Brian R. Leahy
Director

MEMORANDUM

Edmund G. Brown Jr.
Governor

TO: Randy Segawa
Environmental Program Manager I
Environmental Monitoring Branch

FROM: Jing Tao, Ph.D.
Environmental Scientist
Environmental Monitoring Branch
916-324-4201

Original signed by

DATE: July 29, 2013

SUBJECT: TIME SERIES ANALYSIS AND FORECASTING OF VOLATILE ORGANIC COMPOUND EMISSIONS FROM NONFUMIGANT PESTICIDES USED IN SAN JOAQUIN VALLEY DURING THE OZONE SEASON

BACKGROUND

Under California's State Implementation Plan (SIP), the Department of Pesticide Regulation (DPR) must track and control Volatile Organic Compound (VOC) emissions from pesticides used by agriculture and commercial structural applications in five ozone nonattainment areas (NAAs). The San Joaquin Valley (SJV) NAA has a SIP goal of 18.1 tons/day during the ozone season (May – October), a 12% reduction from its VOC emissions in 1990.

Fumigant and nonfumigant pesticides are two groups of VOC contributors and have different emission calculation methods. While increased emissions were observed in some NAAs recently, DPR, growers, registrants, and others have taken steps to reduce emissions in these areas. However, the measures such as voluntary use of the reformulated products and fumigant regulatory restrictions may not be sufficient to meet the SIP goal for years with the highest pesticide use in the SJV NAA. Almost three quarters of SJV pesticide VOC emissions are derived from nonfumigants. Therefore, DPR has regulated to reduce VOC emissions from nonfumigant pesticides used in the valley. The regulation triggers a prohibition on most uses of high-VOC nonfumigant products applied to seven crops, when the SJV VOC emissions exceed 17.2 tons/day (95% of 18.1 tons/day) (Neal et al, 2013). One difficulty for the timely implementation of this regulation is that the prior year's VOC data are usually not available until late spring or early summer. Consequently the VOC emissions of the current year must be forecasted from the historical VOC data from two years prior because they are the most recent available.

This memorandum documents the development of a time series model to predict the monthly SJV nonfumigant VOC (VOC_{NF}) emissions. The prediction estimate is then used to evaluate the probability of the SJV VOC emissions exceeding 17.2 tons/day during the ozone season of 2012 – 2013.

DATA

The VOC emission data are estimated based on pesticide use reports complied by county agricultural commissioners and DPR every year. The monthly VOC_{NF} emissions from 2000 to 2011 (total 144 data points) are used for the time series analysis in this memorandum (Figure 1).

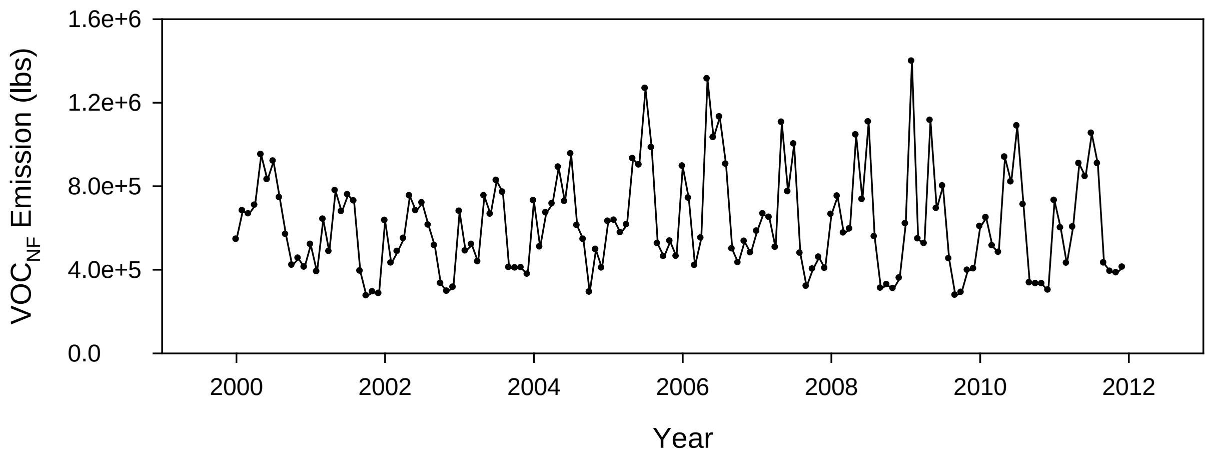

Figure 1. Monthly VOC_{NF} (lbs) of the San Joaquin Valley in 2000 – 2011

TIME SERIES MODEL

Time series modeling has been used to analyze and predict VOC_{NF} emissions of the Ventura NAA since 2009 (Tao, 2013). The model has the form:

$$X_t = m_t + s_t + y_t$$

where X_t is the monthly VOC_{NF};

m_t is the trend estimated from the linear regression of deseasonalized VOC_{NF} on t;

s_t is the seasonal component, monthly in this study with $\sum_{j=1}^{12} s_j = 0$. The detrended VOC_{NF} were averaged for each month over the analyzed time and then centered to obtain the estimate;

y_t is the residual series that are fitted with an autoregressive integrated moving average (ARIMA) process;

t is the year as time index.

The notation used to denote a specific seasonal ARIMA model is

$$\text{ARIMA}(p,d,q) \times (P,D,Q)_L$$

where p is the order of nonseasonal autoregressive component;
d is the order of nonseasonal differencing;
q is the order of nonseasonal moving average process;
P is the order of seasonal autoregressive component;
D is the order of seasonal differencing;
Q is the order of seasonal moving average process; and
L is the seasonal length.

The classical decomposition algorithm method is used to develop the time series model. With this method, the data is first smoothed using a moving average filter. Then the seasonal component is estimated and removed from the original data. The deseasonalized data is used to fit the potential linear trend. After removing the trend and the seasonal component, the residuals are fitted with ARIMA models. A proper ARIMA model should achieve normal distributed white noise residuals, a sequence of uncorrelated random variables with constant variance and mean zero. When predicting, ARIMA model will be used to calculate the prediction variance and one part of the prediction estimate. The fitted trend and seasonal component also estimate their own part of prediction. All three parts are then combined to be the final prediction for the monthly VOC_{NF} emissions of 2012 – 2013. Statistical software package R is used to process the modeling and prediction.

MODEL ESTIMATES AND DISCUSSIONS

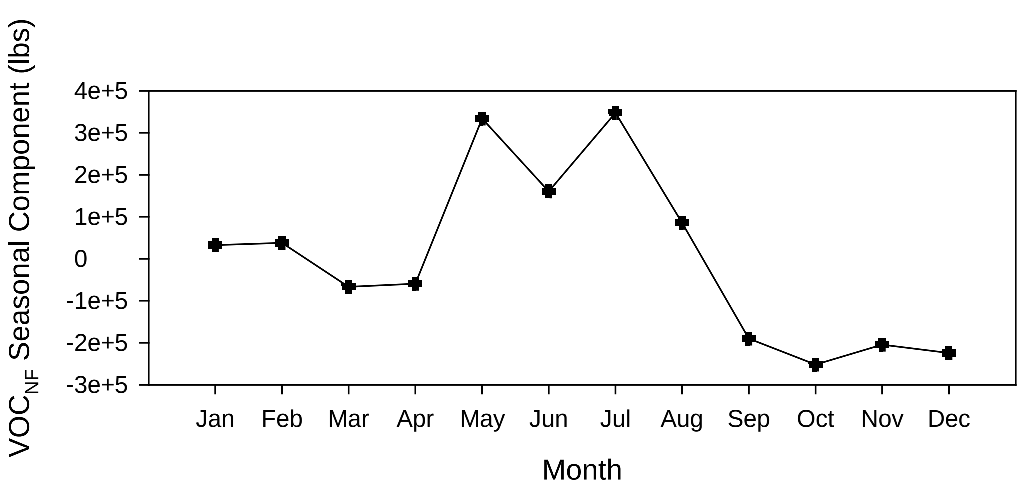

Figure 2. Seasonal component estimates (lbs) of the monthly VOC_{NF}

The general seasonal component $\{s_t\}$ is estimated first (Figure 2) and removed for the further modeling. As shown in Figure 2, May – August are the highest VOC_{NF} emission months. Then a linear regression model $\{m_t\}$ is estimated to fit the deseasonalized data:

$$m_t = -8920189 + 4755 \times t$$

A positive slope of the regression suggests an increasing trend of VOC_{NF} emissions in the valley. However, with the adjusted R^2 0.0095 and the P-value 0.126, this linear regression is not statistically significant. Hence the linear trend does not exist in the SJV VOC_{NF} emissions from 2000 to 2011.

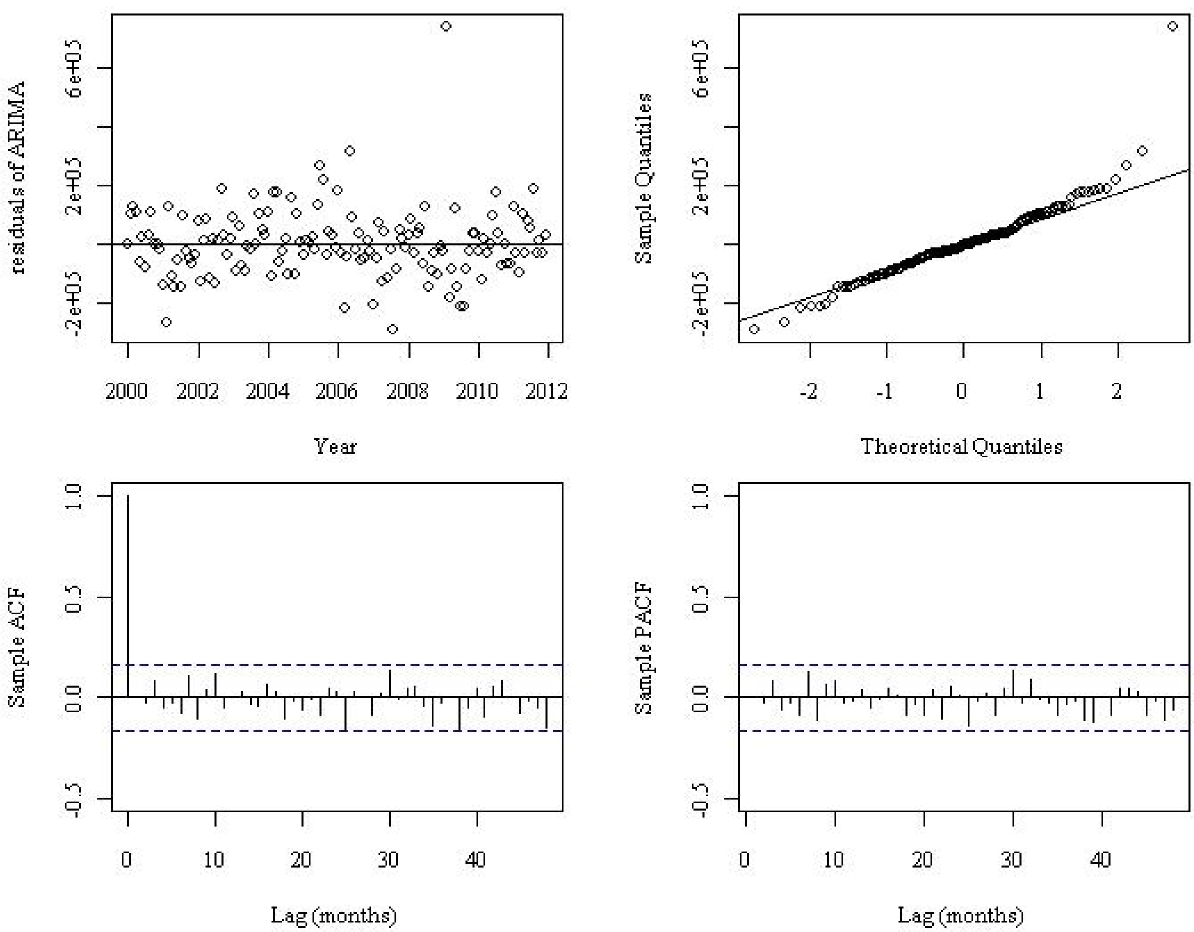

Figure 3. Residual plots of ARIMA(1,1,1) × (0,0,1)$_{12}$ for the deseasonal VOC_{NF}

Since the trend is not statistically significant, ARIMA model is fitted with the deseasonalized data, instead of the linear regression residuals. ARIMA(1,1,1) × $(0,0,1)_{12}$, is estimated and has the equation:

$$(1 - \varphi B)y_t = (1 + \Theta B^{12})(1 + \theta B)w_t$$

Where B is differencing term; $By_t = y_{t-1}$ and $B^{12}\, y_t = y_{t-12}$. The autoregression coefficient φ is 0.1579. The nonseasonal moving average coefficient θ is -0.8998. The seasonal moving average coefficient Θ is 0.1923. And w_t is Gaussian white noise residuals ~N $(0, \sigma^2 = 1.43\times10^{10})$.

An adequate ARIMA model should have the residuals close to white noise. Four ARIMA residual plots are shown in Figure 3 for model diagnostic checking. The top two plots present that the residuals are evenly distributed around mean 0 and very close to the normal distribution. The bottom two plots, sample ACF and PACF with dash lines showing 95% confidence interval, demonstrate that the correlations between time lags of residual series are not significantly different from 0. The Ljung-Box statistic is a function of the accumulated sample autocorrelation up to any specified time lag. It is used to test the "overall" randomness of time series model residuals based on a number of lags. Table 1 exhibits that all the calculated Ljung-Box statistics are non-significant, suggesting the independence of residuals within all the tested lags. All these tests indicate that the ARIMA(1,1,1) × $(0,0,1)_{12}$ has achieved independent random residuals and no further modeling is needed.

Table 1. Ljung-Box Chi-Square statistic of ARIMA(1,1,1) × $(0,0,1)_{12}$ residuals

Lag	12	24	36	48	72
Chi-Sqaure	8.65	14.06	28.98	45.46	65.58
df	9	21	33	45	69
P-Value	0.47	0.87	0.67	0.45	0.59

The time series analysis intends to discern the pattern of the data collected over time. The approach used here focuses on modeling future points of the time series data as a parametric function of the past values. The model solely estimates the correlation between the current data and its past values. It does not include any effect from factors such as unusual weather change, irregular crop rotation, recently updated agricultural practice, comparatively new pesticide regulation, and so on. These factors can directly or indirectly affect the pesticide use and then could significantly change VOC emissions in a short time. However, the correlations between these factors and emissions are poorly understood and cannot be used to model the VOC emissions currently. Also the time series analysis in this memorandum cannot forecast the future fluctuations of VOC emissions caused by these factors.

SJV VOC EMISSION FORECAST

The prediction estimates and standard errors of the ARIMA(1,1,1) × $(0,0,1)_{12}$ combine with the seasonal component to complete forecasting the monthly VOC_{NF} emissions for 2012 – 2013. The forecast results show similar pattern with the historical data (Figure 4).

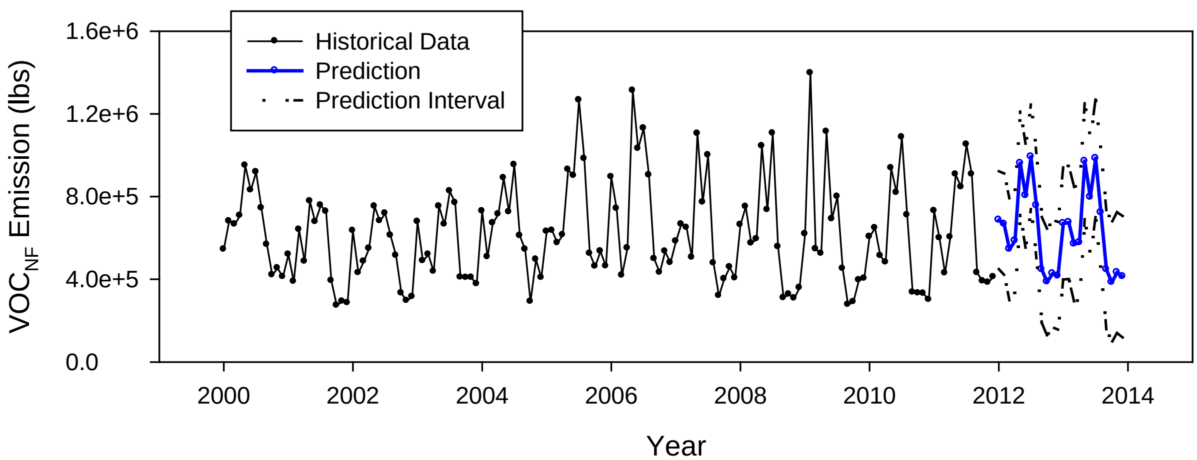

Figure 4. Monthly VOC_{NF} predictions (lbs) of the San Joaquin Valley

The predictions for each month of the 2012 – 2013 ozone seasons are listed in Table 2. The total amounts of the season emissions and their standard errors are then calculated from the monthly predictions. The average daily VOC_{NF} emissions of the ozone season are also calculated accordingly.

To further compare the predictions to the historical data, Table 3 summarizes the average daily emissions of the ozone seasons from 2004 to 2011 and also the base year 1990. Although the SJV VOC_{NF} emissions increased from 9.9 to 12.4 tons/day in 2009 – 2011, the increased amount was not significant compared to the fluctuation of 2004 – 2008. The VOC_{NF} emission predictions for 2012 and 2013 are 11.9 and 11.8 tons/day, falling within the data range of 2004 – 2011 emissions, 11.0 – 14.5 tons/day. As analyzed in the time series modeling, the current evidence is insufficient to support that the SJV VOC_{NF} emissions would continuously increase in 2012 – 2013.

As for the possibility of the SJV VOC emissions reaching the trigger 17.2 tons/day in 2012 – 2013, the contributions from both nonfumigant and fumigant pesticides need to be estimated. However, it is difficult to use the past data to predict fumigant VOC (VOC_{FUM}) emissions in SJV. In 2008, DPR implemented regulations that require use of low emitting fumigation methods in the valley. Subsequently, the VOC_{FUM} emissions during the ozone season were considerably reduced from average 6.6 tons/day in 2004 – 2007 to 3.4 tons/day in 2008 (Table 3). But from 2008 to 2011, the fumigant uses rebounded and the emissions were on the increase. The data of these four years' ozone season can be fitted to a linear regression with an adjusted R^2 0.949:

$$VOC_{FUM} \text{ (tons/day of the ozone season)} = -639.09 + 0.32 * year,\ 2008 \leq year \leq 2011$$

Table 2. VOC_{NF} prediction estimates and standard errors of the San Joaquin Valley during the 2012 and 2013 ozone season

Predicted Monthly Emission (lbs)	2012		2013	
	Est.	**Std. Er.**	**Est.**	**Std. Er.**
May	960302.9	126462.7	970342.5	141754.5
June	804413.6	127263.3	796714.4	142768.1
July	991671.3	128058	984117.9	143774.4
August	756229.3	128847.5	722119.6	144773.7
September	446375.2	129632.3	446419.1	145766.1
October	387152.9	130412.3	384741.2	146751.8
Ozone Season Emission (lbs)	4346145.2[a]	314644.5[b]	4304454.7[a]	353399.8[b]
Ozone Season Emission (tons/day)	**11.9**[c]	**0.9**[c]	**11.8**[c]	**1.0**[c]

a. The estimate of ozone season emission = $\Sigma_{May-Oct}$ (monthly emission estimate)
b. The standard error of ozone season emission = $(\Sigma_{May-Oct}$ (monthly emission standard error^2) $)^{1/2}$
c. The emission (tons/day) = ozone season emission (lbs) / 2000 (lbs/ton) / 183 days

Table 3. Fumigant and nonfumigant pesticide VOC emissions (tons/day) of the San Joaquin Valley during the ozone seasons

Emission (tons/day)	1990	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Fumigant	5.5	6.4	6.9	6.8	6.1	3.4	3.9	4.1	4.4	--	--
Nonfumigant	15.0	11.0	13.8	14.5	11.0	11.2	9.9	11.5	12.4	**11.9**	**11.8**
Source	*Annual report on VOC emissions (Neal et al, 2013)*									*Prediction*	

Table 4. The probability of VOC emissions reaching 17.2 tons/day, trigger level of the San Joaquin Valley, during the 2012 and 2013 ozone season

Year	VOC Trigger *(VOC_T, tons/day)*	Fumigant VOC *(VOC_{FUM}, tons/day)*	Nonfumigant VOC *(VOC_{NF}, tons/day)*			t-statstic	P *(VOC_{NF} > L)*
			L $=VOC_T - VOC_{FUM}$	Pred.	S.E.	*df = 141*	
2012	17.2	3.9 [a]	13.3	11.9	0.9	1.56	6%
		4.4 [b]	12.8			1.00	16%
		4.8 [c]	12.4			0.56	29%
2013	17.2	3.9 [a]	13.3	11.8	1.0	1.50	7%
		4.4 [b]	12.8			1.00	16%
		5.1 [c]	12.1			0.30	38%

a. *Average emission of 2008 – 2011*
b. *Maximum emission in 2008 – 2011*
c. *Forecast from the trend of 2008 – 2011 : VOC_{FUM} = -639.09 + 0.32 * year*

Given the effort of DPR regulations, it is rational to believe that the VOC_{FUM} emissions in the 2012 and 2013 ozone seasons will not reach the emission level before 2008. The emissions from 2008 to 2011 are a comparatively reasonable reference to evaluate the VOC_{FUM} emissions in 2012 – 2013. Since four years data are not enough for time series analysis, three methods are used to approximately estimate the scope of VOC_{FUM} emissions in 2012 – 2013 (Table 4): 1) the average emissions of 2008 – 2011; 2) the emissions of 2011, the most currently available year; and 3) the estimate from the linear regression of 2008 – 2011. The VOC_{NF} emission targets are then calculated by reducing the VOC_{FUM} from the VOC trigger level. Using the estimated predictions and standard errors, the probability of VOC_{NF} higher than the target level can be examined through t-distribution. As shown in Table 4, if the VOC_{FUM} emissions of 2012 – 2013 remain the same level with 2011 or lower, the probability of the total VOC emissions exceeding 17.2 tons/day is not higher than 16%. If the VOC_{FUM} is increasing at the same pace of 2008 – 2011, the probability could reach 38% in 2013.

REFERENCE

Neal, R., Spurlock, F., Segawa, R. 2013. Annual Report on Voloatile Organic Compound Emissions from Pesticides: Emissions for 1990 – 2011. AIR2013-2. Department of Pesticide Regulation, Environmental Monitoring Branch, P.O.Box 4015, Sacramento, California 95812.

Tao, J. 2013. Time Series Analysis and Forecasting of Ventura County Nonfumigant Pesticide Volatile Organic Compound Emissions – 2012 Update. January 28, 2013, Memorandum to Randy Segawa.