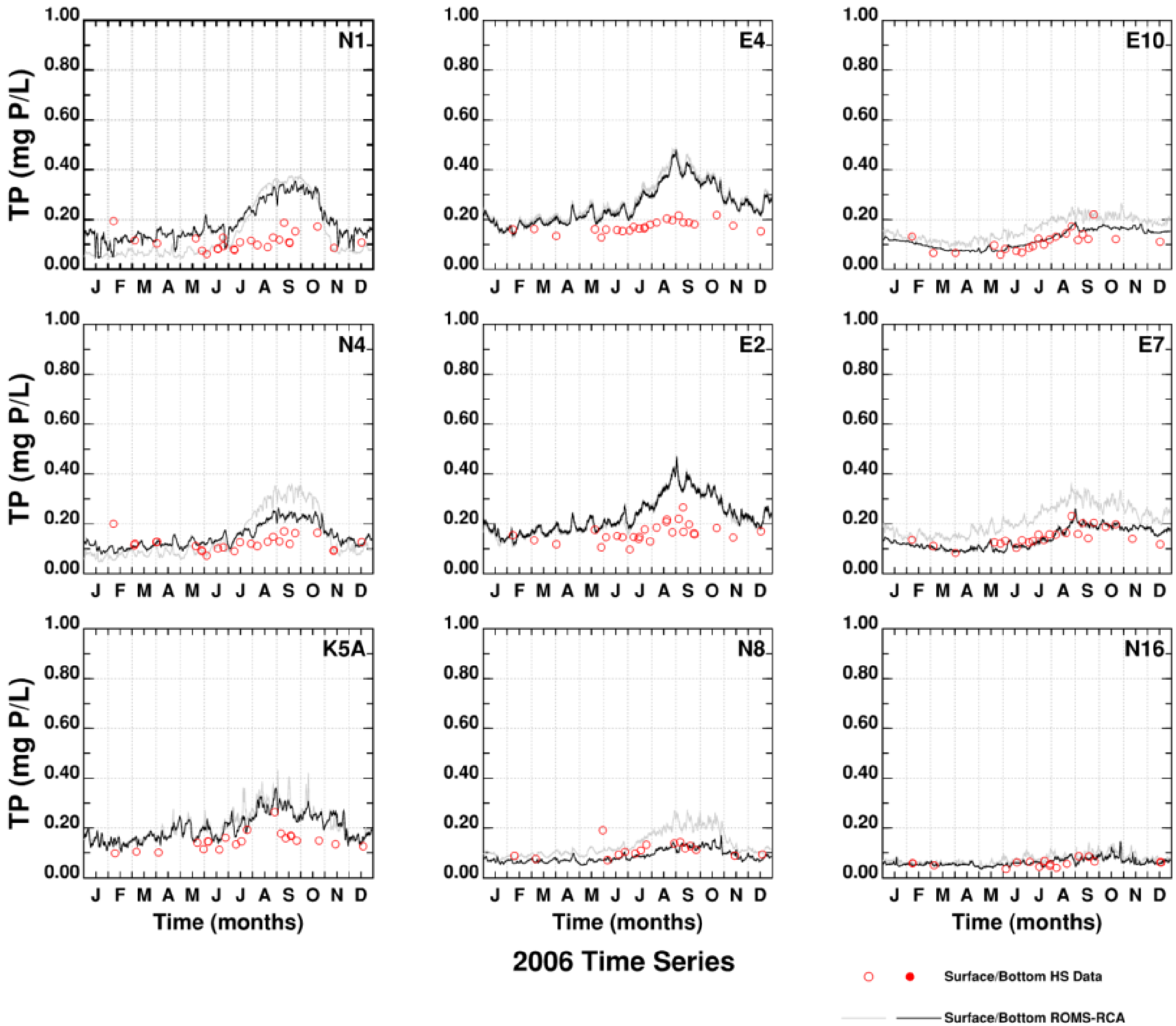

Figure 37. Model versus Data Comparisons for Total Phosphorus at Nine Representative DEP HS Stations in 2006

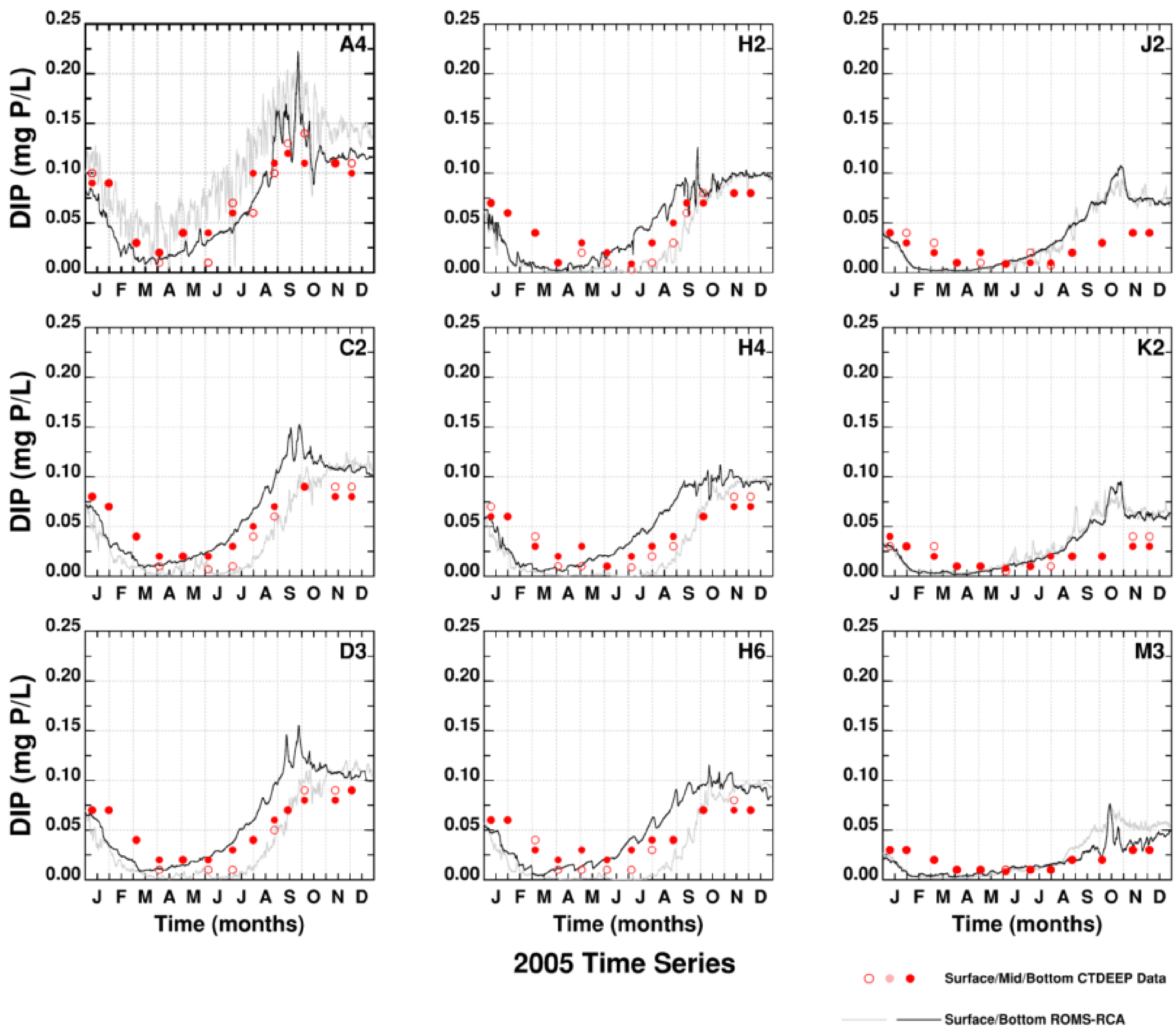

Figure 38. Model versus Data Comparisons for DIP at Nine Representative CTDEEP Stations in 2005

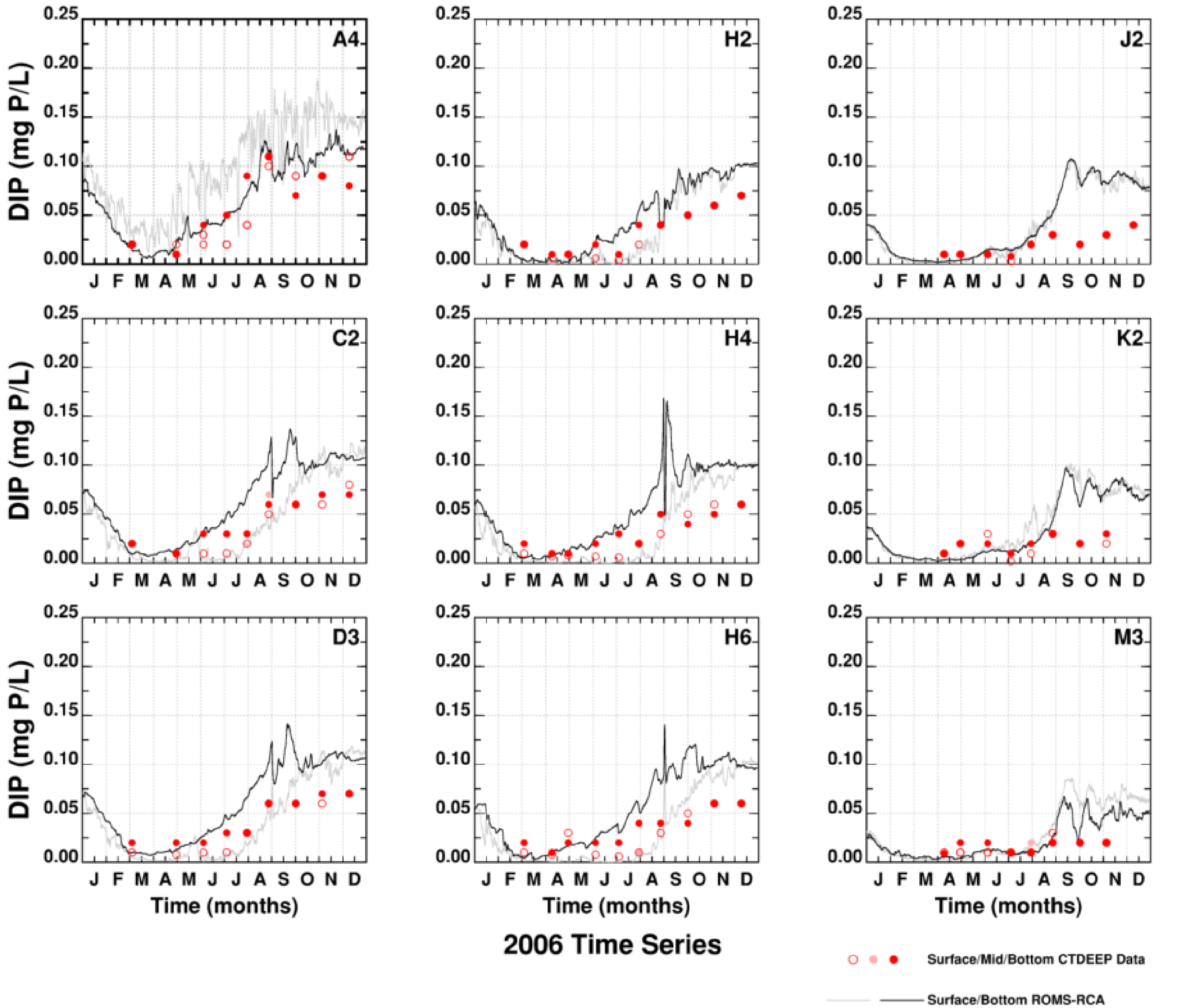

Figure 39. Model versus Data Comparisons for DIP at Nine Representative CTDEEP Stations in 2006

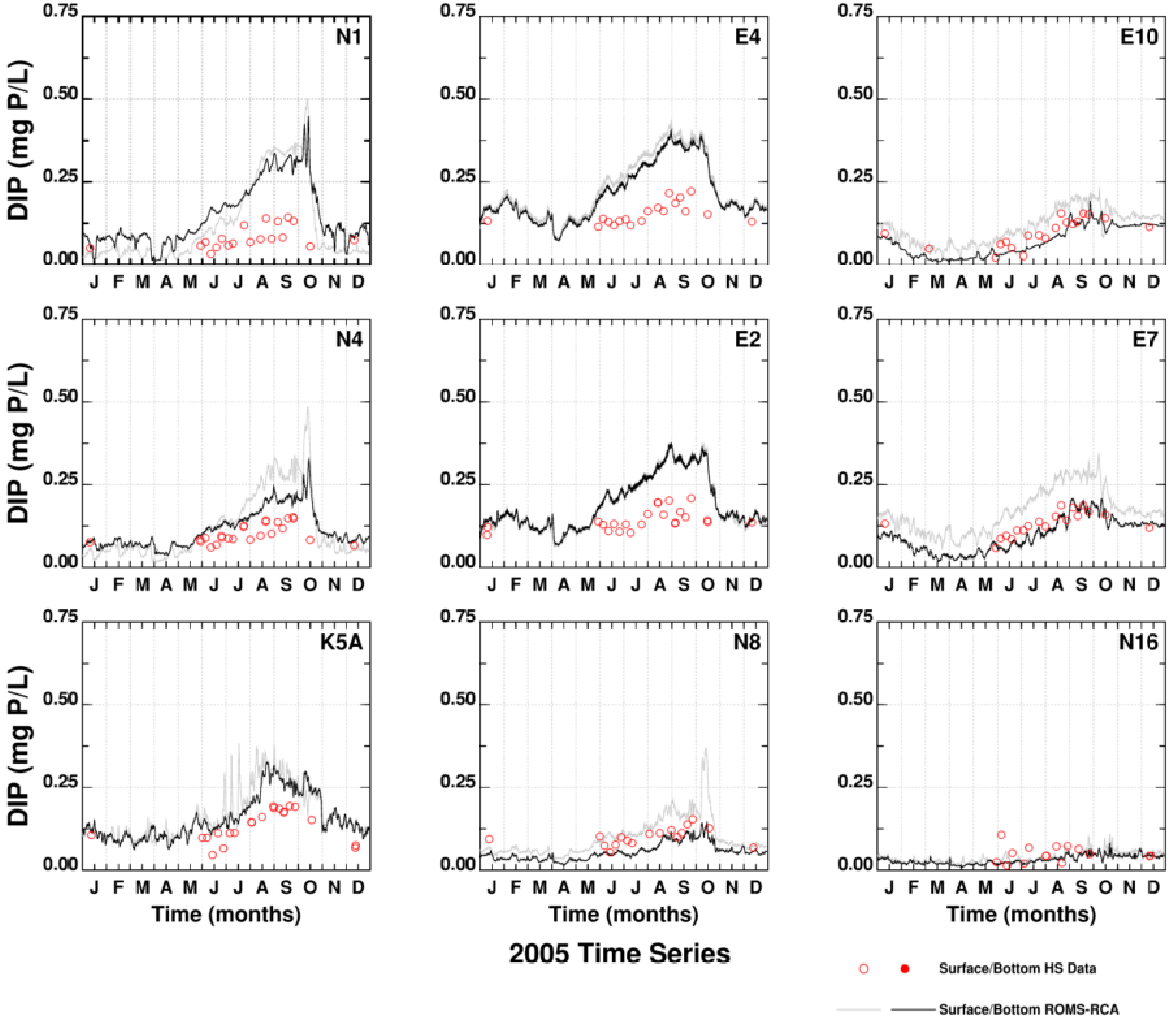

Figure 40. Model versus Data Comparisons for DIP at Nine Representative DEP HS Stations in 2005

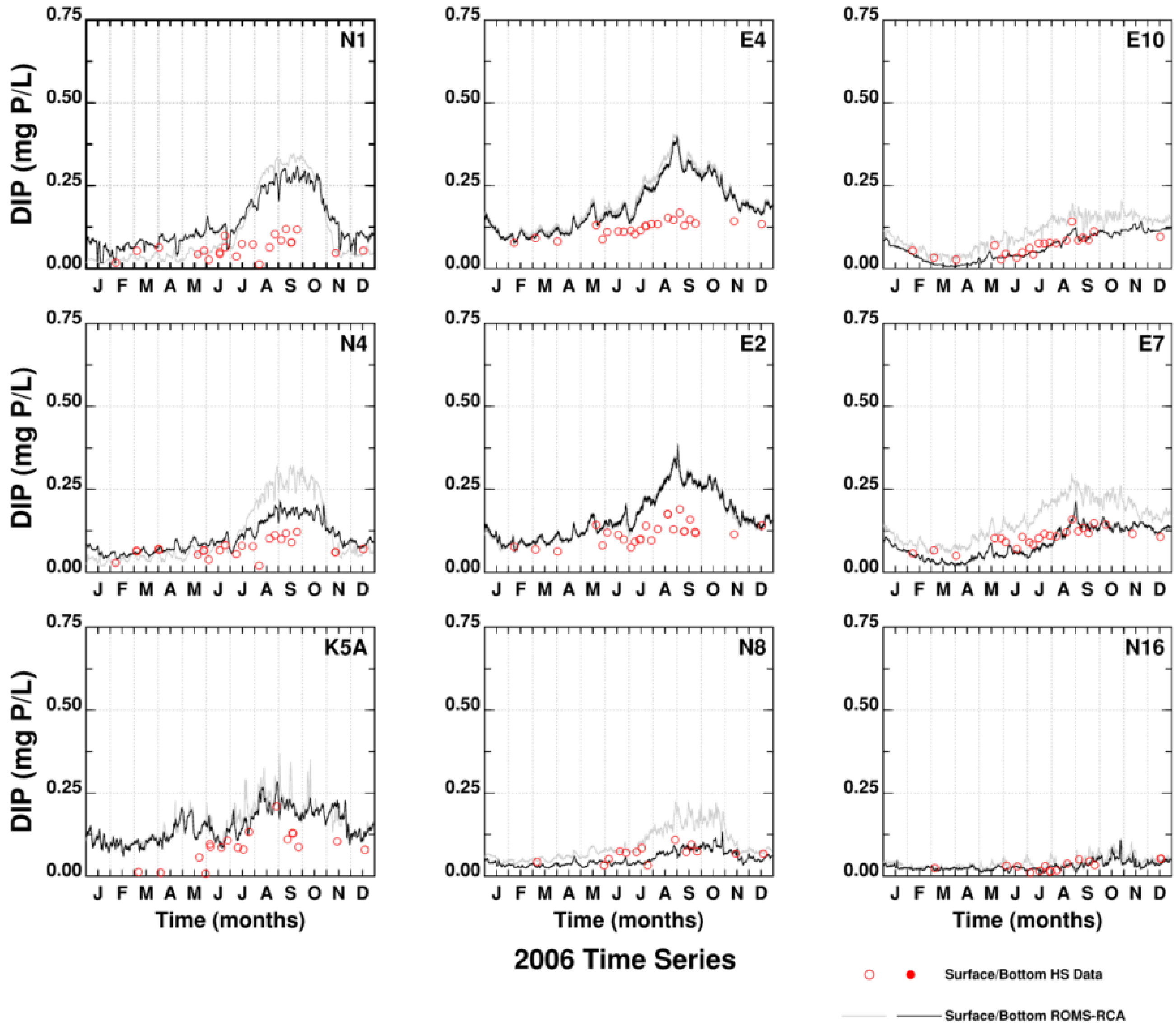

Figure 41. Model versus Data Comparisons for DIP at Nine Representative DEP HS Stations in 2006

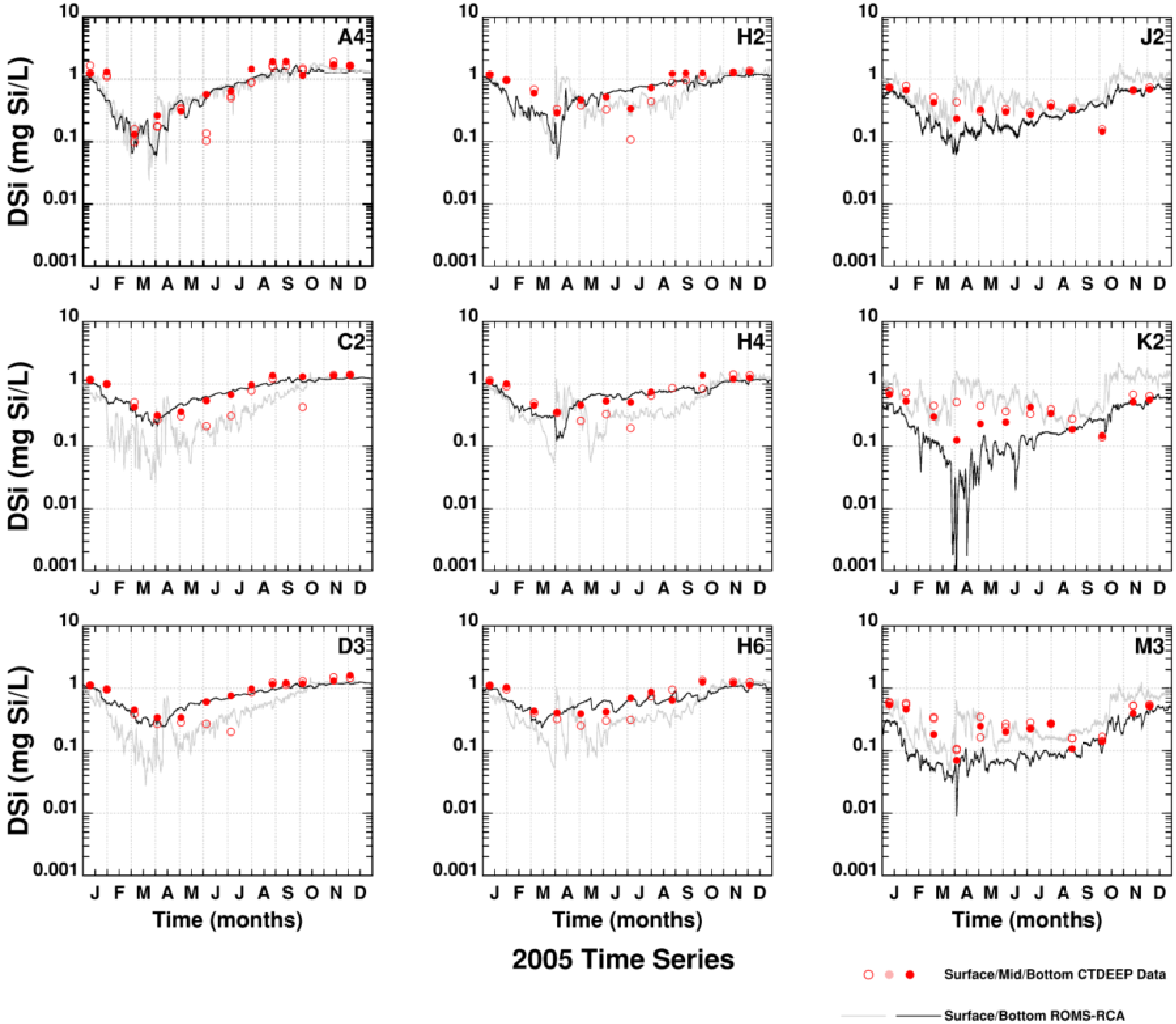

Figure 42. Model versus Data Comparisons for Dissolved Silica at Nine Representative CTDEEP Stations in 2005

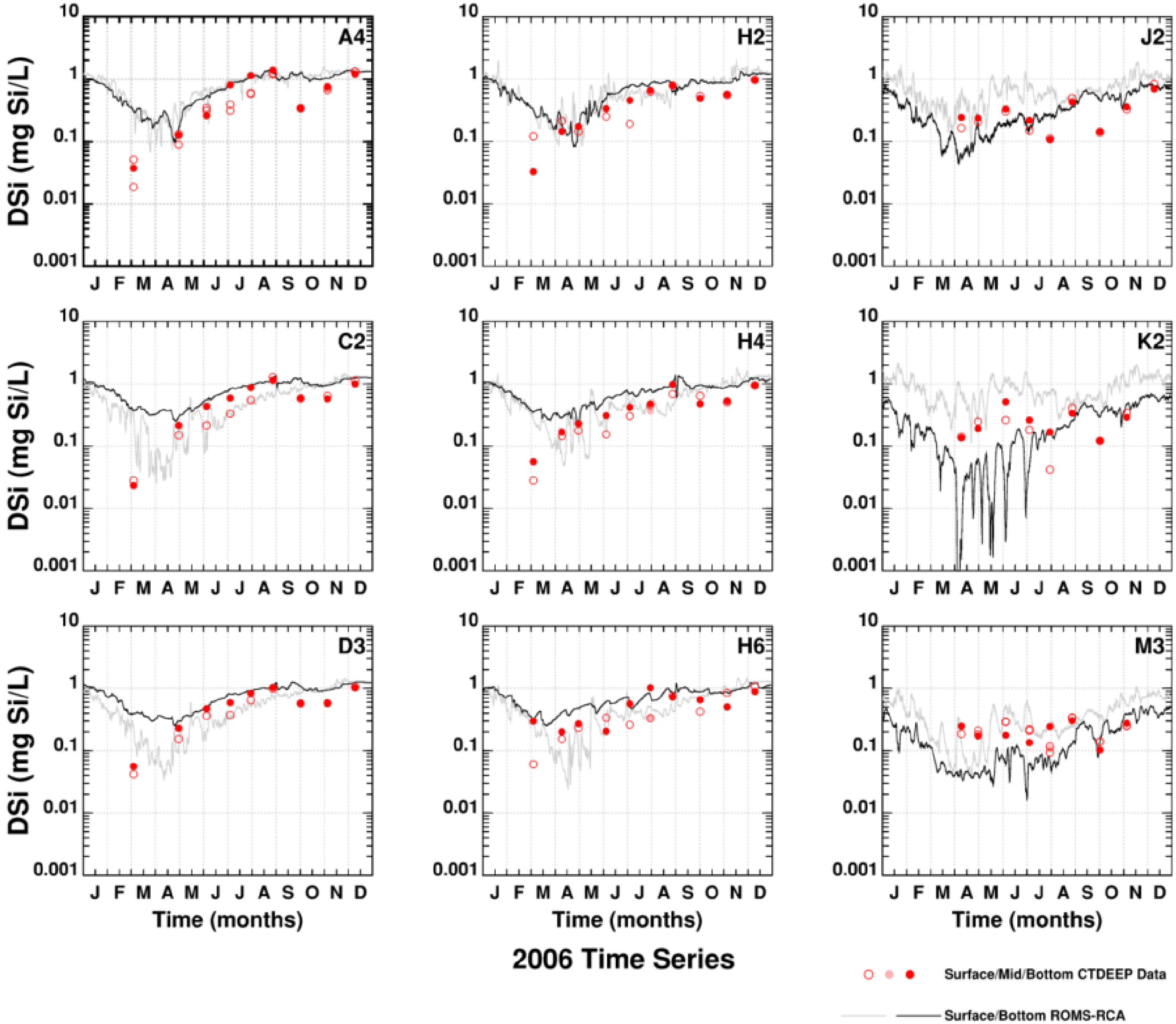

Figure 43. Model versus Data Comparisons for Dissolved Silica at Nine Representative CTDEEP Stations in 2006

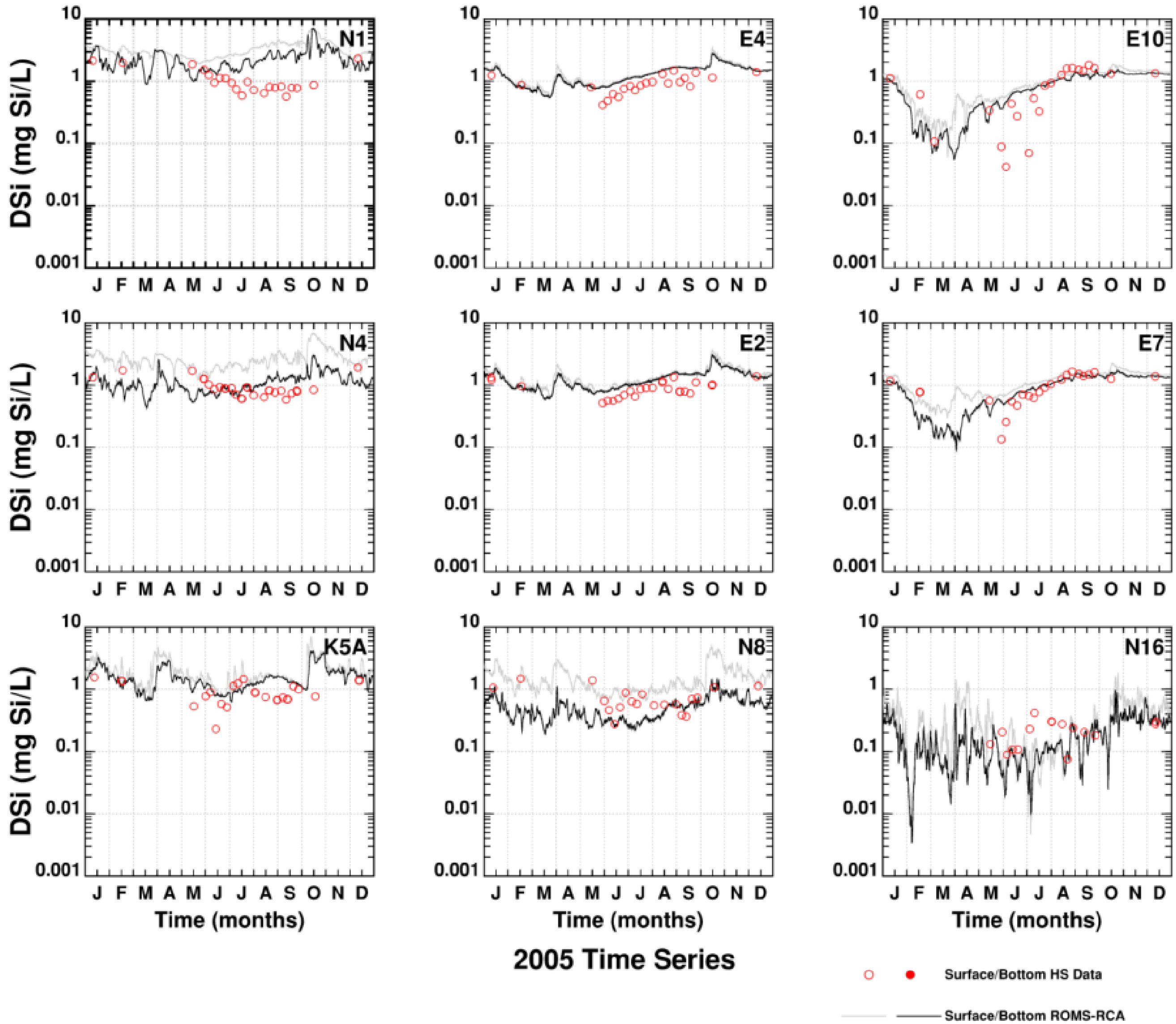

Figure 44. Model versus Data Comparisons for Dissolved Silica at Nine Representative DEP HS Stations in 2005

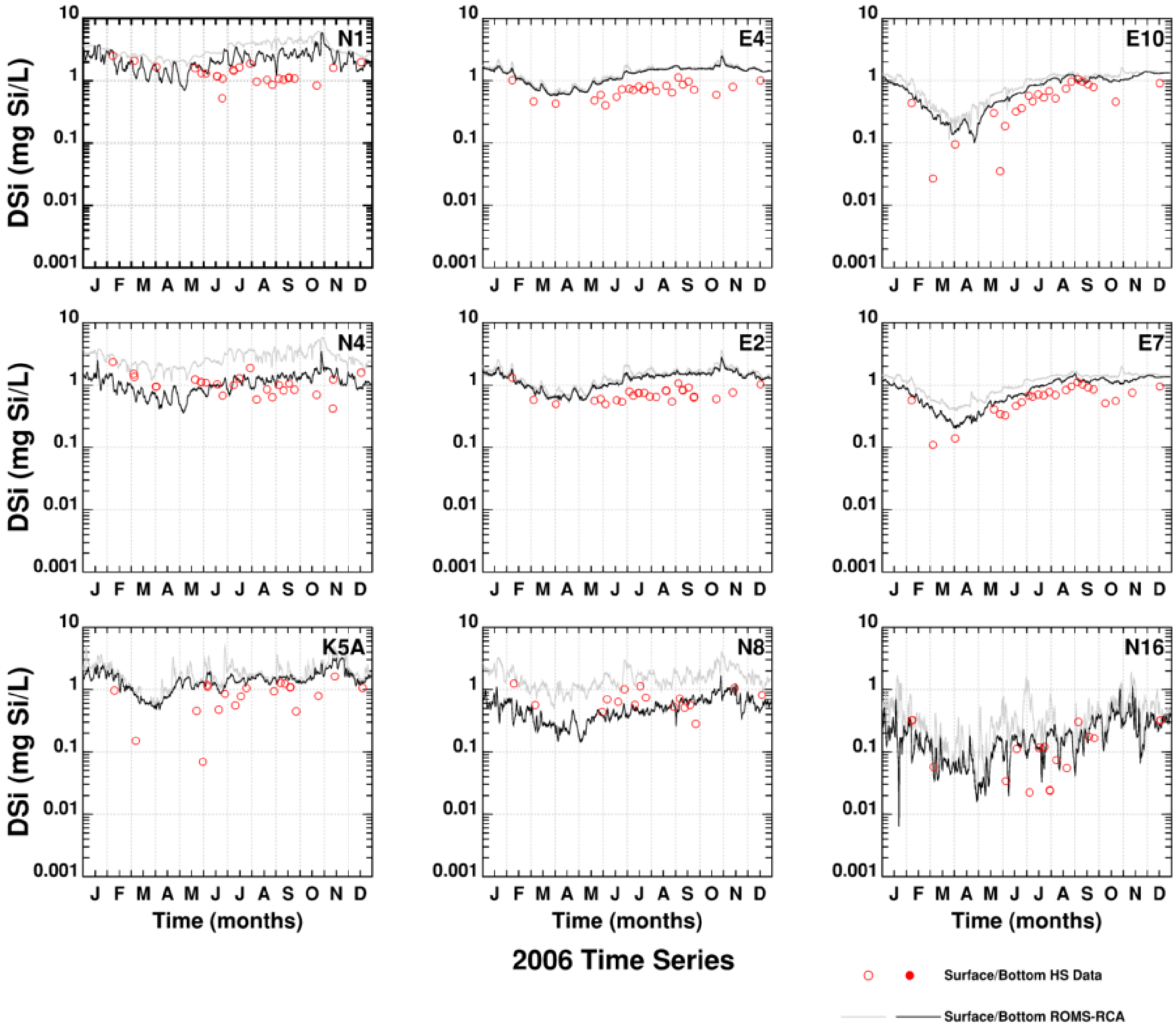

Figure 45. Model versus Data Comparisons for Dissolved Silica at Nine Representative DEP HS Stations in 2006

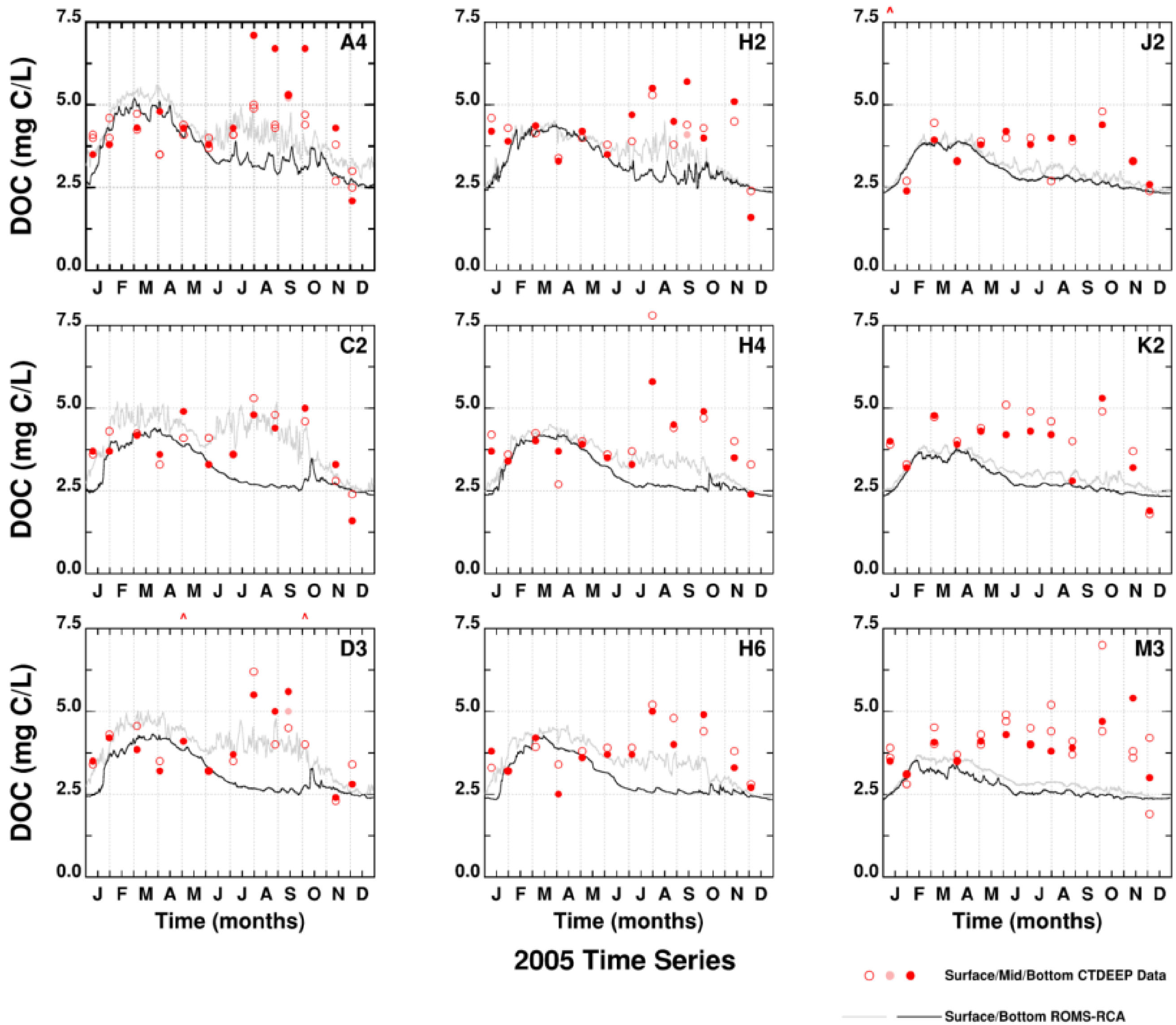

Figure 46. Model versus Data Comparisons for Dissolved Organic Carbon at Nine Representative CTDEEP Stations in 2005

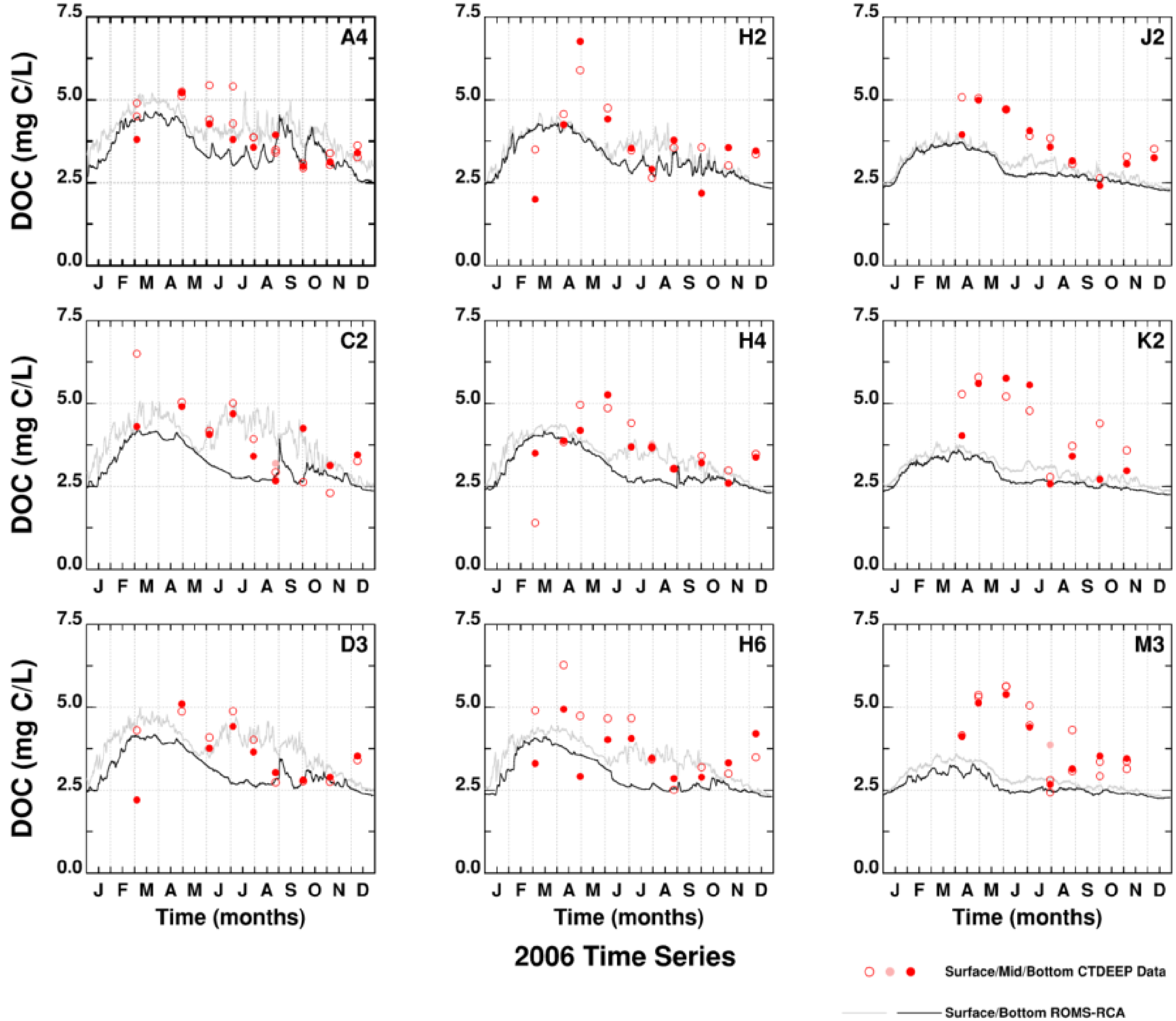

Figure 47. Model versus Data Comparisons for Dissolved Organic Carbon at Nine Representative CTDEEP Stations in 2006

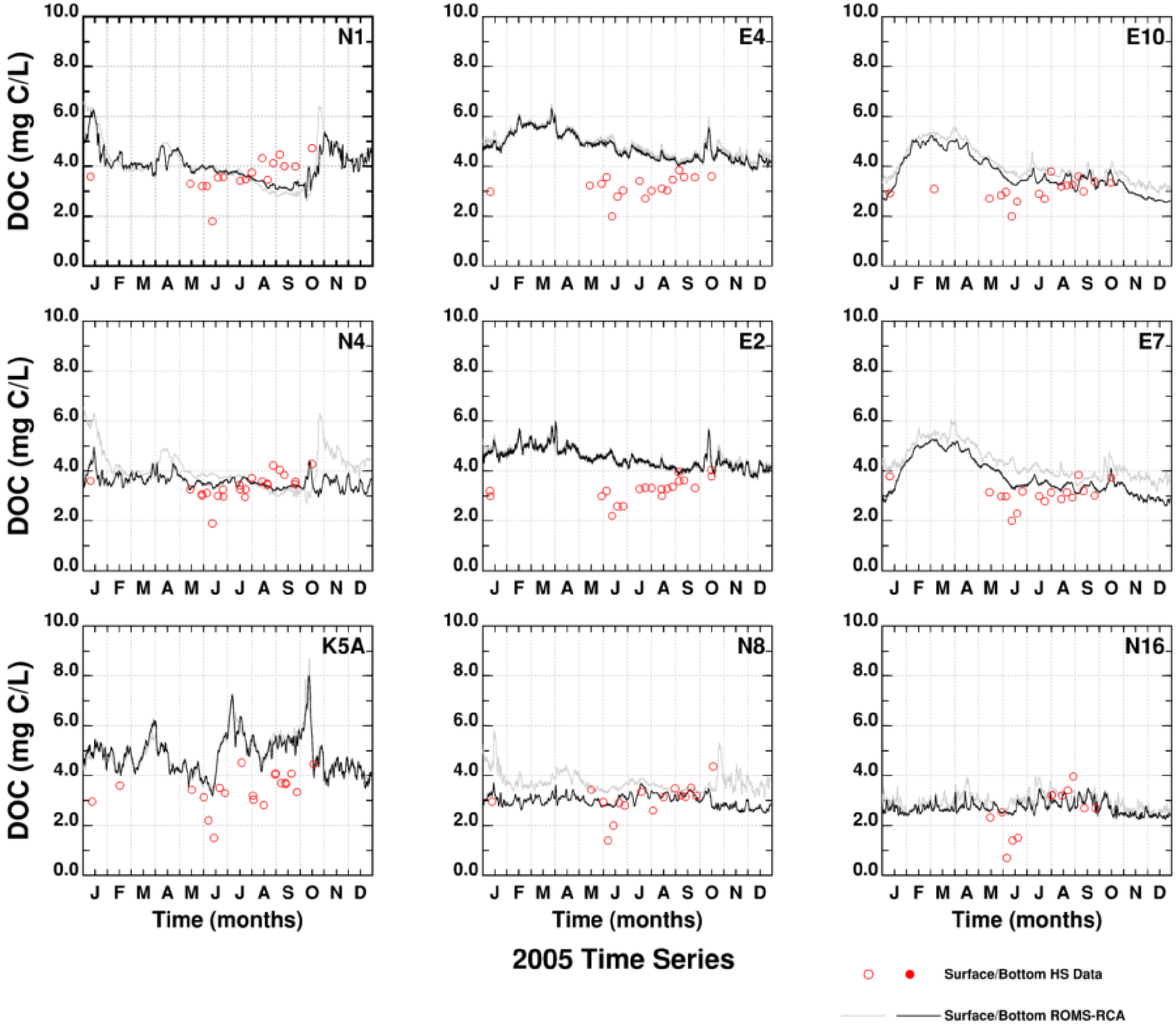

Figure 48. Model versus Data Comparisons for Dissolved Organic Carbon at Nine Representative DEP HS Stations in 2005

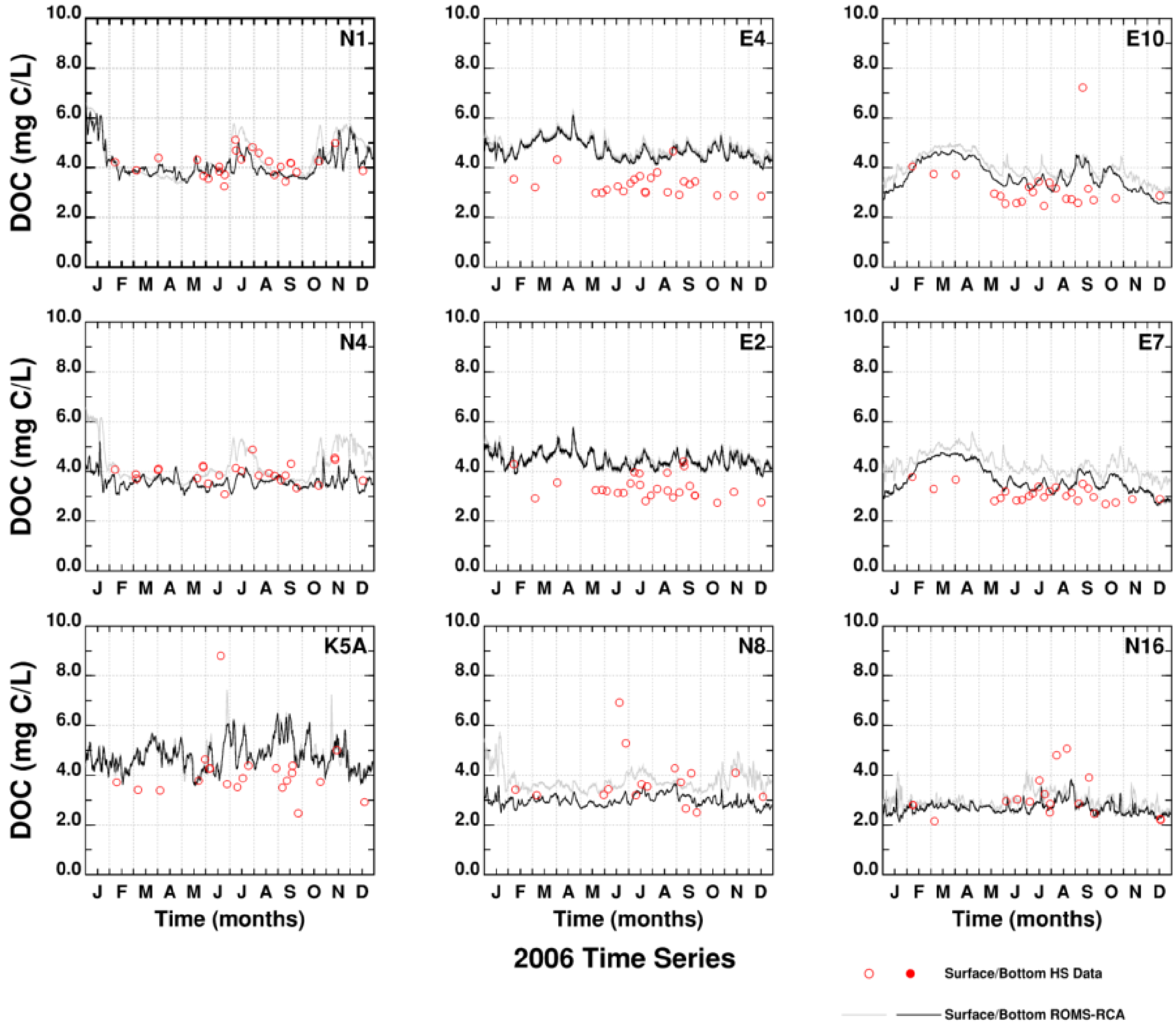

Figure 49. Model versus Data Comparisons for Dissolved Organic Carbon at Nine Representative DEP HS Stations in 2006

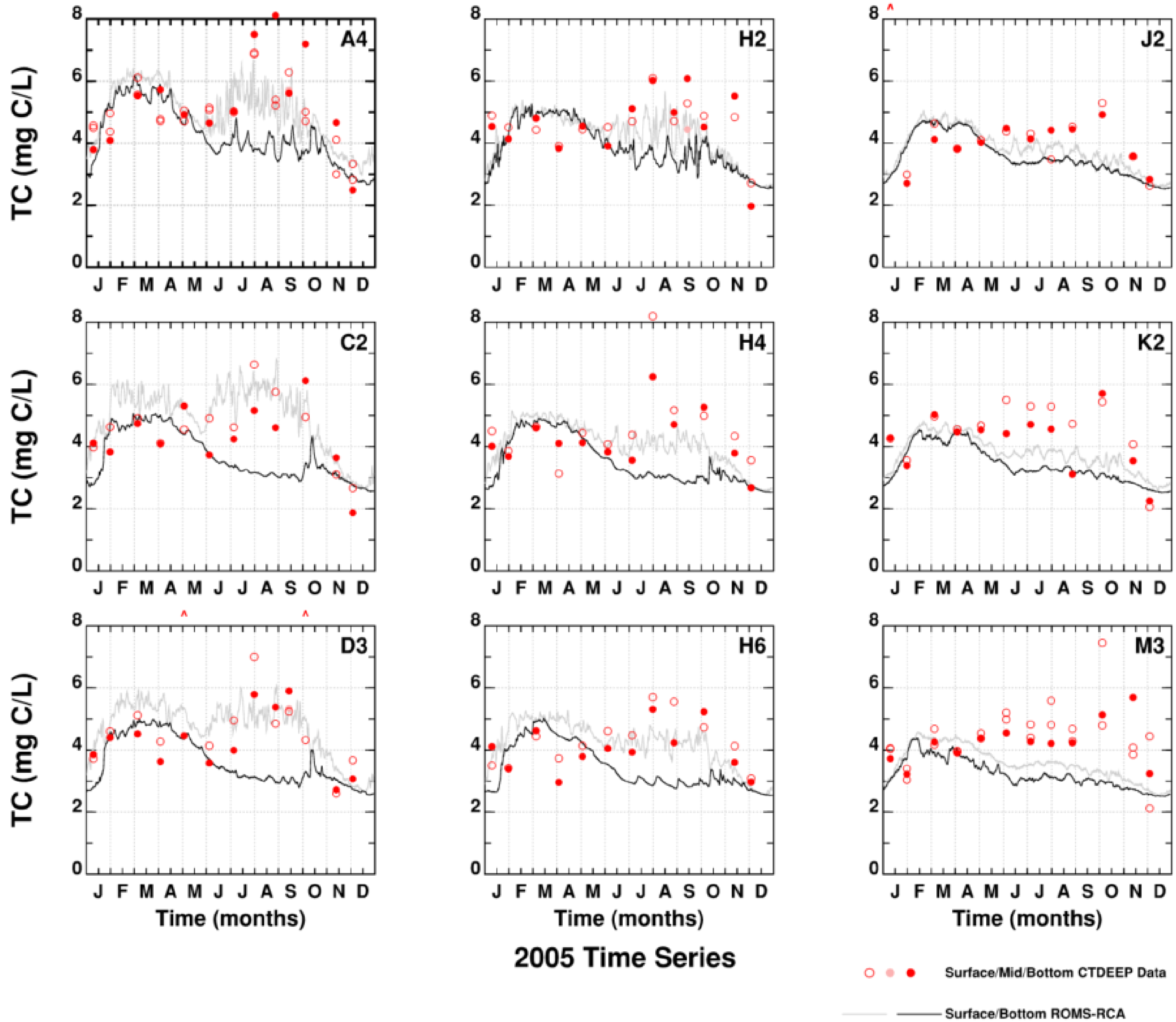

Figure 50. Model versus Data Comparisons for Total Organic Carbon at Nine Representative CTDEEP Stations in 2005

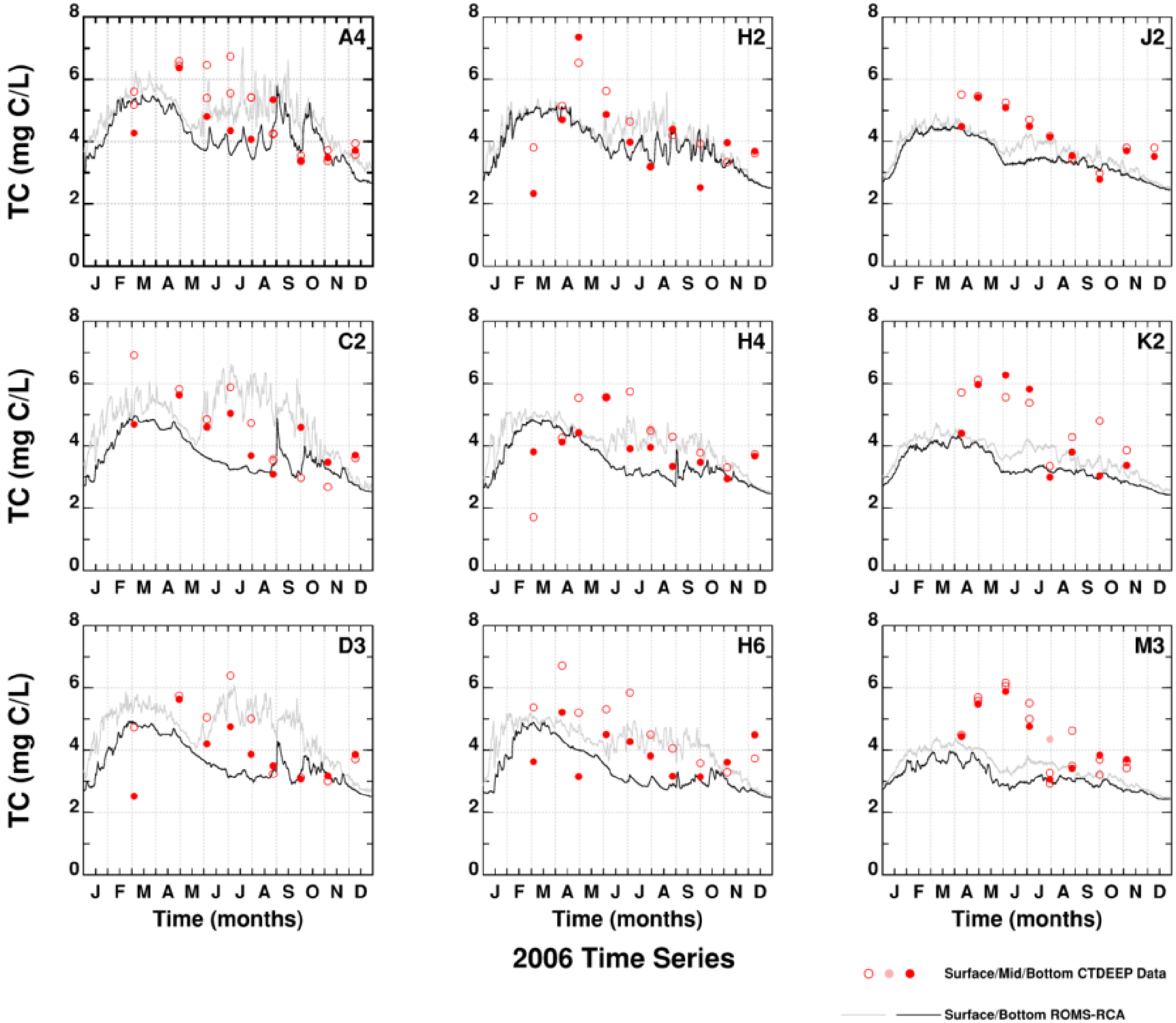

Figure 51. Model versus Data Comparisons for Total Organic Carbon at Nine Representative CTDEEP Stations in 2006

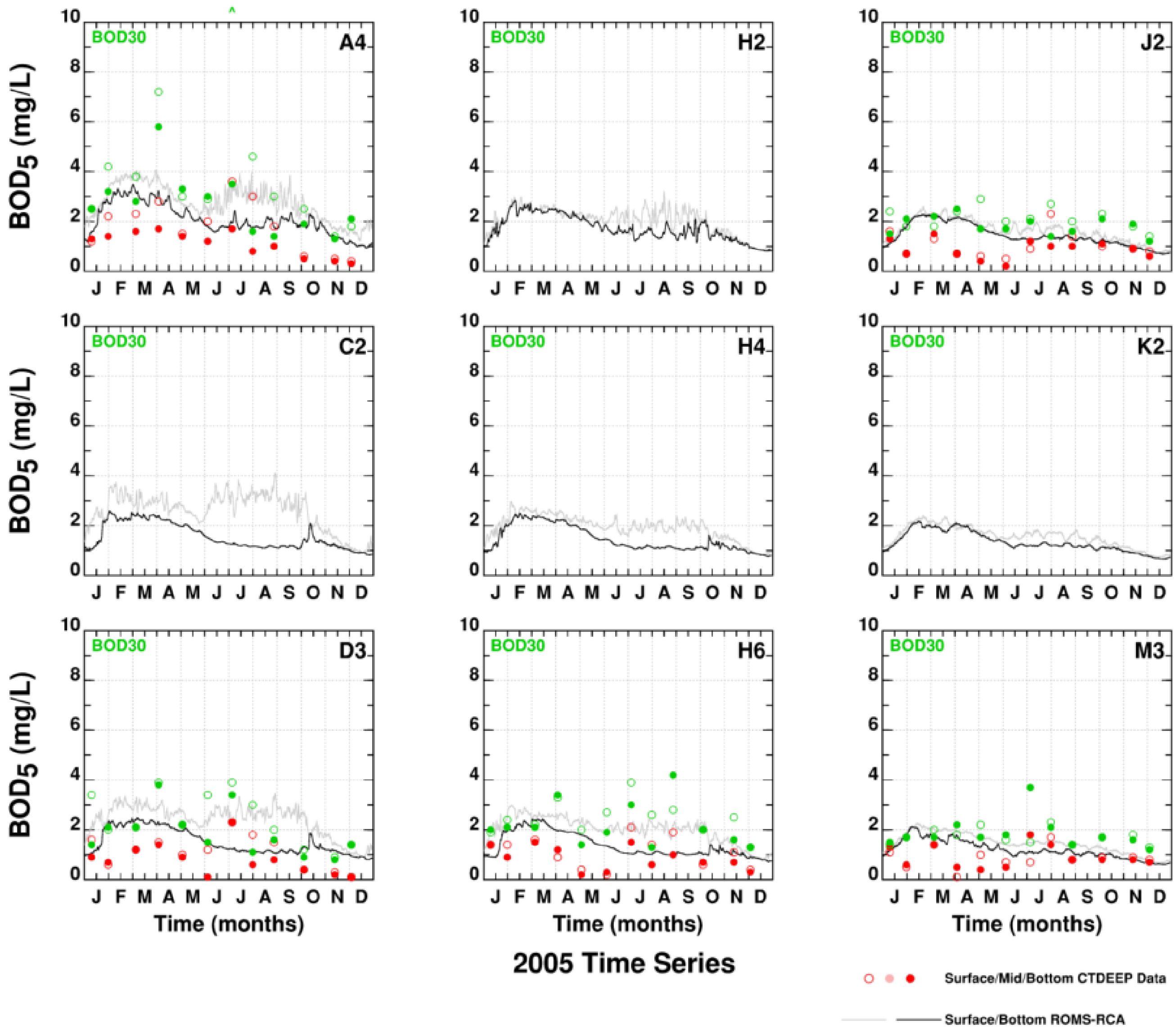

Figure 52. Model versus Data Comparisons for 5-day BOD at Nine Representative CTDEEP Stations in 2005

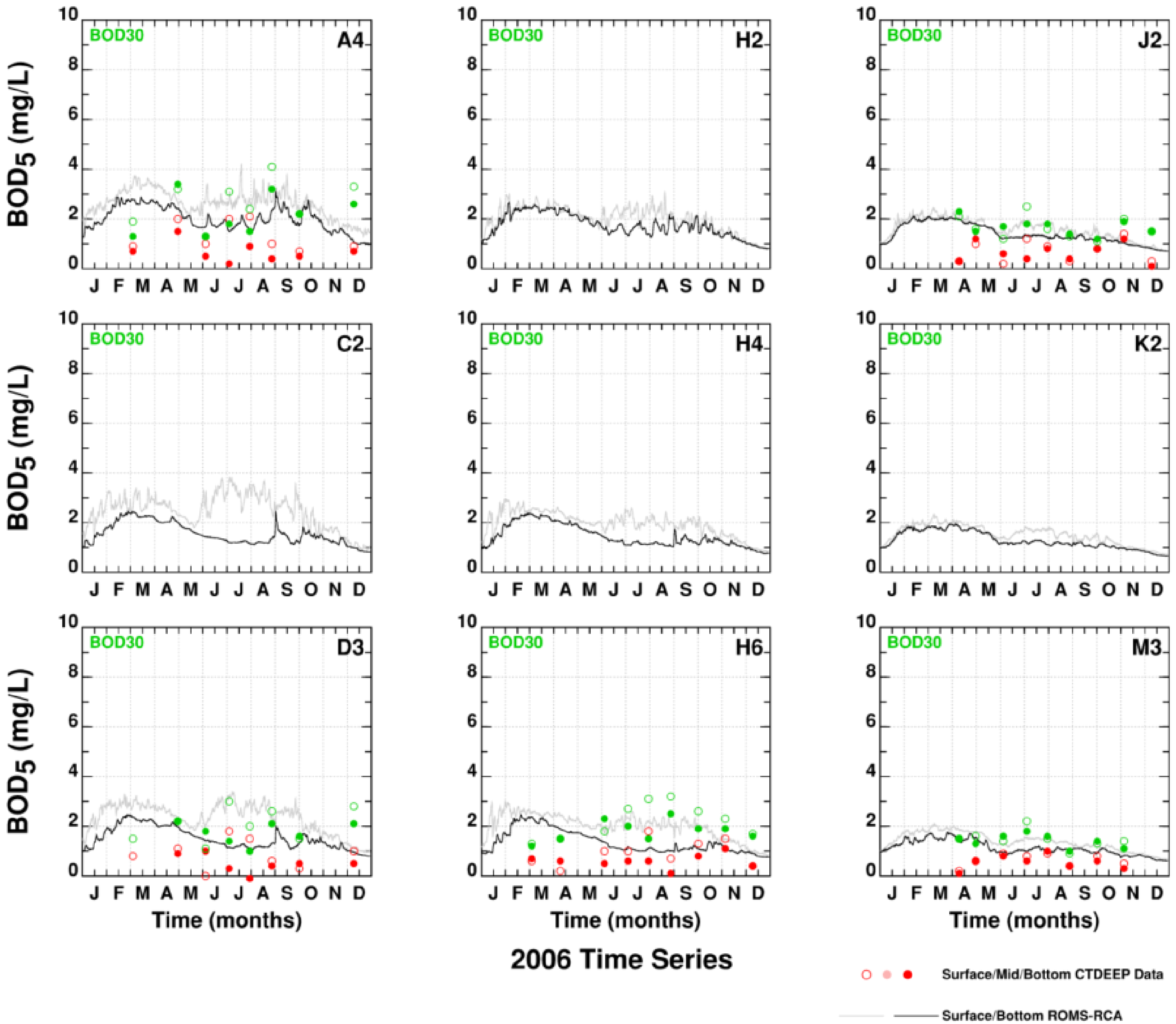

Figure 53. Model versus Data Comparisons for 5-day BOD at Nine Representative CTDEEP Stations in 2006

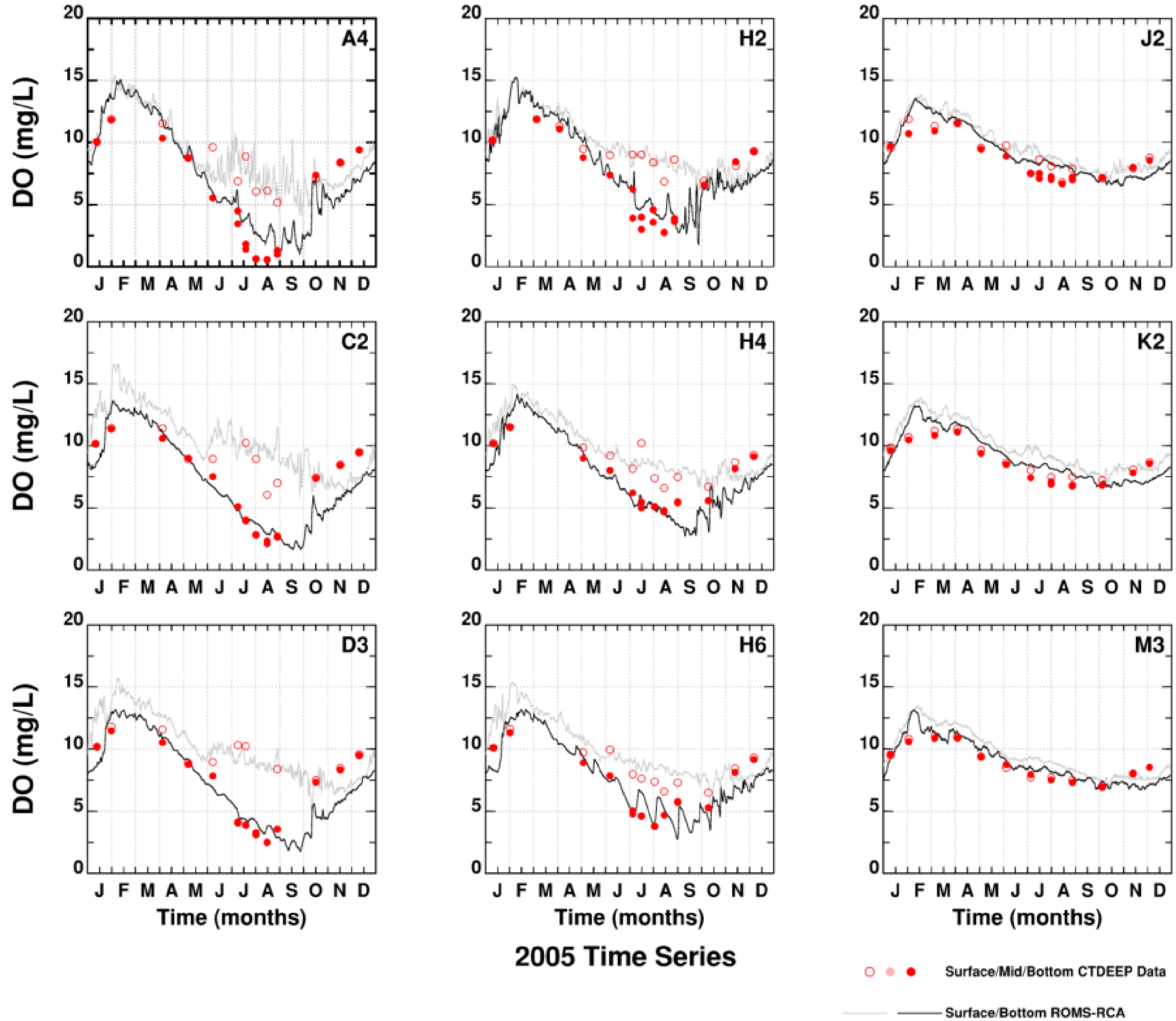

Figure 54. Model versus Data Comparisons for Dissolved Oxygen at Nine Representative CTDEEP Stations in 2005

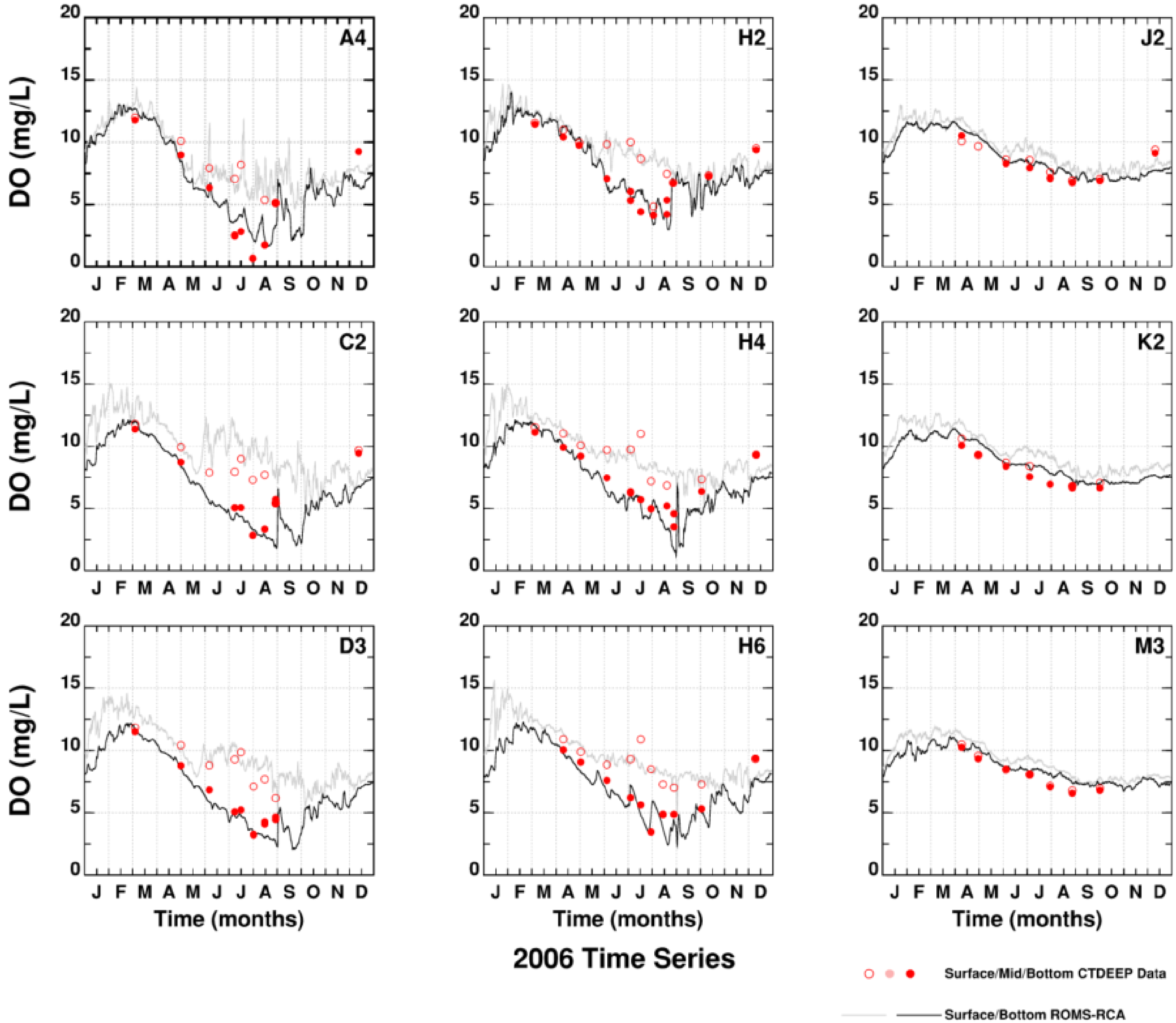

Figure 55. Model versus Data Comparisons for Dissolved Oxygen Carbon at Nine Representative CTDEEP Stations in 2006

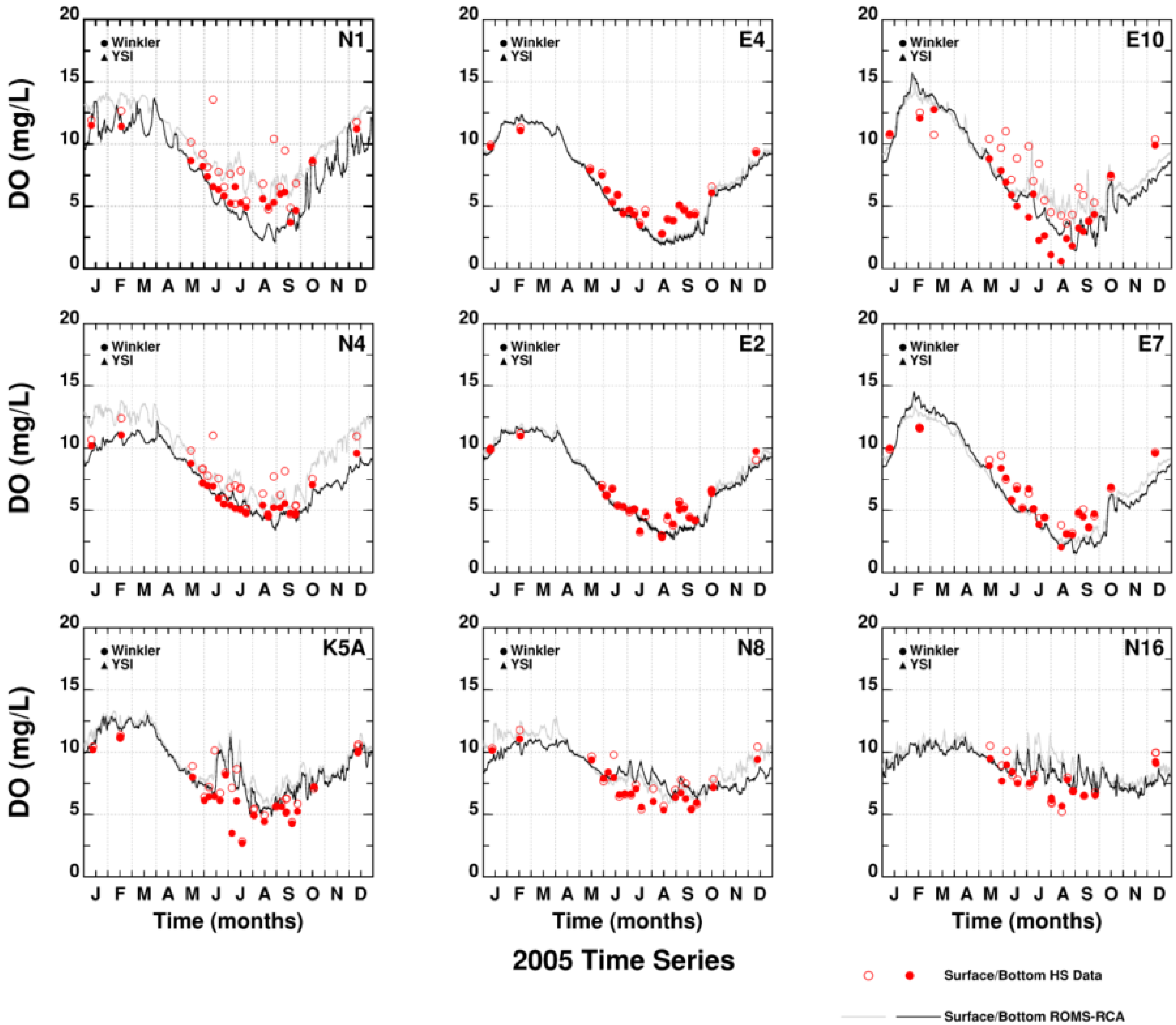

Figure 56. Model versus Data Comparisons for Dissolved Oxygen Carbon at Nine Representative DEP HS Stations in 2005

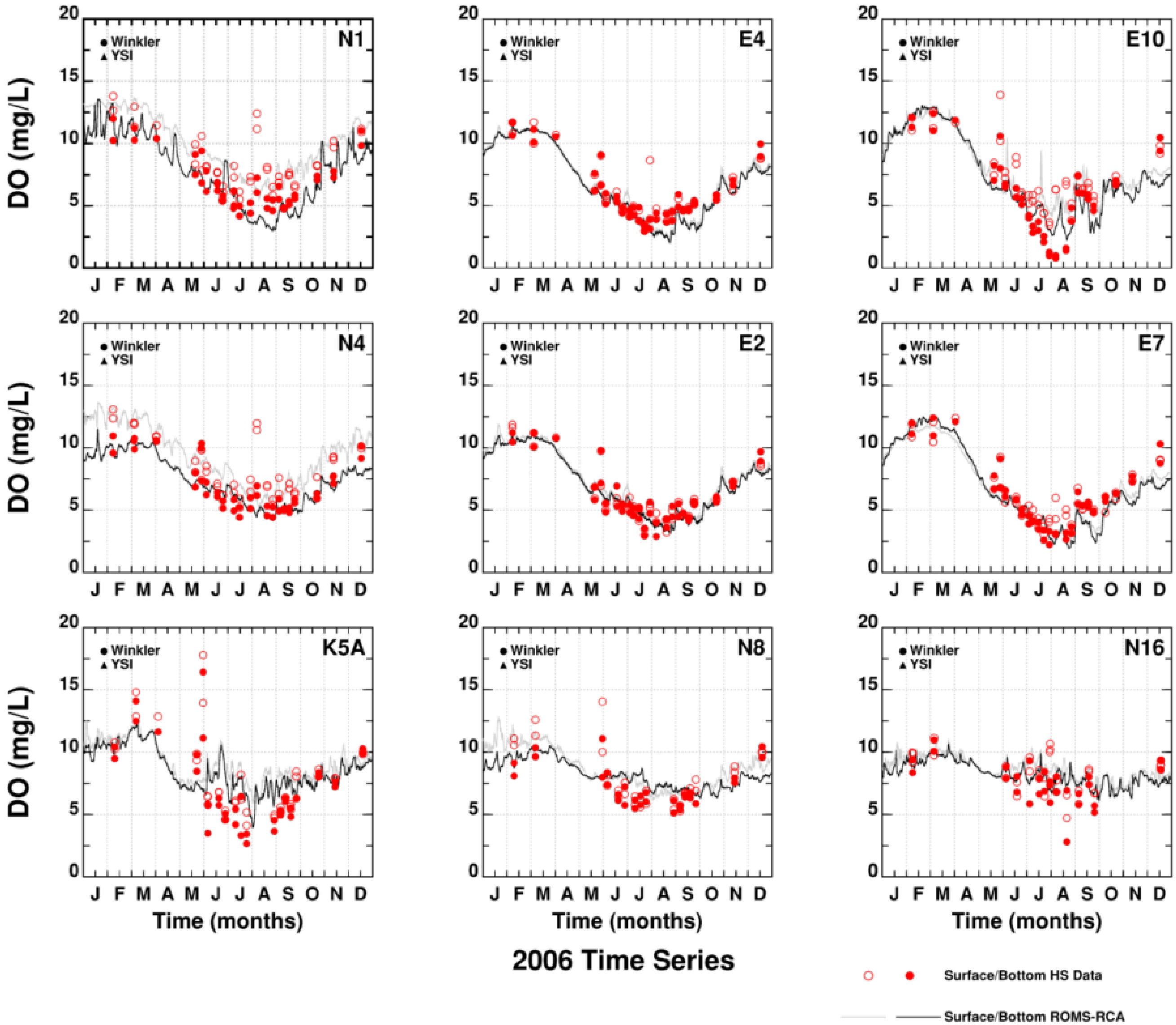

Figure 57. Model versus Data Comparisons for Dissolved Oxygen Carbon at Nine Representative DEP HS Stations in 2006

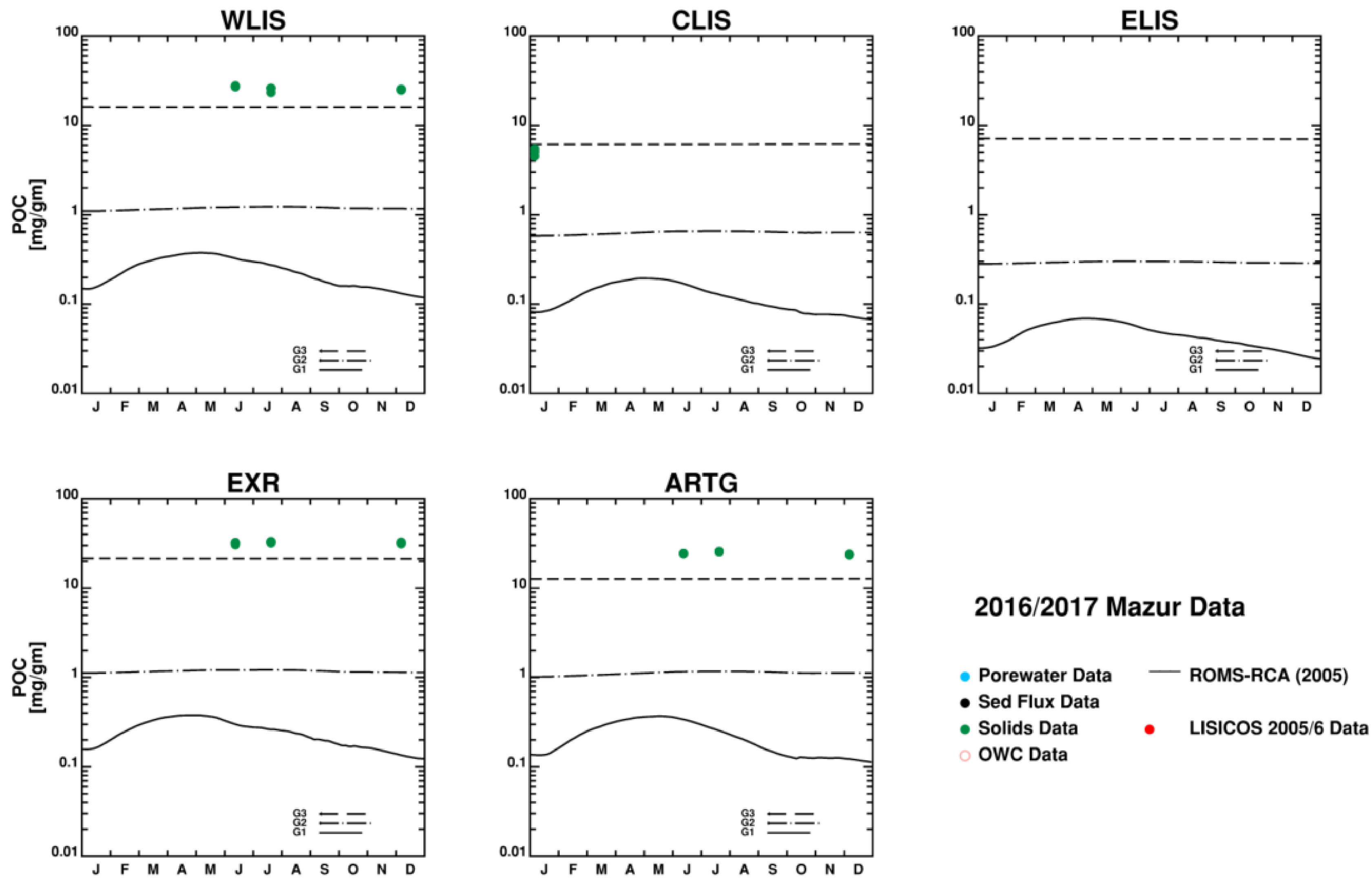

Figure 58. Model versus Data Comparisons for Sediment Particulate Organic Carbon in 2005

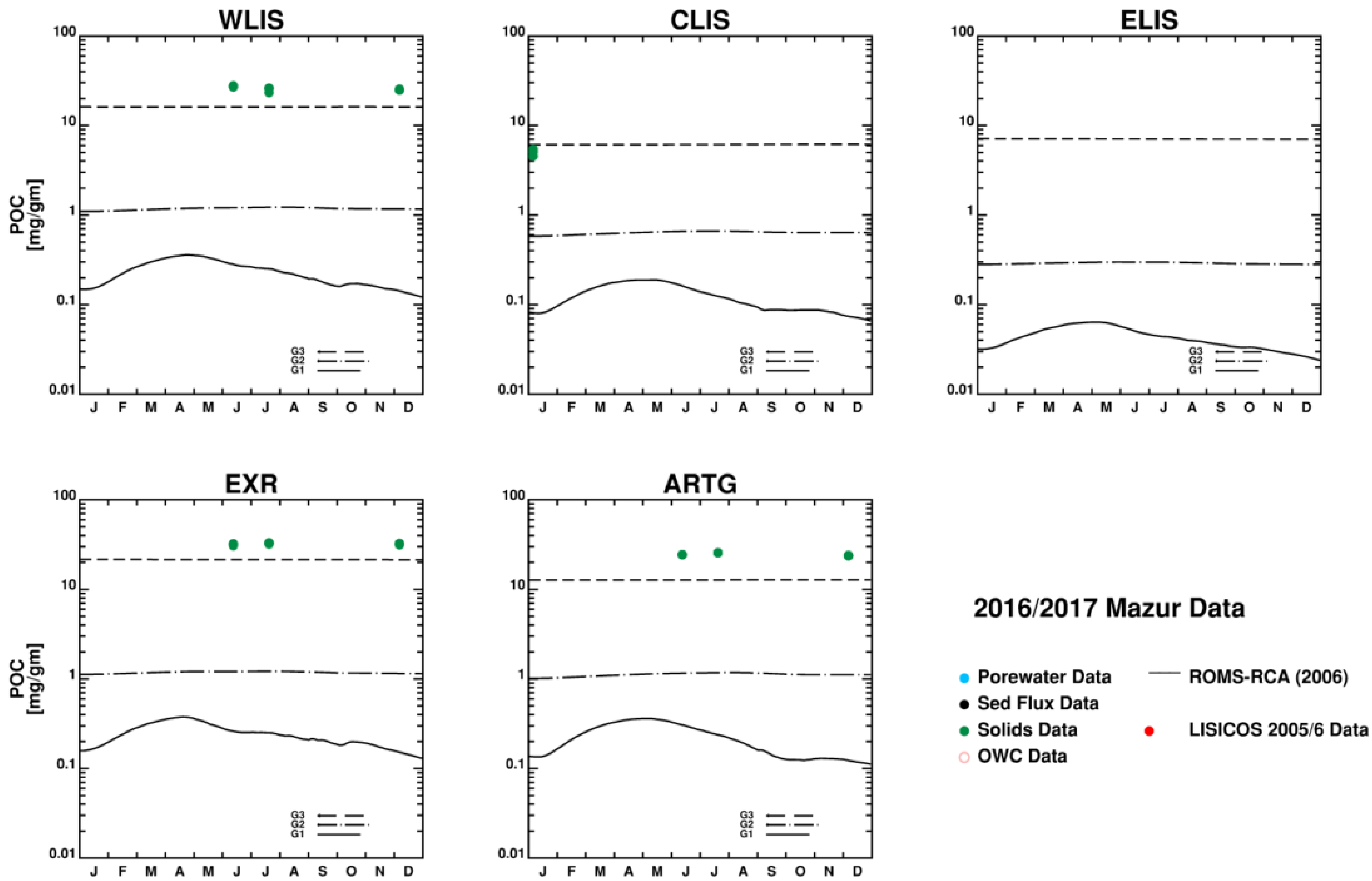

Figure 59. Model versus Data Comparisons for Sediment Particulate Organic Carbon in 2006

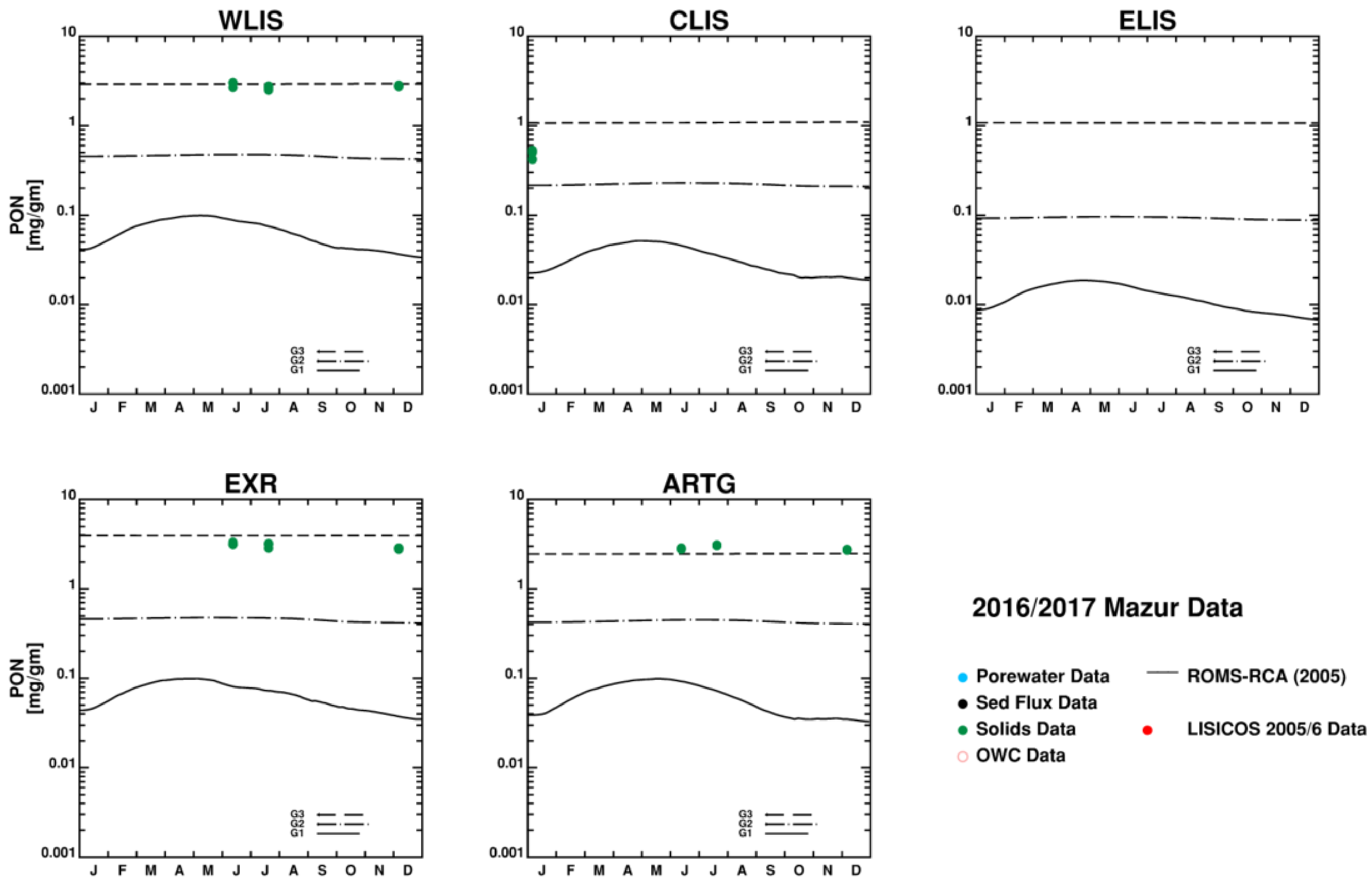

Figure 60. Model versus Data Comparisons for Sediment Particulate Organic Nitrogen in 2005

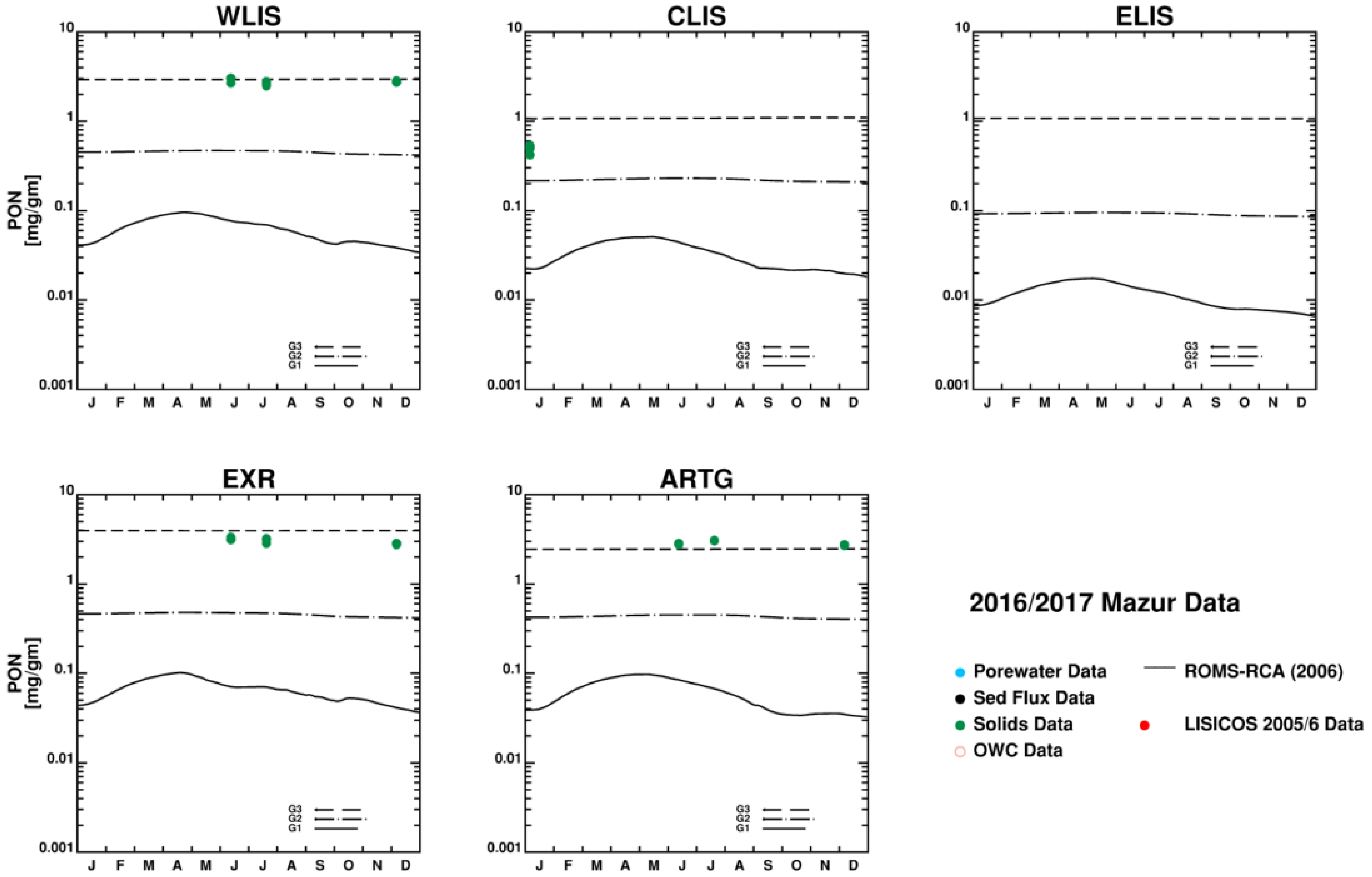

Figure 61. Model versus Data Comparisons for Sediment Particulate Organic Nitrogen in 2006

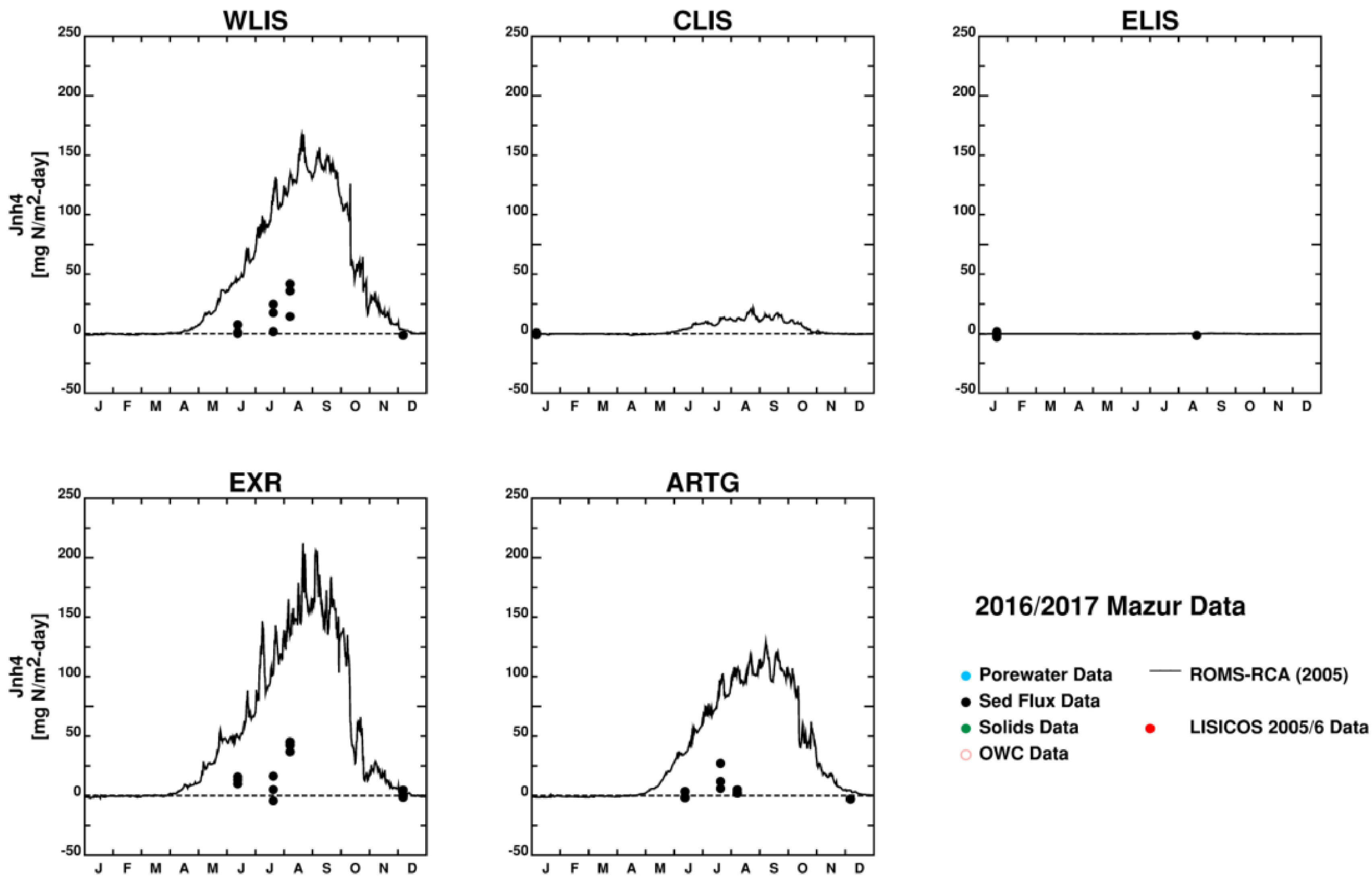

Figure 62. Model versus Data Comparisons for Sediment Ammonium Fluxes in 2005

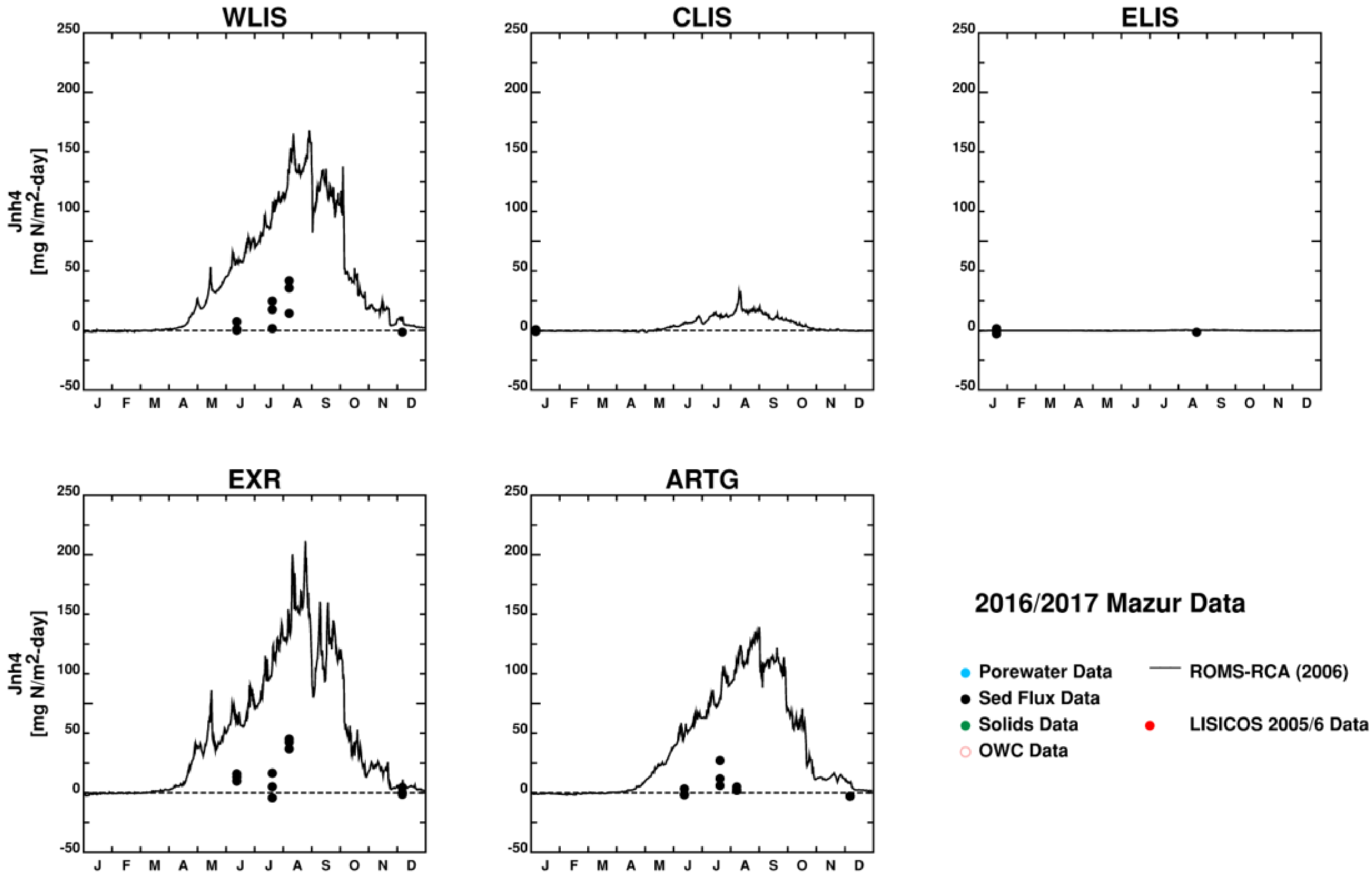

Figure 63. Model versus Data Comparisons for Sediment Ammonium Fluxes in 2006

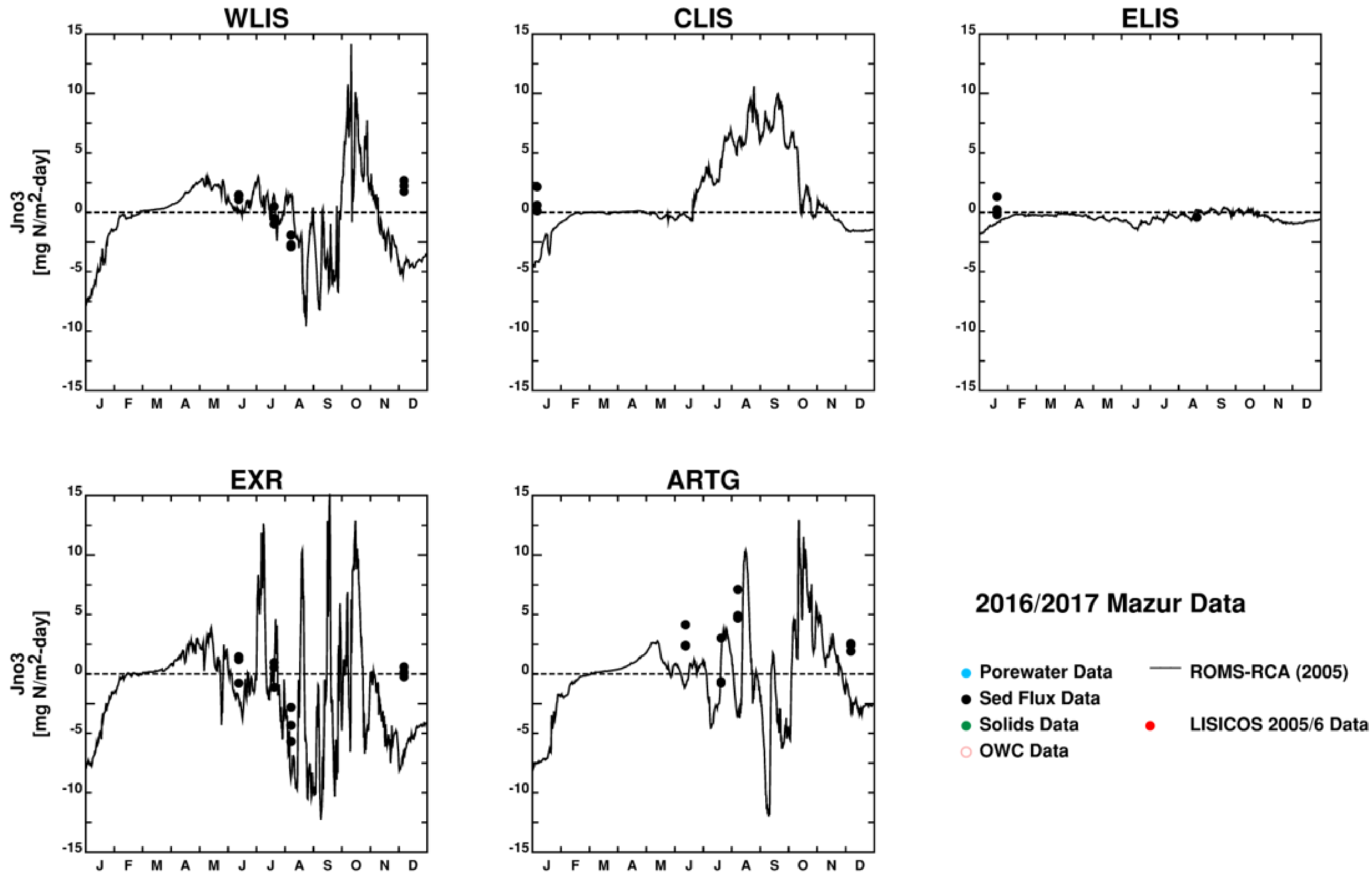

Figure 64. Model versus Data Comparisons for Sediment Nitrate Fluxes in 2005

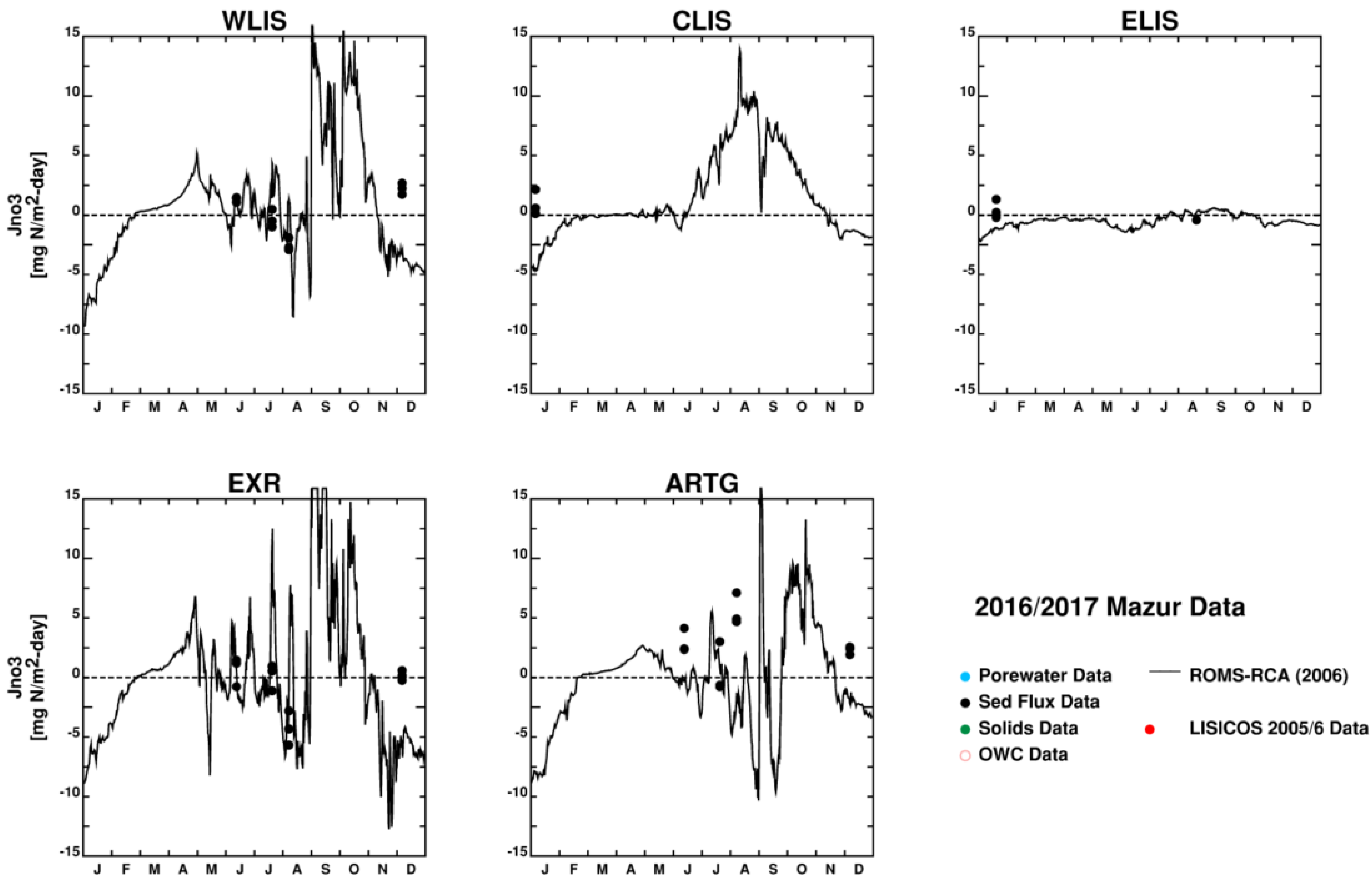

Figure 65. Model versus Data Comparisons for Sediment Nitrate Fluxes in 2006

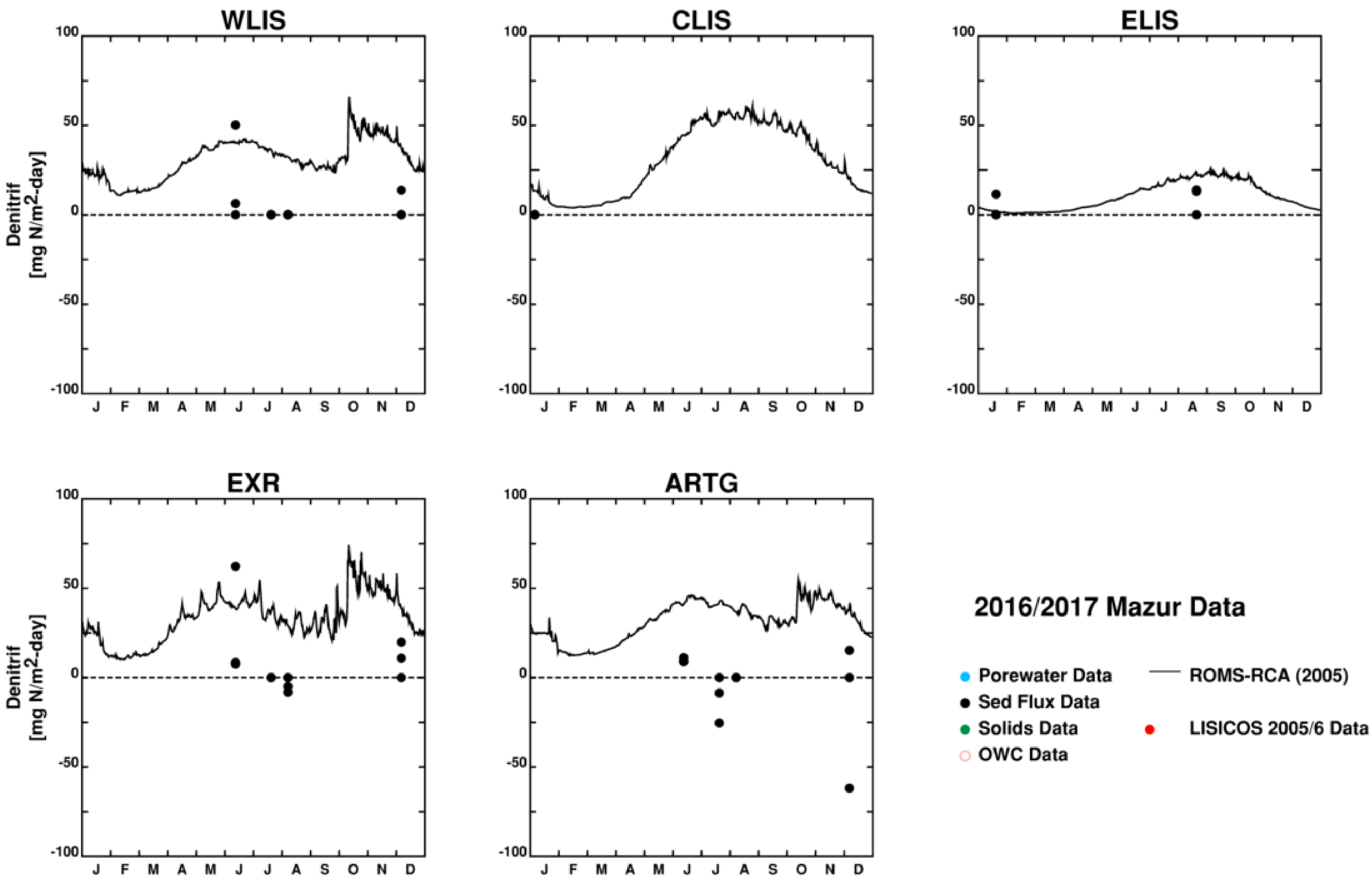

Figure 66. Model versus Data Comparisons for Sediment Denitrification in 2005

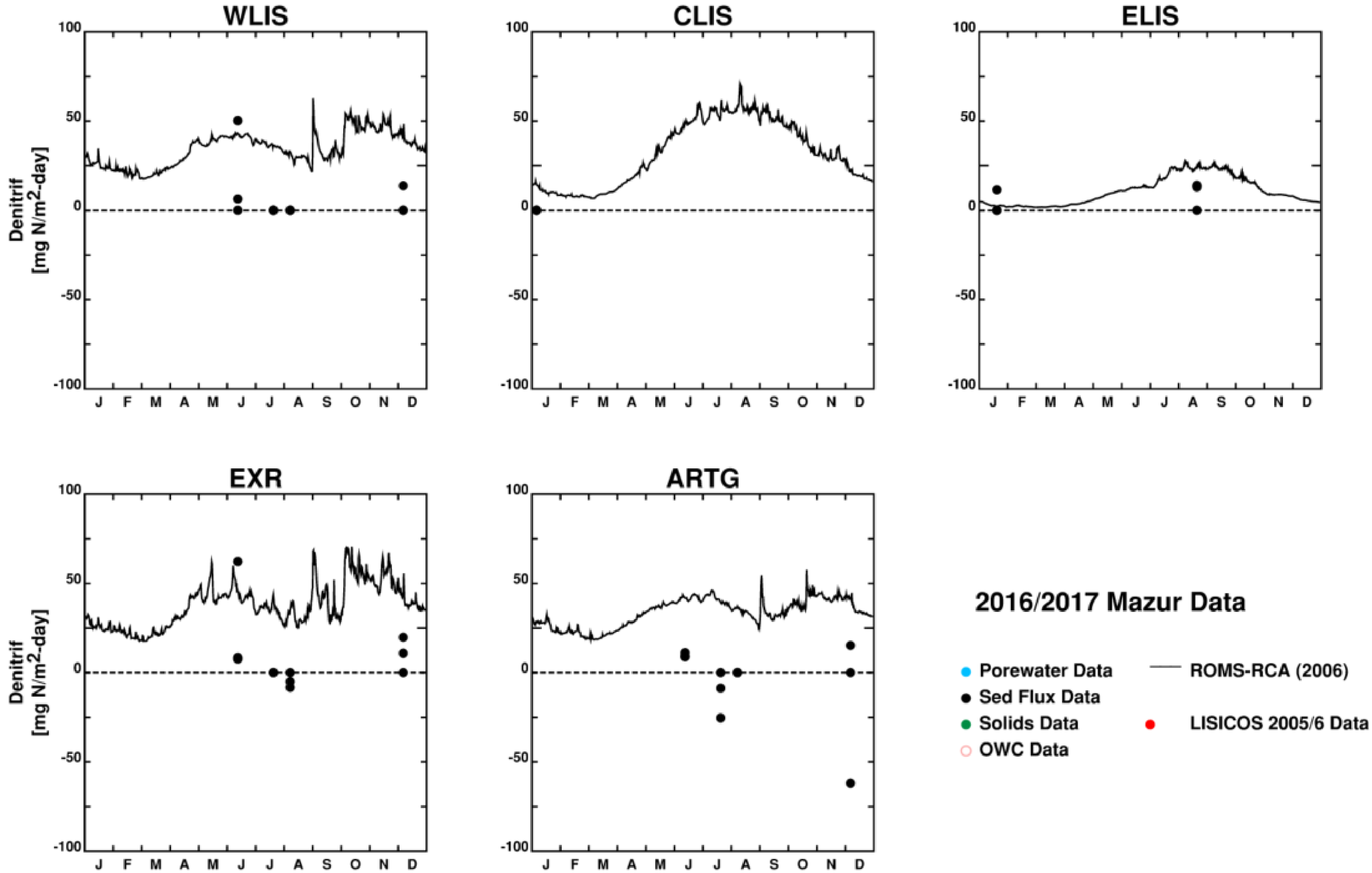

Figure 67. Model versus Data Comparisons for Sediment Denitrification in 2006

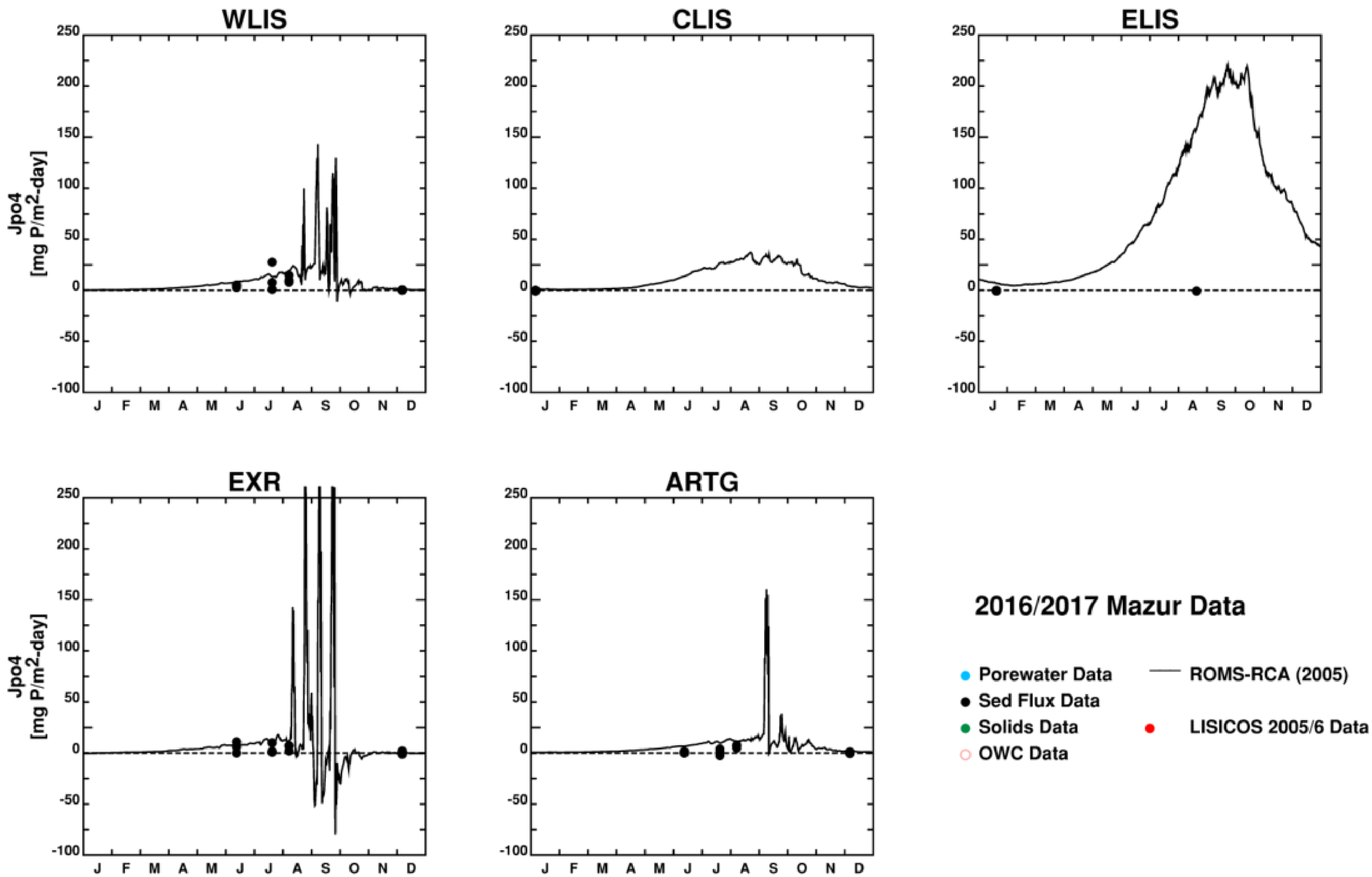

Figure 68. Model versus Data Comparisons for Sediment Phosphate Fluxes in 2005

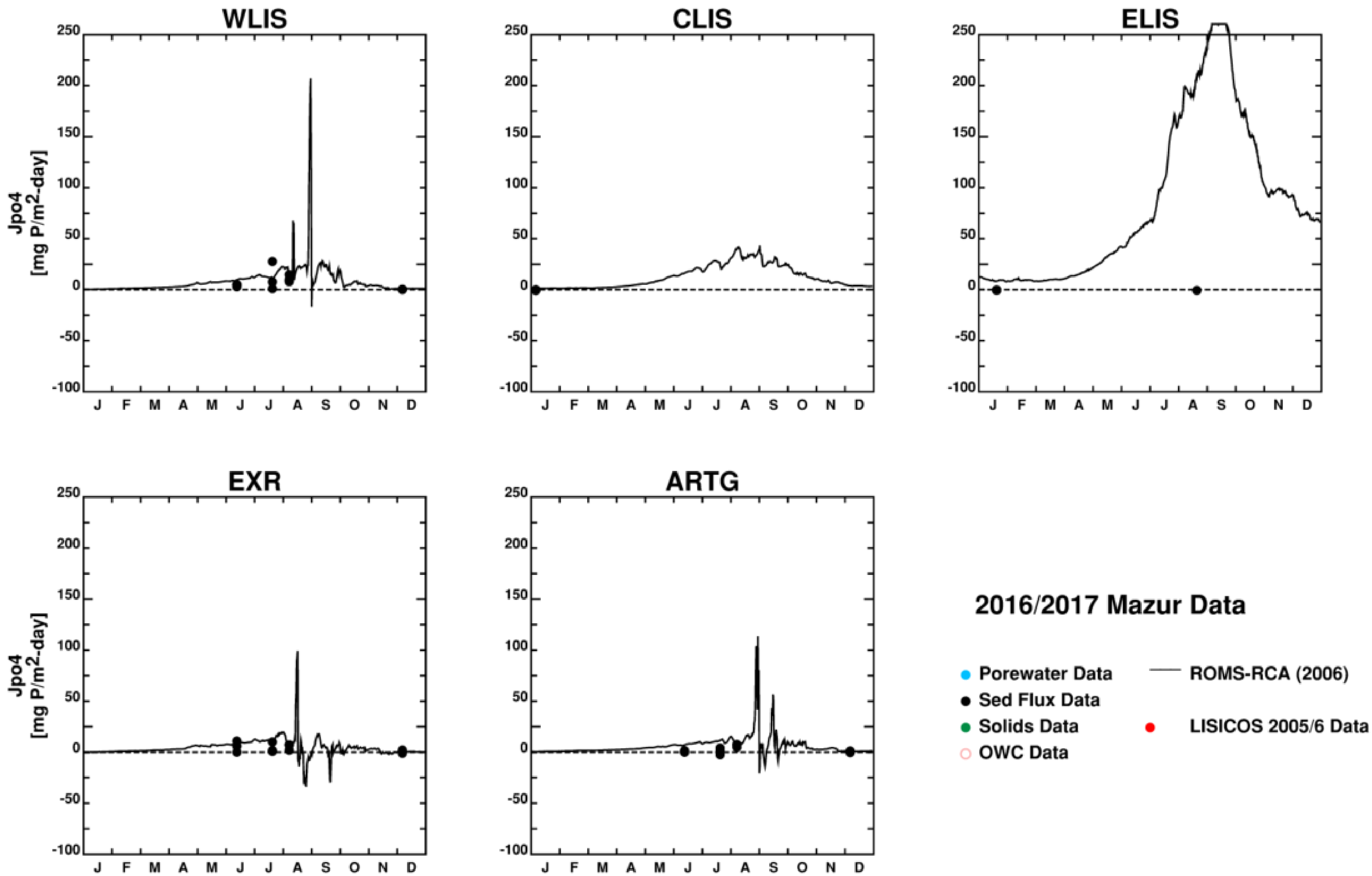

Figure 69. Model versus Data Comparisons for Sediment Phosphate Fluxes in 2006

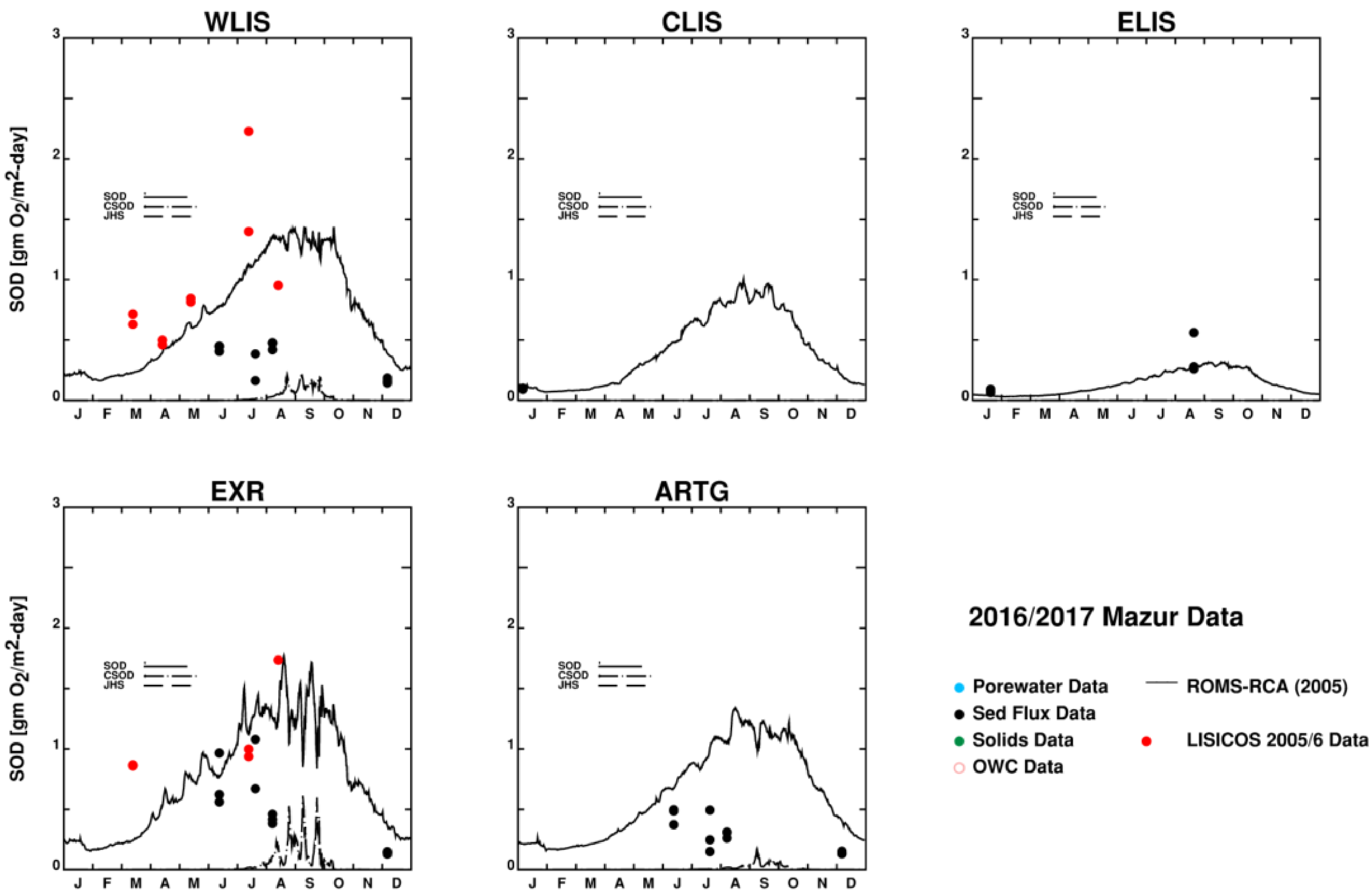

Figure 70. Model versus Data Comparisons for Sediment Oxygen Demand in 2005

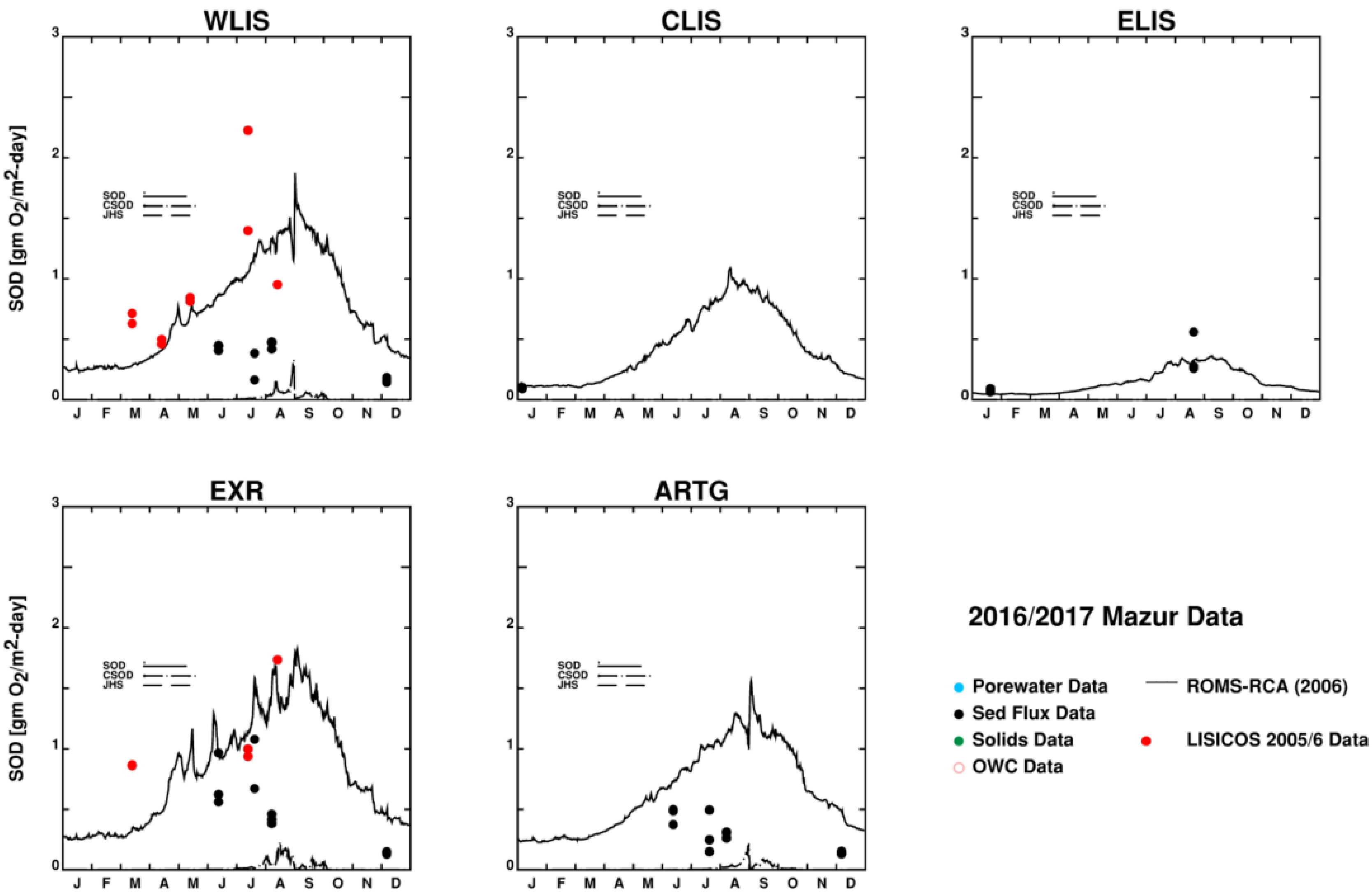

Figure 71. Model versus Data Comparisons for Sediment Oxygen Demand in 2006

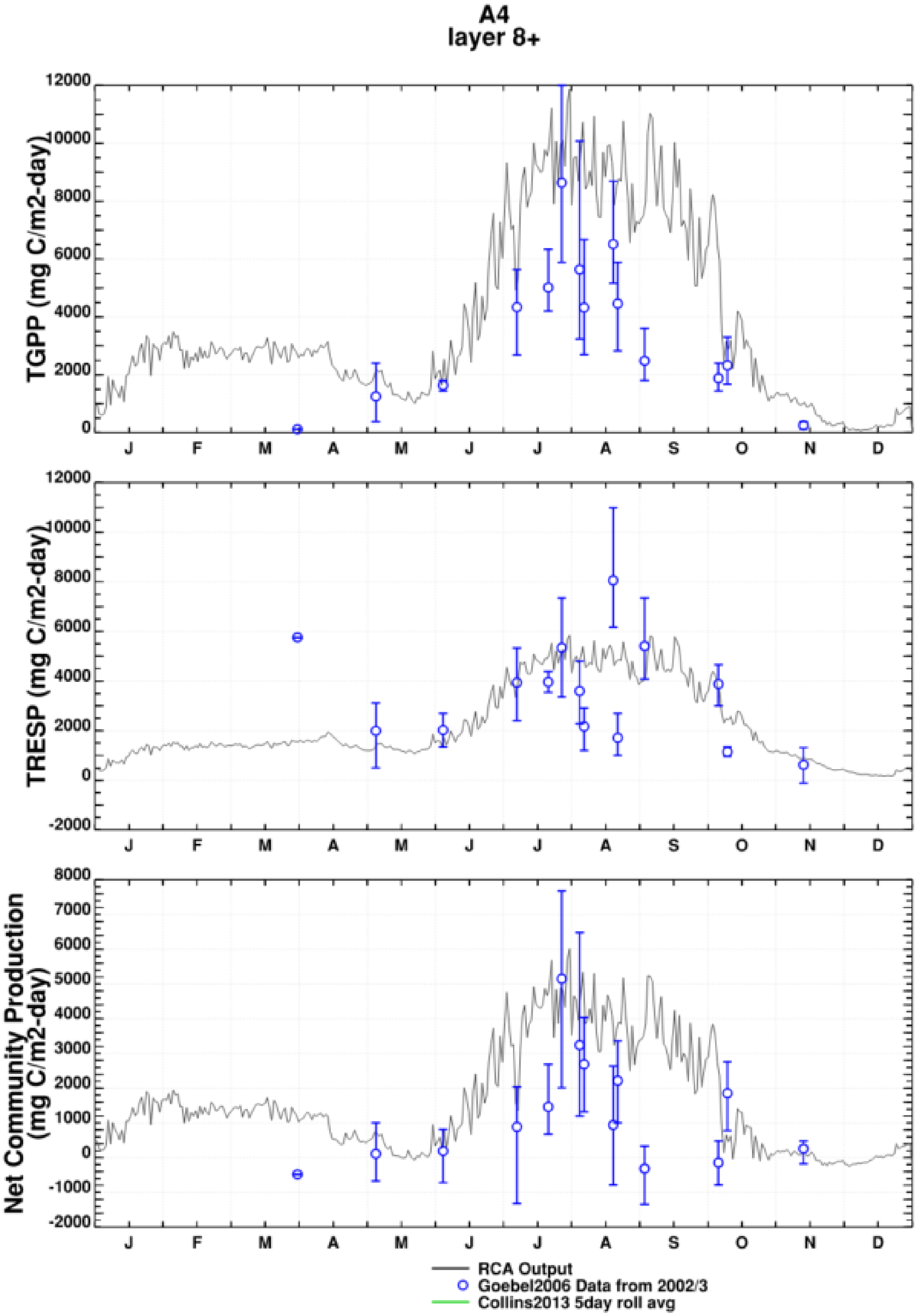

Figure 72. Model versus Data Comparisons for TGPP, R, and NCP at Station A4 in 2005

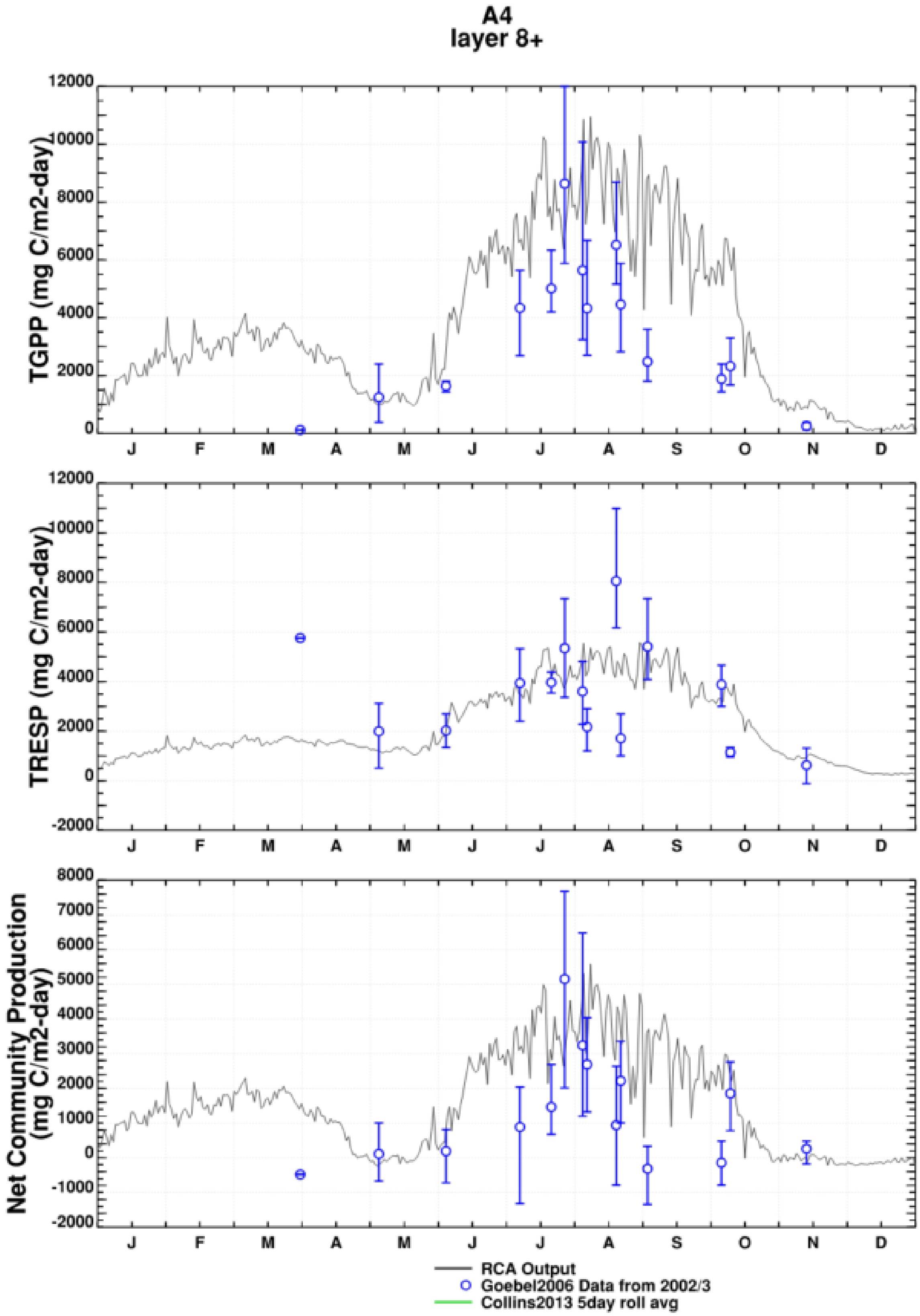

Figure 73. Model versus Data Comparisons for GPP, R, and NCP at Station A4 in 2006

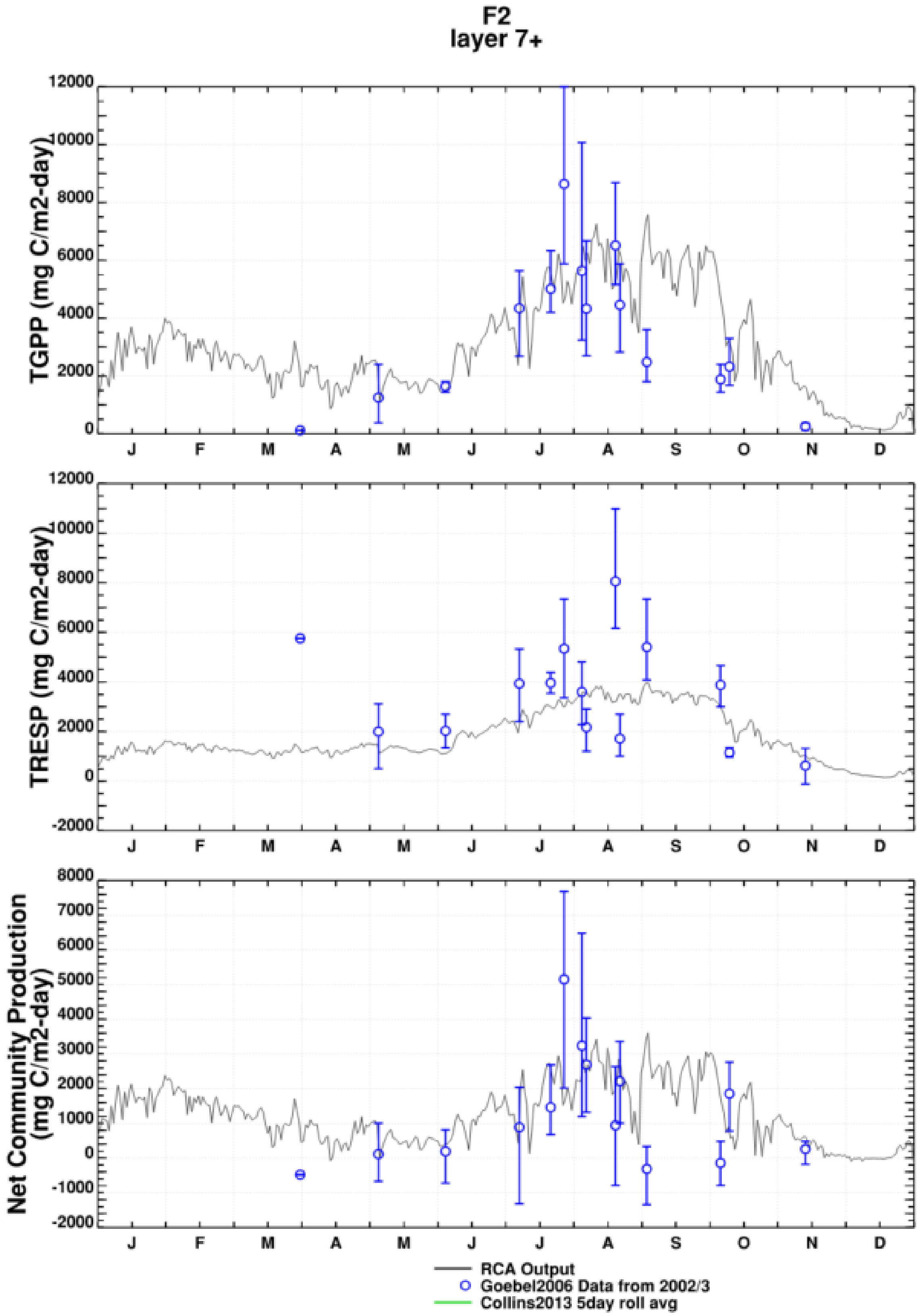

Figure 74. Model versus Data Comparisons for GPP, R, and NCP at Station F2 in 2005

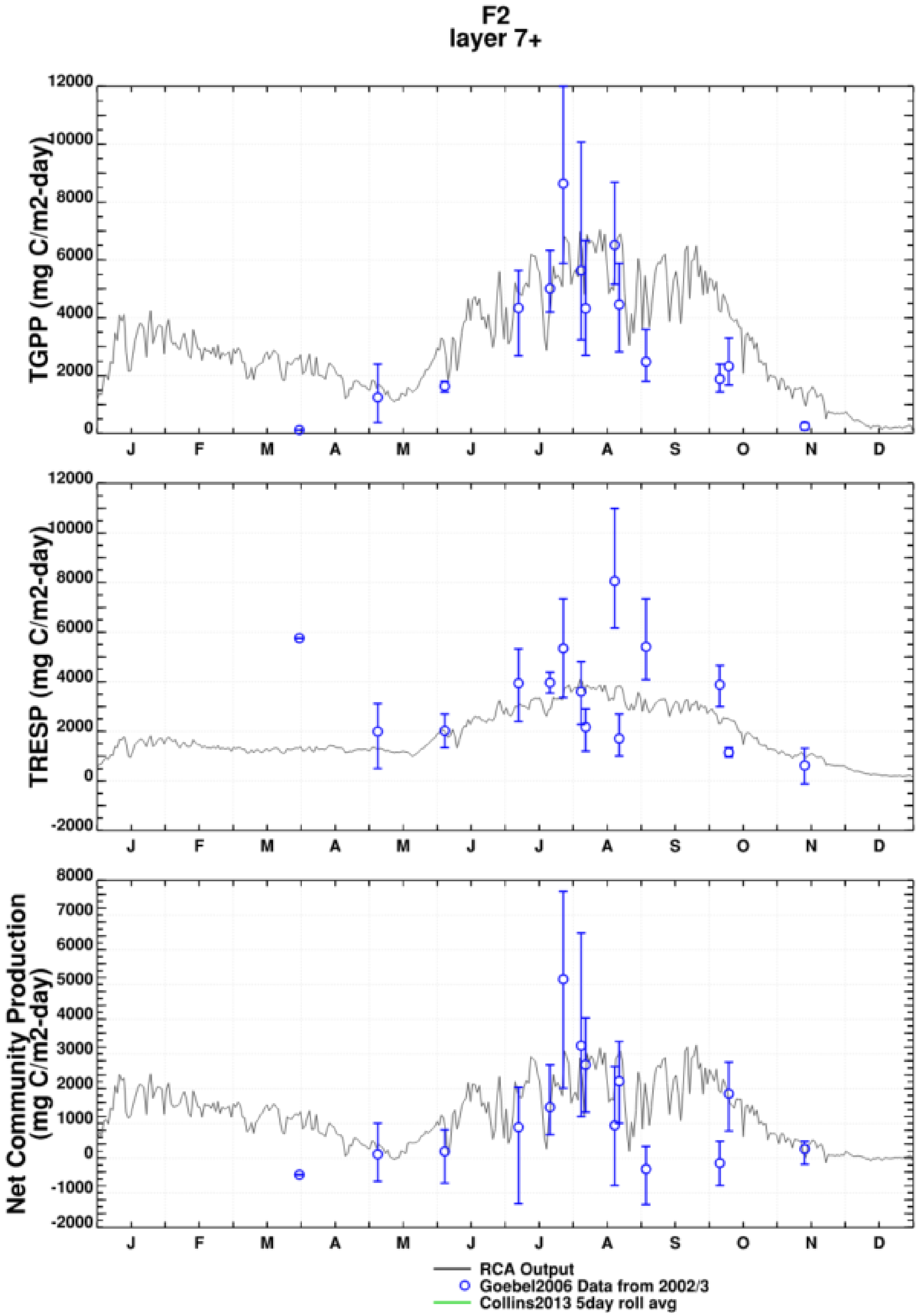

Figure 75. Model versus Data Comparisons for GPP, R, and NCP at Station F2 in 2006

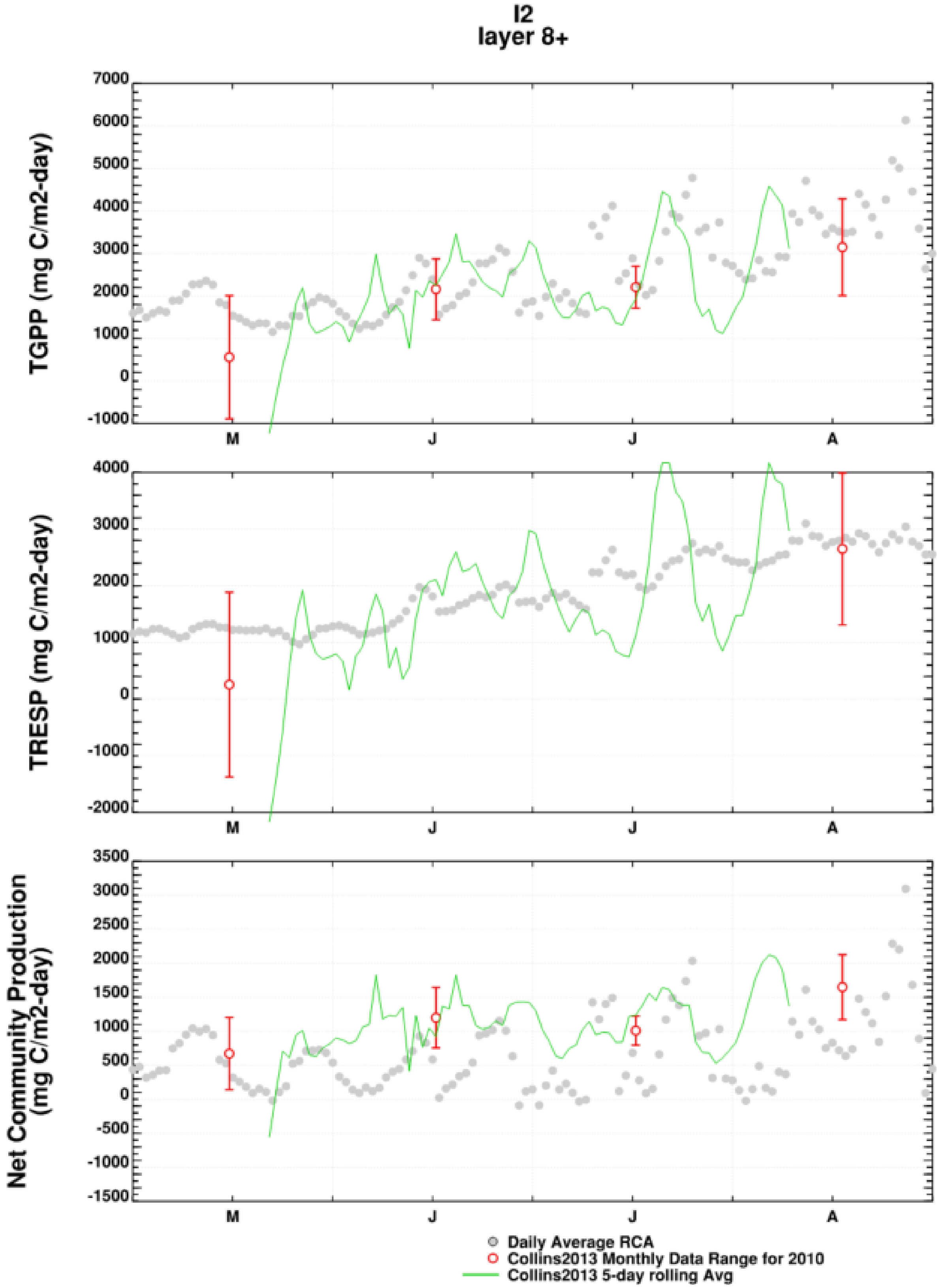

Figure 76. Model versus Data Comparisons for GPP, R, and NCP at Station I2 in 2005

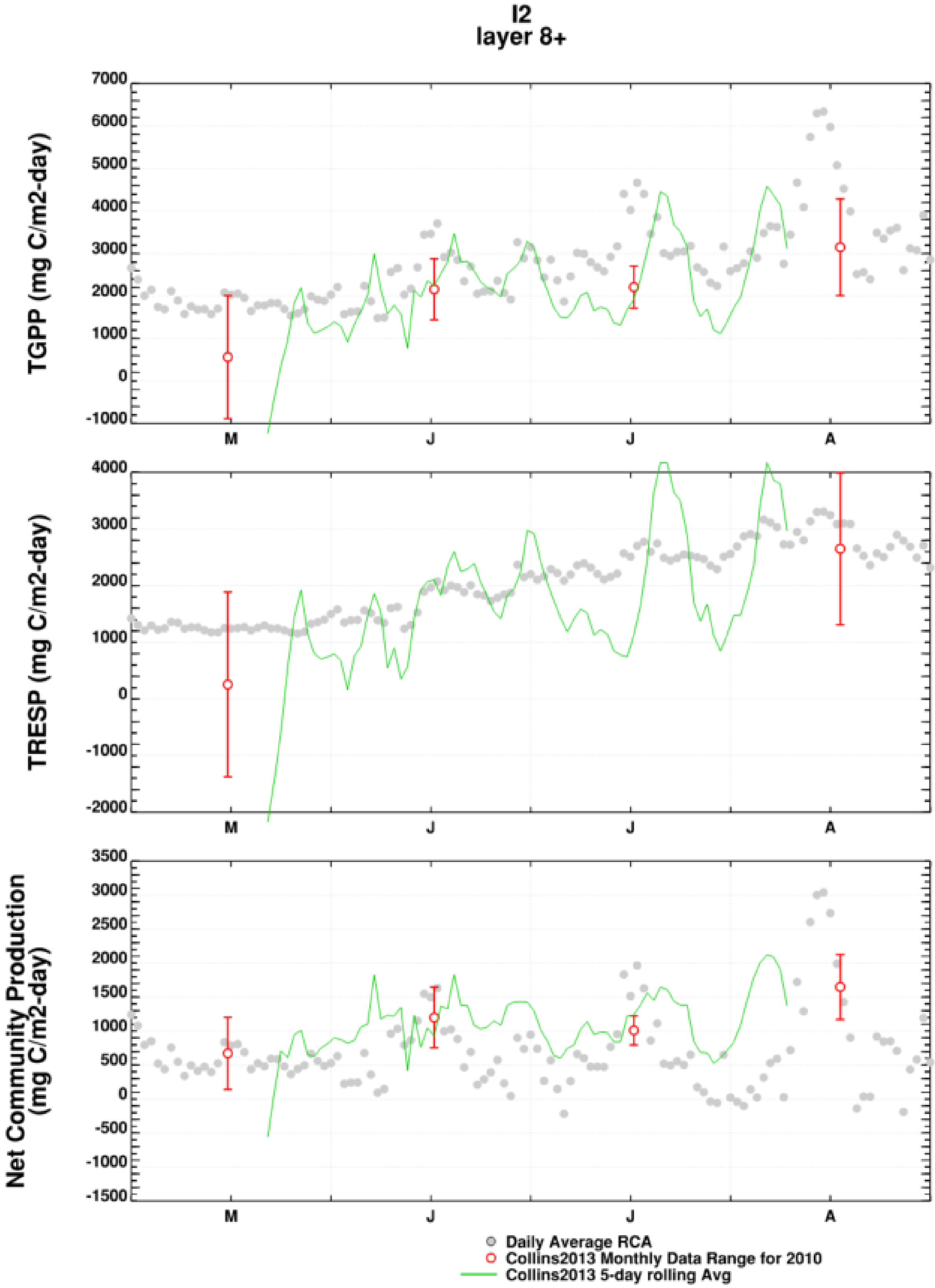

Figure 77. Model versus Data Comparisons for GPP, R, and NCP at Station I2 in 2006

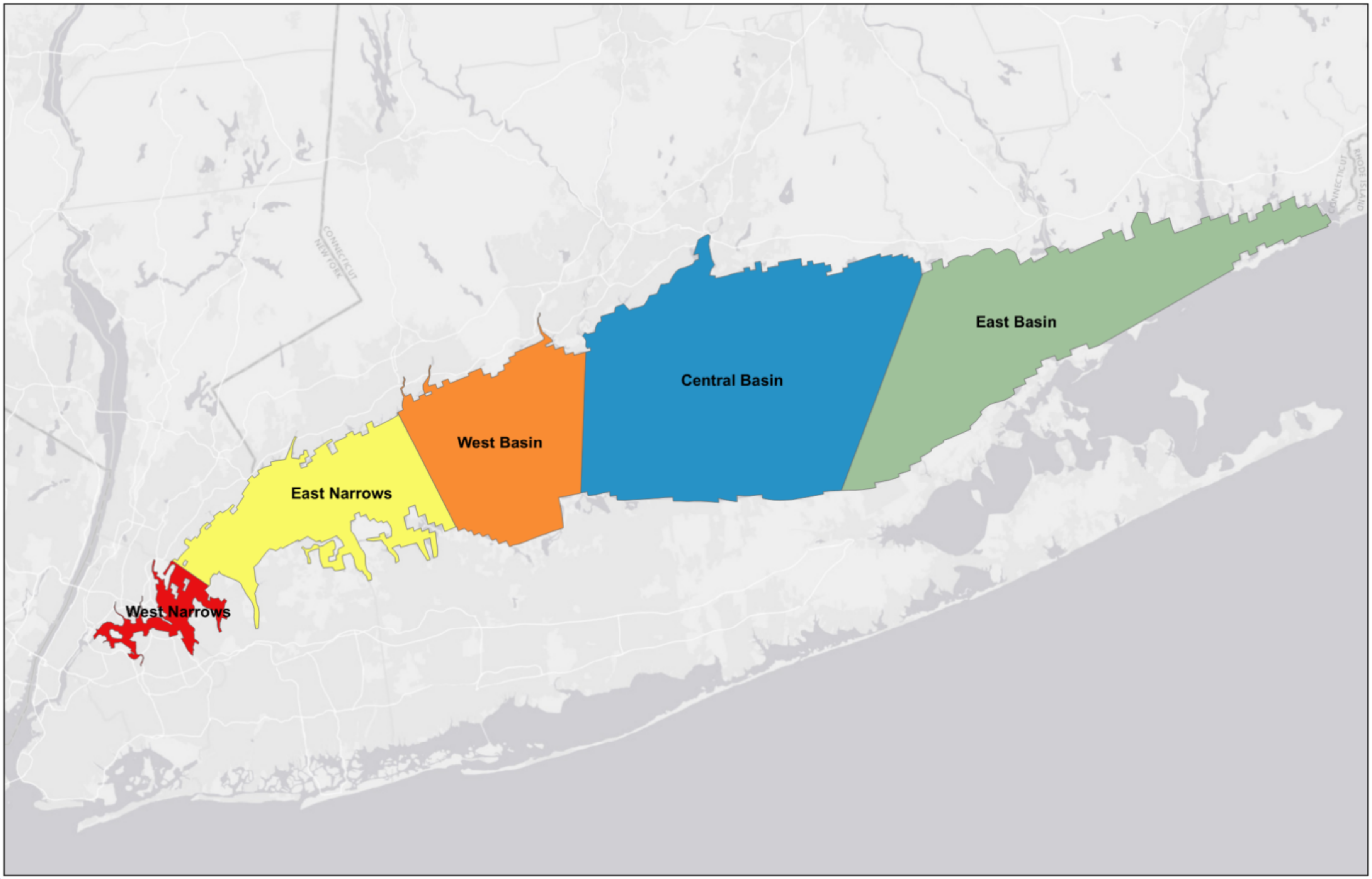

Figure 78. Areas Used for Grouping Stations for Statistical Analysis