MEMORANDUM

To: BPP Steering Committee & Stakeholders

From: Sarah Connick & Connie Liao

Cc: Jim Carleton, Terry Cooke, Betty Pun, Kirsten Rosselot

Re: The use of land cover and land use data as data inputs and its impact on BPP technical studies

Date: February 1, 2006

This memorandum summarizes the Brake Pad Partnership's approach for using specific land use and land cover data in the components of its effort where such data are needed, and provides the underlying rationale in support of each recommendation.

Background

The Brake Pad Partnership is conducting three interrelated environmental transport and fate modeling studies to understand the role of copper in automobile brake pad wear debris on short- and long-term concentrations of copper in the South San Francisco Bay, including an air deposition model, watershed model, and bay model.

The air deposition modeling study will provide information on the amount of copper from brake wear debris and other airborne sources of copper (i.e., industrial emissions) deposited spatially within the study watershed (i.e, Castro Valley Creek watershed) and to the bay.

The watershed model will use the results of the air deposition model and estimates of other land-based sources of copper to determine the concentrations of copper in Castro Valley Creek. The watershed modeling effort will also extrapolate from the Castro Valley watershed to the entire San Francisco Bay area, using additional estimates of copper sources for the region, to model copper loadings from all land immediately surrounding the Bay.

The bay modeling effort will model short- and long-term concentrations of copper in the bay using a hydrodynamic model with inputs of copper from the air deposition and watershed models, and from estimates of copper loadings directly to the bay from point and nonpoint sources.

The Land Use and Land Cover Data Challenge

The major challenge the Partnership faces in carrying out these studies is the lack of perfect information. By necessity, we must rely on available information about which there are uncertainties due to limited numbers of samples, limitations on sample collection and measurement abilities, data collected for different purposes, and, in some cases, quite simply a lack of data. In meeting this challenge, the Partnership is seeking to identify and use the best possible data for use in each component of its work, while assuring consistency and quality across the entire effort.

This challenge has become most acute with regard to identifying the most appropriate land use and land cover (LULC)[1] dataset for use, which is addressed here.

Data Needs

Three components of the Partnership's work require LULC data: air deposition modeling, watershed modeling, and source estimation studies. Each of these three studies has very different needs for LULC data.

Air Deposition Model. The air deposition model only requires LULC data to calculate roadway surface area as a fraction of the land surface area in both the local and regional models. This information is needed to estimate the fraction of deposited material that may be subject to reemissions from roadways due to passing vehicles and wind. None of the other inputs to the air deposition model are based on LULC data.

Watershed Model. The watershed model requires LULC data to divide the Bay watershed into two land "types": "pervious" and "directly connected impervious area" (DCIA). The watershed model also relies indirectly on LULC data used to develop the estimates of nonbrake copper sources to the watershed.

Estimates of Nonbrake Sources of Copper. The estimates of nonbrake sources of copper requires LULC data for (1) estimating roof areas for architectural sources of copper, and (2) for apportioning to the modeled subwatersheds emissions of copper from agricultural applications of pesticides, agricultural algaecide treatment of surface waters, agricultural fertilizers, and industrial runoff.

Existing LULC Data Sources

The following summarizes the LULC datasets available for use in the South San Francisco Bay watershed that are relevant to the Brake Pad Partnership's work.

USGS National Land Cover Data (NLCD). The NLCD datasets are produced by the US Geological Survey using satellite data collected at a 30-meter resolution. The dataset is grid-based, uses a modified Level II Anderson Land Use Classification System[2] that includes 21 land

---

[1] Although the terms land cover and land use is often used interchangeably, their actual meanings are quite distinct. Land cover refers to the surface cover on the ground, whether vegetation, urban infrastructure, water, bare soil or other. Land use refers to the purpose the land serves, for example, recreation, wildlife habitat, or agriculture.

It is important to distinguish this difference between land cover and land use, and the information that can be ascertained from each. The properties measured with remote sensing techniques relate to land cover, from which land use can be inferred, particularly with ancillary data or a priori knowledge.

Source: http://www.ccrs.nrcan.gc.ca/ccrs/learn/tutorials/fundam/chapter5/chapter5_18_e.html

[2] In 1976, the U.S. Geological Survey published a numerical land-use classification system designed to meet the varied needs of federal, state, regional, county and local governments (Anderson et al., 1976; paper available at: http://landcover.usgs.gov/pdf/anderson.pdf.).

cover categories. For additional information on the NLCD land categories, see: http://landcover.usgs.gov/classes.asp. The first NLCD dataset was completed for the entire country in 1996 and is based on satellite data collected circa 1992—it is referred to as the 1992 NLCD dataset.

The 2001 NLCD dataset is based on satellite data collected circa 2001. The dataset was slated for completion nationwide in 2005; however, components of it are currently available for select areas. The 2001 NLCD dataset has several enhancements over the 1992 dataset through the generation of additional land-cover attributes, including percent tree cover, and percent impervious surface cover data layers. Percent impervious and forest cover data layers are currently available for the San Francisco Bay Area; however, the land cover data layer is not yet available. For additional information on the development of this database, see: http://landcover.usgs.gov/pdf/NLCD_pub_august.pdf.[3]

ABAG Existing Land Use in 2000.[4] The existing land use data for 2000 was published by the Association of Bay Area Governments (ABAG) in 2002 and is available for nine counties. It is not a raster-based dataset. The dataset uses a Level IV Anderson Land Use Classification System—the first three levels are standard Anderson Land Use Classifications, the fourth level was developed by ABAG in collaboration with the U.S. Geological Survey.[5]

The land use data that is common for all counties is derived from a combination of aerial photography, interviews with city and county personnel, and field checking. Urban land use data are aggregated from parcel data, and non-urban land uses are derived from aerial imagery at approximately 10 meters resolution. In addition, assessors' data—i.e., ownership, not land use information—have been incorporated for San Francisco, San Mateo, Marin, Napa, and Contra Costa counties. Assessors' data are not incorporated for Alameda, Santa Clara, Sonoma, and parts of Solano counties. Data for some non-urban areas have been augmented using the 1992 NLCD dataset.

Although the ABAG dataset uses a four-level classification scheme, because it was derived from a variety of data sources having varying degrees of resolution and detail, the classification scheme is not applied uniformly over the spatial extent of the dataset. In other words, a given

---

This "nested" classification system, four levels of numerical subdivision are used to express various levels of detail of map data. One digit subdivisions, such as "1, Urban or Built-Up Areas," are designed to capture data that could be obtained from satellite images and that might be of interest to federal and state agencies. Two digit subdivisions, such as "12, Commercial and Services," are designed to capture data that could be obtained from aerial photographs and that might be of interest to state or regional agencies. Three digit subdivisions, such as "123, Education," are designed to capture data that could be obtained from more detailed aerial photographs with some field checking and that might be of interest to county or regional agencies. Four digit subdivisions, such as "1231, Elementary and Secondary Schools" often require field checking and are of most interest to local governments. (Source: ABAG Land Use 2000 documentation.)

[3] Homer, Collin, Chengquan Huang, Limin Yang, Bruce Wylie and Michael Coan, August 2005. Development of a 2001 National Landcover Database for the United States. SAIC Corporation, USGS/EROS Data Center, Sioux Falls, SD 57198. Corresponding author: Collin Homer, 605-594-2714; homer@usgs.gov.

[4] According to Jeanne Perkins at ABAG, the complete dataset for the region could be made available to the Brake Pad Partnership for $5,000. Portions of the dataset could be made available for lesser amounts.

[5] Napier and others, 1992 as cited in Perkins, J.B., Chuaqui, B., and Smith, K., 2002. *Existing Land Use in 2000: Data for Bay Area Counties*: Association of Bay Area Governments, Oakland, CA, 42 pp.

location on the ground that could potentially be assigned a fourth-level classification (say class 1231 "Elementary and Secondary Schools") might be assigned only a first- or second-level classification (e.g., class 12 "Commercial and Services") due to the limitations of that location's source data. Meanwhile an identical site in a different location may have the full fourth-level classification. As a result, the potential exists for underestimating the area of the more detailed land use classifications.

County Assessors Data. Assessor data and GIS basemaps showing each parcel are available for all 9 Bay Area counties. The costs to obtain the data and the data quality vary from county to county.

Topologically Integrated Geographic Encoding and Referencing (TIGER). Although not technically LULC data, TIGER® is the system and digital database developed by the U.S. Census Bureau to support its mapping needs for the Decennial Census and other Bureau programs. The TIGER files are a digital database of line-based geographic features, such as roads, railroads, rivers, lakes, legal boundaries, census statistical boundaries, etc. covering the entire United States. The database contains information about these features such as their location in latitude and longitude, the name, the type of feature, address ranges for most streets, the geographic relationship to other features, and other related information. They are the public product created from the Census Bureau's TIGER database. The most recent version is the 2003 TIGER/Line® Files.

ESRI. Although not technically LULC data, ESRI's road databases contain shape files and include information on segment length, road-type (e.g., interstate, state highway, and surface street), and road name. Some ESRI datasets do not include all roads.

Alameda County Road Database. Alameda County maintains a line-based database of publicly maintained roads within the unincorporated areas of the county for the purpose of projecting road maintenance needs. Road information in this database includes segment length; type of surface, base, and subbase; shoulder/curb type; rights of way; dates of improvements; and average daily traffic. Average traffic count data are distilled from traffic counts, which are conducted periodically to estimate rates of wear and tear on pavement. These data are not spatially referenced and thus are not available as a GIS layer.

LULC Data Sources for the Brake Pad Partnership—Discussion and Recommendations

Although it would be desirable from a consistency standpoint to use the same LULC dataset in each component of the Partnership's work, no single LULC dataset meets the different needs of each of the models and estimation efforts. The following includes a discussion of the advantages and disadvantages of the available LULC datasets in the context of each model and in relation to the overall Brake Pad Partnership effort. It also includes a recommendation for which data to use in each case, along with a listing of additional work that is being performed to improve our understanding of the underlying data.

Air Deposition Model

Discussion. The sole LULC data input to the air deposition modeling effort is roadway surface area as a fraction of the land surface area in both the local and regional models. This information could be extracted from each of the available LULC datasets with varying levels of precision.

Roadway surface area could be estimated from the NLCD dataset, by assuming that a fraction of the land use type “transportation/commercial/industrial” comprises road surfaces. Given the 30-meter resolution of these data, however, they are likely insufficient to account for local roads, which are much narrower than 30 meters. This approach would only provide a lower bound estimate on road surface areas.

Roadway surface area could also be estimated from the ABAG dataset in those counties for which assessors’ data are available. Assessors’ data show ownership parcels, but not roads and public rights of way (e.g., road shoulders or sidewalks, drainage ditches, landscaping and street trees, medians and islands, etc.). One could estimate roadway surface area by assuming a certain fraction of the areas that are not private parcels to be roads. The ABAG dataset lacks assessor’s data for Alameda County, in which the Castro Valley watershed is located, as well as for Santa Clara County and unincorporated areas of Sonoma and Solano counties.

Roadway surface area can also be calculated from one of the TIGER or ESRI datasets, which provide information on road lengths. Multiplying road lengths by standard road widths would yield roadway surface area. Use of a Geographic Information System (GIS) to perform this calculation would allow one to account for curves, further increasing the accuracy of the estimates. The TIGER dataset does not distinguish by types of road, and thus a weighted average width for all road types would need to be developed to conduct this calculation. The ESRI data set distinguishes among road types and is the most current roadway dataset. This approach would require the development of estimates of road widths for each road type.

Roadway surface can similarly be calculated from the Alameda County road database. This database is only available for Alameda County, however, and would not support the regional air deposition model.

Recommended Approach and Rationale. The Brake Pad Partnership opted to use the ESRI database and GIS to calculate roadway surface area as a percentage of land surface area for the air deposition local and regional models. Of the options available, the ESRI derived data are preferred because they (1) are specific to roads and thus provide the most direct measurement of roadway surface area, (2) distinguish among road types, providing the opportunity to develop a more accurate estimate, (3) are available for the entire Bay Area watershed, and (4) are readily available. Use of the ESRI data is expected to result in greater precision in the air deposition modeling results with respect to the material reemitted from roadways by passing vehicles and wind.

The Brake Pad Partnership is conducting an evaluation of the ESRI data to identify the percent error and determine whether that error is random or systematic. If the investigation identifies meaningful systematic error (e.g., >10%), the air deposition modeling should account for it using a data adjustment factor. The following analyses are underway:

- Comparison of road width estimates used in the road area calculations for the air deposition model to actual road widths for a statistically significant sample of road segments in Castro Valley, by either (1) using road width data from Alameda County Department of Public Works, or (2) using Alameda Countywide Clean Water Program (ACCWP) GIS data for roads in the watershed.
- Check to see if the ESRI data include all roads, using Castro Valley as a test case, using a GIS overlay provided by ACCWP for the comparison.
- Comparison of road areas estimated from the ESRI data to road area estimates that have already been prepared for other subwatershed areas, to the extent such data are available in appropriate formats.

Watershed Model

Discussion. The direct LULC data input to the watershed model is the fraction of each of the modeled subwatersheds for the local Bay watershed that is "directly connected impervious area" (DCIA) versus other impervious area and "pervious area." The watershed model also relies indirectly on LULC data for source loading terms of copper from nonbrake sources. The relative merits of the available LULC data for direct use in the watershed model are discussed in this section; the indirectly used data are discussed under "Estimates of Nonbrake Sources of Copper" below.

DCIA is not a parameter in any of the LULC datasets available for the San Francisco Bay Area, and needs to be derived from existing LULC for the purposes of the watershed model. To estimate DCIA, one needs to first estimate the overall percentage imperviousness (obtained directly from 2001 NLCD, or derived based on land uses if using 1992 NLCD or ABAG data) and then apply a correction factor to the impervious area to estimate the portion of the impervious area that is DCIA.

DCIA can be estimated from total imperviousness, in turn estimated from the 1992 NLCD dataset by developing and applying a percent imperviousness "correction factor" for each of the land uses,[6] and then applying an overall correction factor to obtain DCIA.

DCIA can also be estimated from the percent imperviousness data layer in the 2001 NLCD dataset. The percent imperviousness data layer is derived from the 30-meter resolution Landsat data, in combination with 1-meter resolution IKONOS satellite data, and digital orthophoto

---

[6] For example, EPA's ATtILA software uses such correction factors to estimate percent imperviousness from land use. Similarly, the TR55 manual (the USDA's curve number method for urban areas), lists conversion factors derived in the 1970s to estimate total imperviousness from nominal land use categories.

The USDA Curve Number method is a simple and widely used method for determining the approximate amount of runoff from a rainfall event in a particular area. Although the method is designed for a single storm event, it can be scaled to find average annual runoff values. The curve number is based on the area's hydrologic soil group, land use, treatment, and hydrologic condition.

quadrangle (DOQ)[7] images from aerial photography. Comparison of predicted percent imperviousness with measured percent imperviousness in three areas of the U.S. gave errors ranging from 8.8 to 11.4 percent.[8] DCIA can then be estimated using a single correction factor to estimate the portion of the impervious area that is DCIA.

DCIA can also be estimated from the ABAG dataset. With its four levels of land use coding, the ABAG dataset contains a wealth of detailed land use information. For example, non-mixed-use residential is divided into five categories from less than 1 unit per acre to more than 8 units per acre, and mobile home parks. This greater detail in land use description is driven by ABAG's need for data for economic modeling, however, and adds only indirect value for environmental parameters such as percentage imperviousness. This greater detail is not applied uniformly throughout the dataset, and is inconsistent from county to county. As with the 1992 NLCD dataset, DCIA can be derived from the ABAG dataset by calculating percent imperviousness using a correction factor for each of the land use categories,[9] and then applying a single correction factor to estimate the portion of impervious area that is DCIA.

DCIA cannot be estimated from the TIGER, ESRI, or Alameda County Road datasets.

Recommended Approach and Rationale – Total impervious surface area estimate. The Brake Pad Partnership recommends using the 2001 NLCD dataset to estimate total impervious surface area for the watershed model.

Use of any available dataset will require some estimation method to derive DCIA. Estimating DCIA from either 1992 NLCD or 2001 ABAG data would require first the application of appropriate correction factors for each land cover or land use category. The 2000 ABAG dataset is more current than the 1992 NLCD dataset. Although the ABAG data are coded to a much higher level of specificity of land use, that additional information does not directly translate toward a better understanding of imperviousness, and would require extensive development of appropriate correction factors. When asked about the utility of ABAG data for deriving percentage impervious surface, two separate sources familiar with the dataset[10] suggested the Partnership would likely be better off using the 1992 NLCD dataset. Neither of these individuals was familiar with the 2001 NLCD dataset on percentage imperviousness.

The 2001 NLCD data are more up-to-date than the 1992 data, and provide the percentage imperviousness data layer. The NLCD data have flaws—for example, since these data have a

---

[7] The USGS scans photos, almost exclusively from the National Aerial Photography Program (NAPP), an interagency effort coordinated by the USGS for map revision and a variety of other purposes. Each of these photos covers an area a little more than 5 miles on a side, representing about one-fourth of a standard, 7.5-minute USGS topographic quadrangle map. Source: http://geography.usgs.gov/digitalbackyard/faqsnew.html.

[8] Limin Yang, Chengquan Huang, Collin G. Homer, Bruce K. Wylie and Michael J. Coan 2005. An approach for mapping large-area impervious surfaces: Synergistic use of Landsat 7 ETM+ and high spatial resolution imagery. Manuscript submitted to Canadian Journal of Remote Sensing (research paper). Corresponding author: Limin Yang, USGS/EROS Data Center, Sioux Falls, South Dakota 57198. Tel: (605) 594-6039 E-mail: lyang@usgs.gov.

[9] No such correction factors are now available for ABAG's four-digit level of land use coding and it would be a significant time and resource commitment to develop them.

[10] Jeff Kopellas, San Francisco Regional Water Quality Control Board, and Jeanne Perkins, Association of Bay Area Governments.

30-meter resolution they might have difficulty accounting for local roads. Tree coverage may cause impervious area (e.g., tree-lined roads) to be mis-classified. Since roads typically account for a significant fraction of DCIA, it is possible that the NLCD data underestimates total impervious area. Based on best professional judgment, the 2001 NLCD dataset is likely to provide the best accuracy within the spatial scale of the watershed model, and would entail the least amount of additional estimation and data manipulation.

The Brake Pad Partnership is evaluating the 2001 NLCD data to identify the percent error, determine whether that error is random or systematic, and if the error is systematic, determine how it can be accounted for in the watershed modeling. If the investigation identifies meaningful systematic error (e.g., >5%), the watershed modeling should account for it using a data adjustment factor. The following evaluations are underway:

- Comparison of the 2001 NLCD percentage imperviousness estimates to previous estimates for Santa Clara County Watersheds, San Mateo County Watersheds and Alameda County Watersheds developed by the Santa Clara Valley Urban Runoff and Pollution Prevention Program, the San Mateo Countywide Stormwater Pollution Prevention Program, and ACCWP.

Recommended Approach and Rationale – DCIA estimate. The Brake Pad Partnership recommends using a correction factor for the NLCD impervious area estimate to estimate DCIA for the watershed model. Estimating DCIA from the NLCD dataset will require the use of a correction factor for the percentage of impervious surface that is directly connected to the storm drainage system, including creeks.

Based on best professional judgment, the 2001 NLCD dataset is likely to provide the best accuracy within the spatial scale of the watershed model, and would entail the least amount of additional estimation and data manipulation.

The Brake Pad Partnership is evaluating a proposed DCIA estimation method to identify the percent error, determine whether that error is random or systematic, and if the error is systematic, determine how it can be accounted for in the watershed modeling. If the investigation identifies meaningful systematic error (e.g., >5%), the watershed modeling should account for it using a data adjustment factor. The following evaluations are underway:

- A review of previous methods for DCIA estimation, the proposed procedure and the basis for its selection, and comparison of previously used methods for estimating DCIA with the proposed method using the 2001 NLCD data, including:
    - Comparison of estimated DCIA to estimated road area from ESRI database and URS analysis (for Castro Valley only and bay wide).
    - Comparison of estimated DCIA to the “inverse area” from assessor parcel data for readily available watersheds (i.e., that area in a county that is not included in parcels but is primarily roads and therefore primarily DCIA).
    - If data are available and different than previous data sets, comparison of the total estimated road area to the total estimated road area in the Bay watersheds provided by others (e.g., Jeff Kapellas) based on ABAG’s 2000 data set.

Estimates of Sources of Copper

Discussion. Estimates of copper loadings to each modeled subwatershed within the San Francisco Bay watershed are required inputs to the watershed model.[11] With the exception of copper in domestic water discharged to storm drains, whose estimates are directly dependent on population within the modeled subwatersheds, the Partnership's approach is to apportion estimated emissions to each modeled subwatershed either according to population or land use, depending on the source in question. LULC data are used to apportion the following sources of releases of copper to the modeled subwatersheds based on relative amounts of land use:[12]

- architectural sources of copper,
- agricultural applications of pesticides, including algaecidal applications to surface waters,
- agricultural applications of fertilizer, and
- industrial runoff.

Architectural releases of copper require special mention. The methodology for estimating releases of copper from architectural sources relies in part on a value for the fraction of land covered by roofs. The fraction of roof cover was developed based on ABAG's definitions for residential and commercial/industrial/institutional land use.

U.S. Census data from 2000 were used to estimate releases of copper in domestic water discharged to storm drains and to apportion population-based sources (copper released from pesticides applied to urban land, pool, spa, and fountain algaecides, copper leached from pressure-treated lumber, copper in nonagricultural fertilizer applications, and copper in algaecides used in nonagricultural rights of way, recreation areas, and public health) to the modeled subwatersheds.

In the counties where 2001 ABAG data are available, either the 2001 ABAG or the 1992 NLCD dataset could be used for performing these calculations. The advantage of the 2001 ABAG dataset is that it is more recent. It also contains land use classifications at a much higher level of specificity of land use, although the greater specificity does not directly translate into a better

---

[11] The bay model also uses estimates of direct emissions to the bay from algaecide treatment of surface waters adjacent to the Bay and marine antifouling coatings, but these estimates are not influenced by LULC data. Non-brake air releases of copper are small and also are not influenced by LULC data.

[12] a) For architectural copper, land use data was used on residential and commercial/transportation/industrial surface area in my estimates of roof area.
b) For agricultural applications of pesticides, county-based data were apportioned to the subwatersheds based on agricultural surface area (crops only). Since total agricultural area is such a small percentage of the Bay Area total, no futher refining the estimates for agricultural pesticides is needed.
c) For algaecide treatment of surface waters, agricultural uses of surface water algaecides were apportioned to the subwatersheds according to agricultural (crop) land use area.
d) For fertilizers, state-based agricultural data were apportioned to the subwatersheds based on agricultural (crop) land use area.
e) For industrial runoff, estimates of copper releases are based on commercial/transportation/industrial land use area within the subwatersheds.

understanding of sources of copper from these land uses. In addition, the greater detail is not applied uniformly throughout the dataset, and is inconsistent from county to county.

An important consideration is whether use of the more recent data would decrease the uncertainties associated with the overall estimates of copper sources. For those sources that are apportioned based on land use, more accurate land use information would allow for greater accuracy in the apportionment of sources among the modeled subwatersheds, however, it would not improve the accuracy of the total loadings within the entire watershed.

Recommended Approach and Rationale. The Brake Pad Partnership has opted to use the 1992 NLCD dataset to apportion to the modeled subwatersheds emissions of copper from agricultural applications of pesticides, agricultural algaecide treatment of surface waters, agricultural fertilizers, and industrial runoff, based on relative amounts of land uses. For estimating roof areas for architectural sources of copper, the Brake Pad Partnership has opted to use Bay Area wide ABAG data to calculate copper releases from architectural features, and 1992 NLCD data is used to apportion the releases to the modeled subwatersheds. Since the roof cover fractions were developed based on the Bay Area wide ABAG data set, it would be inappropriate to combine these roof cover fractions with a different land use data set.

The Brake Pad Partnership will further evaluate the potential impact of any reallocation of loads among watersheds that could have resulted from land use and population changes since 1992 using a sensitivity analysis when running the Bay model.

Discussion of Effects of LULC Data Source Recommendations on the Brake Pad Partnership's Technical Effort

In addition to considering which data are most appropriate for a specific element of the Partnership's work, it is important that the data selections support the overall goal of the project, which is to understand the role of copper from brake pads in short- and long-term concentrations of copper in the lower San Francisco Bay. This section aims to address how the use of different data sources recommended here may affect how the project components relate to one another and how that would affect the overall study results.

In an ideal world, all the data inputs would cover the same time periods. Road, land use, meteorological, air deposition, and water quality data for the air deposition and watershed models would be specific to each year modeled, and corresponding copper inventories, land use, and population data would be used to produce loading estimates by year. Given the reality of expense, time, and physical constraints, however, the Partnership needs to work with the data that are now available, and take into consideration how and when differences in the dates associated with the data used may affect the study results. In evaluating the uncertainty introduced by the use of different datasets, it is also important to take into account the degree of the uncertainty being introduced by a difference in time relative to the degree of uncertainty associated with other data used in the studies.

The key areas in which the time of original data collection affects the Partnership's studies are the air deposition modeling, watershed modeling, and estimates of sources of copper.

Air Deposition Modeling. The period for which the air deposition model is being run corresponds to the period for which the Brake Pad Partnership collected air deposition monitoring data in the Castro Valley Creek watershed—2003 to 2004. The meteorological data used in the air deposition model corresponds specifically to this time period. The road data from ESRI is the most currently available road data for the bay watershed.

Watershed Modeling. The watershed model relies on multiple data sources covering different timeframes. Air deposition inputs come from the air deposition model for 2003-2004. DCIA is derived from the NLCD land use data circa 2001. Water quality monitoring data and meteorological data used to calibrate and validate the model cover 1989 to 1993 for San Lorenzo Creek, 1989 to 1995 for Alameda Creek, and during the following periods 1989-1996, 2001-2002 and 2003-2004 for Castro Valley Creek.

Estimating Nonbrake Sources of Copper. Copper source estimates are based on 1992 NCLD land use data and 2000 Census data. Copper inventories used in these estimates derive from a variety of sources of data from the late 1990s through 2004.

Discussion

Differences in times of original data collection. Differences in the times that the original data used in this effort were collected have the potential to introduce unintentional biases into the study outcomes. With regard to the discussion here, this potential arises in three areas—(1) estimates of copper loads come from a variety of data sources from the mid-1990s through 2004; (2) changes in land use from 1992 to 2001 may cause the ratios of newly developing versus older urban areas to be misrepresented; and (3) the calibration and validation of the watershed model using land use and water quality monitoring data from different periods may not adequately account for land use changes.

While each of these areas are important to be aware of, the Brake Pad Partnership recognizes that the degree of uncertainty associated with the sampling data far outweighs uncertainty created by the time period differences among the data sets.

Apples and oranges? The Brake Pad Partnership's overall study effort draws on many types of data from multiple of data sources, due in large part to necessity. An important question arises, however, regarding the compatibility of these data sources and whether use of data from different sources may introduce an unintended bias in the project outcomes.

To address this concern, the Brake Pad Partnership intends to conduct a series of "reality checks" across the modeling efforts. The point of this effort is to make sure that the model results are consistent with the data inputs and with each other. The following "reality checks" will be conducted:

- Comparison of copper mass balances for each model. Copper mass should be conserved throughout the modeling effort. Copper mass balances will be checked for each model.