Supplementary material

New satellite-based estimates show significant trends in spring phenology and complex sensitivities to temperature and precipitation at northern European latitudes

1. Data analysis

1.1 Panel data analysis of SOS trend

The panel data analysis technique was used in this study to estimate SOS trends. Panel data analysis is a widely-used econometric modeling tool that can generate more accurate and realistic parameter estimation than conventional time series analysis (Hsiao, 2003). Keenan *et al.* (2014) used panel data of MODIS-derived phenology and estimated significant trends that cannot be found from individual pixels using conventional time series analysis. Failure to estimate reliable phenology trends using individual pixels has also been reported in many other studies (e.g. Julien & Sobrino, 2009; Jeong *et al.*, 2011; Zhang *et al.*, 2014). Here, a panel is defined as an E-OBS climate grid cell of 0.22º resolution (nominally 25 km), and each grid cell covers ~2500 MODIS pixels of 500 m resolution. For each panel with either balanced or unbalanced data, a linear regression model for a SOS variable $y_{i,t}$ at pixel i (i = 1, 2, …, N individual pixels in a panel) of year t (t = 1,2, …, T year) in relation to independent variable t can be written as

$$y_{i,t} = \alpha_i + \beta \cdot t + \mu_{i,t}\,, \qquad \text{(S1)}$$

where α_i is a time-invariant parameter for the pixel that accounts for unknown individual effects, such as topography and land cover type. β is a common trend shared by all pixels in the panel. $\mu_{i,t}$ represents residuals of the linear model. The residual plot and Durbin-Watson test (Durbin & Watson, 1950) were used to check possible heteroskedastic and serial correlations in data.

The regression slope was estimated by centering the variables in Eq. (S1):

$$\tilde{y}_{i,t} = \beta \cdot \tilde{t} + \varepsilon_{i,t}\,, \qquad \text{(S2)}$$

where $\tilde{y}_{i,t} = y_{i,t} - \bar{y}_i$, and $\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{i,t}$,

$\tilde{t} = t - \bar{t}$, and $\bar{t} = \frac{1}{T}\sum_{t=1}^{T} t$. $\varepsilon_{i,t}$ is the residuals, assumed to follow a stationary AR(1) process:

$$\varepsilon_{i,t} = \rho\varepsilon_{i,t-1} + e_{i,t}\,, \qquad \text{(S3)}$$

with $|\rho| < 1$ and $e_{i,t} \sim iid(0,\sigma_e^2)$.

Here it can be shown that $Var(\varepsilon_{i,t}) = \rho^2 Var(\varepsilon_{i,t}) + Var(e_{i,t})$, and we get

$$Var(\varepsilon_{i,t}) = \frac{1}{1-\rho^2} Var(e_{i,t}) . \qquad \text{(S4)}$$

The slope was obtained from ordinary least squares (OLS) estimator $\hat{\beta}$ from Eq. (S2), with a corrected variance $\hat{V}ar(\hat{\beta})$ to account for heteroscedasticity and autocorrelation of a supposed AR(1) process following Wooldridge (2010, p. 61) and Baltagi *et al.* (2011):

$$\hat{\beta} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{Y} . \qquad \text{(S5)}$$

$$\hat{V}ar(\hat{\beta}) = \frac{NT}{NT-1}\frac{1}{(1-\hat{\rho}^2)}(\tilde{X}'\tilde{X})^{-1}\left[\sum_{i=1}^{N}\tilde{x}_i' e_i e_i' \tilde{x}_i\right](\tilde{X}'\tilde{X})^{-1} . \qquad \text{(S6)}$$

Here $NT/(NT-1)$ corrects for the degree of freedom; $1/(1-\hat{\rho}^2)$ corrects for serial correlation of an AR(1) process, $\hat{\rho}$ was estimated from Eq. (S3) (Baltagi et al. 2011); and the rest sandwich form $(\tilde{X}'\tilde{X})^{-1}\left[\sum_{i=1}^{N}\tilde{x}_i' e_i e_i' \tilde{x}_i\right](\tilde{X}'\tilde{X})^{-1}$ corrects for heteroscedasticity. Whether the AR(1) process is stationary ($|\hat{\rho}|<1$) or not was checked.

The modified variance was used to test if the estimated slope $\hat{\beta}$ significantly differs from zero using t-statistics,

$$t = \frac{\hat{\beta}}{\sqrt{\hat{V}ar(\hat{\beta})}} . \qquad \text{(S7)}$$

The slope variance was also used in the consequent meta-analysis of slopes for the whole study area.

1.2 Panel data analysis of SOS climate sensitivity

We used a linear model between first-order differences in SOS and climate variables to estimate climate sensitivities of SOS (Zhou *et al.*, 2001):

$$\Delta y_{i,t} = \alpha_i + \beta_1 \cdot \Delta T_{i,t} + \beta_2 \cdot \Delta P_{i,t} + \varepsilon_{i,t} , \qquad \text{(S8)}$$

where $\Delta y_{i,t}$, $\Delta T_{i,t}$, and $\Delta P_{i,t}$ are the differences in SOS, temperature, and precipitation at pixel *i* between year *t* and year *t-1* respectively. The regression slope β_1 gives the temperature sensitivity and β_2 the precipitation sensitivity. The climate sensitivity β_1 and β_2 were estimated using the panel analysis in the similar way as the aforementioned SOS trend estimation. The heteroscedasticity and autocorrelation corrected variance were using an OLS estimator. These values were then used in significance test and meta-analysis of slopes.

1.3 Meta-analysis of trends and sensitivities

The weighted least squares method (Becker & Wu, 2007) was used to summarize the overall slopes by assuming that the regression slopes of *M* panels are independently and normally distributed:

$$\hat{\beta}_{meta} = \frac{\sum_{i=1}^{M} w_i \hat{\beta}_i}{\sum_{i=1}^{M} w_i}, \tag{S9}$$

where the weight is given by the reciprocal of slope variance: $w_i = 1/\hat{V}ar(\hat{\beta}_i)$, and the variance of the meta-slope is

$$Var(\hat{\beta}_{meta}) = M / \sum_{i=1}^{M} w_i \, . \tag{S10}$$

Note: the variance of meta-slope is incorrect in Eq (3) of Becker and Wu (2007).

1.4 Partial Correlation

Using the panel data, the partial correlation coefficients between the phenology metrics and climate variables were estimated from the first-order difference of variables. The analysis was done using Matlab (Ver 15b, MathWorks, Inc.) The Matlab scripts are shown in section 4 supplementary scripts.

References

Baltagi, B.H., Kao, C. & Na, S. (2011) Test of Hypotheses in Panel Data Models When the Regressor and Disturbances are Possibly Nonstationary. *Advances in Statistical Analysis*, **95**, 329-350.

Becker, B.J. & Wu, M.-J. (2007) The Synthesis of Regression Slopes in Meta-Analysis. *Statistical Science*, **22**, 414-429.

Durbin, J. & Watson, G.S. (1950) Testing for Serial Correlation in Least Squares Regression: I. *Biometrika*, **37**, 409-428.

Hsiao, C. (2003) *Analysis of panel data*, 2nd edn. Cambridge University Press.

Jeong, S.-J., Ho, C.-H., Gim, H.-J. & Brown, M.E. (2011) Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982–2008. *Global Change Biology*, **17**, 2385-2399.

Julien, Y. & Sobrino, J.A. (2009) Global land surface phenology trends from GIMMS database. *International Journal of Remote Sensing*, **30**, 3495-3513.

Keenan, T.F., Gray, J., Friedl, M.A., Toomey, M., Bohrer, G., Hollinger, D.Y., Munger, J.W., O/'Keefe, J., Schmid, H.P., Wing, I.S., Yang, B. & Richardson, A.D. (2014) Net carbon uptake has increased through warming-induced changes in temperate forest phenology. *Nature Clim. Change*, **4**, 598-604.

Wooldridge, J.M. (2010) *Econometric Analysis of Cross Section and Panel Data*, 2nd edn. The MIT Press, Massachusetts Institute of Technology.

Zhang, X., Tan, B. & Yu, Y. (2014) Interannual variations and trends in global land surface phenology derived from enhanced vegetation index during 1982–2010. *International Journal of Biometeorology*, **58**, 547-564.

Zhou, L., Tucker, C.J., Kaufmann, R.K., Slayback, D., Shabanov, N.V. & Myneni, R.B. (2001) Variations in northern vegetation activity inferred from satellite data of vegetation index during 1981 to 1999. *J. Geophys. Res.*, **106**, 20069-20083.

2. Supporting figures

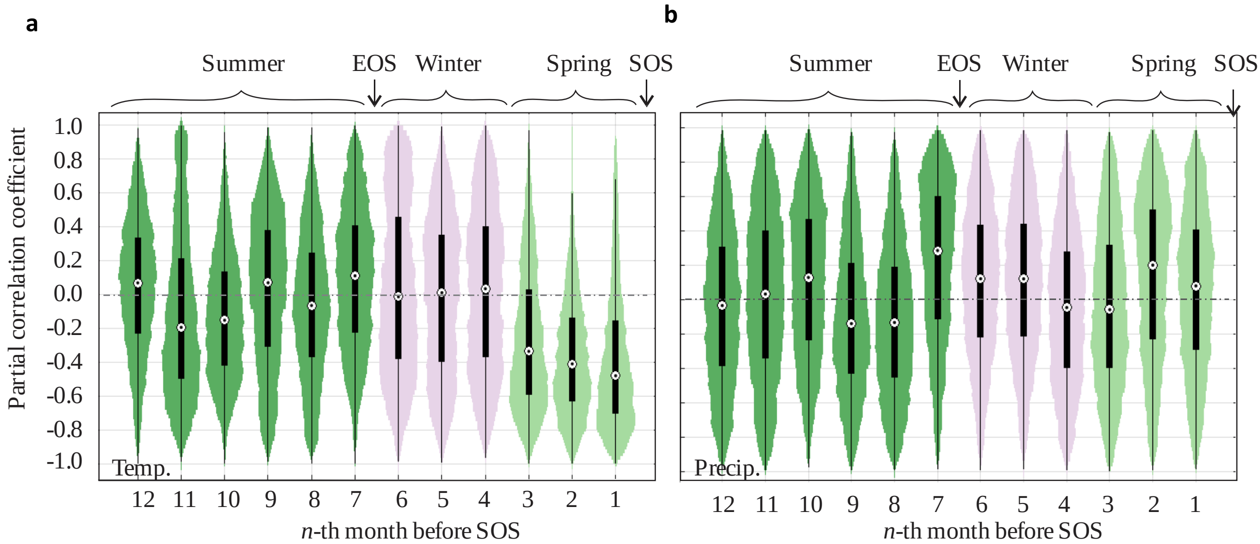

Fig. S1 Violin plots of partial correlation of the SOS to a temperature and b precipitation of the nth month (n = 1, 2, … ,12) before the mean SOS during 2000–2016. The shaded areas of the violin plot denote distribution, the white dots denote the median values, and the bottom and top edges of the black bars denote the 25th and 75th percentiles respectively. The three-month pre-season of SOS (termed Spring) had the highest negative correlation with the SOS and the remaining 9 months of a full yearly cycle were divided into before and after the EOS (termed Summer and Winter respectively). Note that the EOS position is only for illustration purpose and may not necessarily be at the 6~7th month before the SOS.

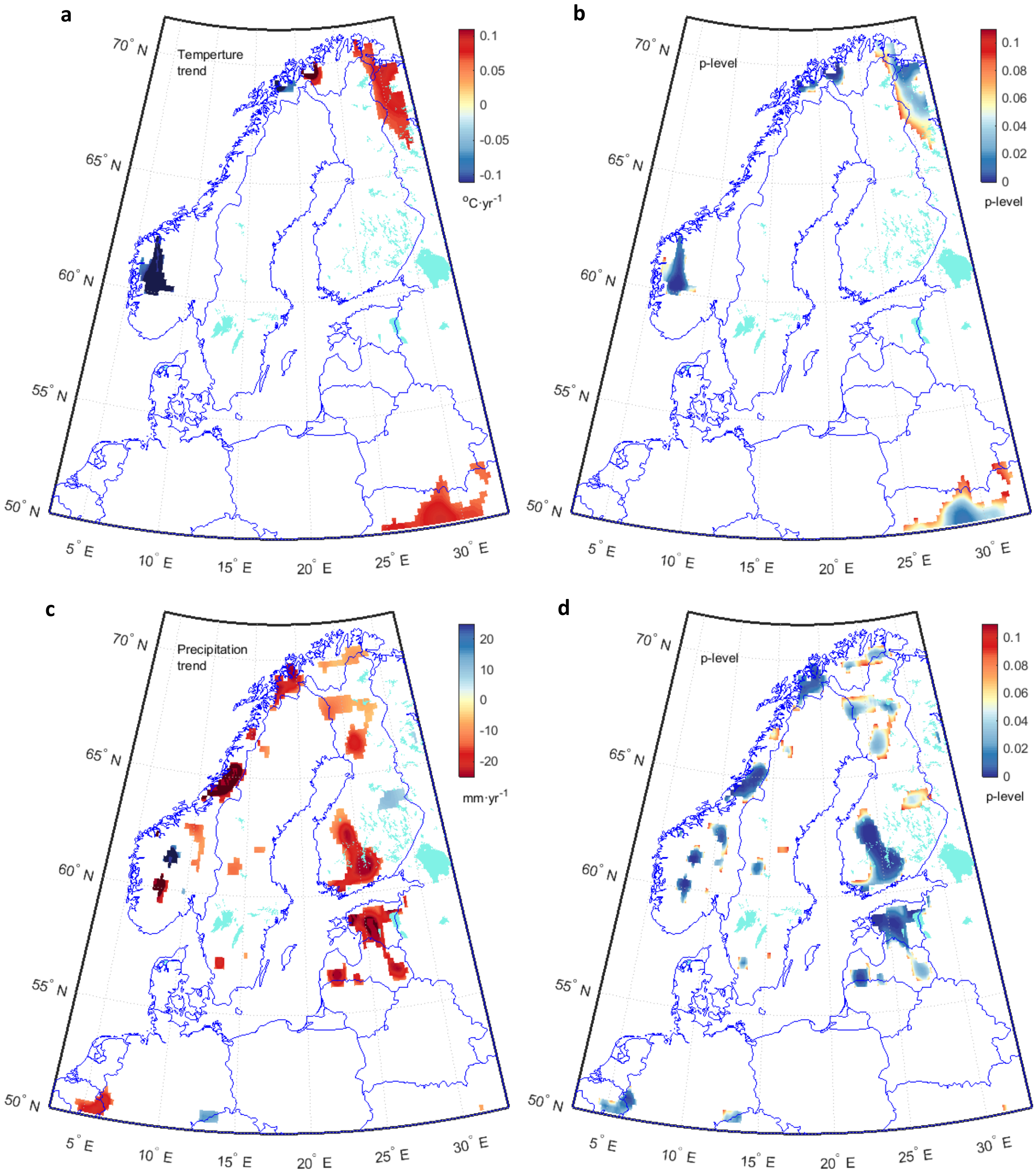

Fig. S2 Mean air temperature **a** trend (°C·yr^{-1}) and **b** significance level, and total precipitation **c** trend ((mm·yr^{-1}) and **d** significance level during the year preceding the mean SOS over 2000-2016. Only the trends significant at $p \leq 0.1$ are shown.

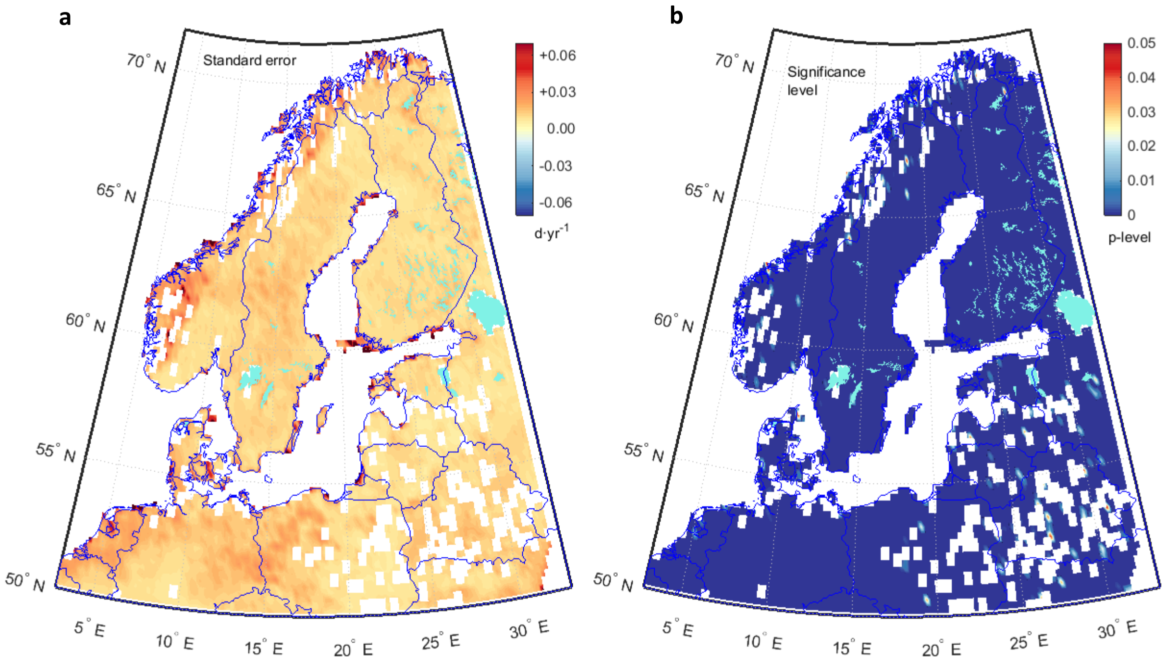

Fig. S3 **a** Standard error (d·yr^{-1}) and **b** significance level of SOS trend during 2000-2016. The standard error was estimated from the variance of the regression slope based on panel analysis, being corrected for heteroscedasticity and autocorrelation. The panels with insignificant trends ($p > 0.05$) are not shown.

3. Supporting Tables

Table S1 Temperature and precipitation trends (mean ± SE, $p > 0.1$) during periods before the SOS from 2000 to 2016.

Land cover types	Temperature trend (°C·yr^{-1}) during				Precipitation trend (mm·yr^{-1}) during			
	Spring	Winter	Summer	Year	Spring	Winter	Summer	Year
ALL	0.04 ± 0.07	0.03 ± 0.08	0.01 ± 0.04	0.03 ± 0.04	-0.94 ± 1.56	-0.82 ± 2.57	-1.15 ± 3.14	-2.62 ± 5.04
ENF	0.06 ± 0.07	0.01 ± 0.09	0.01 ± 0.04	0.02 ± 0.04	-0.73 ± 1.43	-1.37 ± 2.84	0.02 ± 3.38	-2.46 ± 5.44
DBF	0.05 ± 0.07	0.03 ± 0.08	0.01 ± 0.04	0.03 ± 0.04	-0.79 ± 1.52	-1.06 ± 2.68	-2.13 ± 3.00	-3.47 ± 5.07
CRO	0.00 ± 0.09	0.04 ± 0.07	0.02 ± 0.03	0.03 ± 0.04	-1.87 ± 1.69	-0.32 ± 2.23	0.27 ± 3.29	-1.71 ± 4.76
WET	0.06 ± 0.06	0.03 ± 0.08	0.00 ± 0.04	0.03 ± 0.04	-0.70 ± 1.35	-1.70 ± 2.90	-3.53 ± 2.81	-6.09 ± 4.88

Note: ALL − all land covers, ENF – evergreen needle leaf forest, DBF – deciduous broadleaf forest, shrubland and mixed forest, CRO –cropland and grassland, WET – wetland. Spring – the three months preceding SOS, Winter – from EOS to the start of Spring, Summer – from SOS to EOS, Year – the full year preceding SOS.

Table S2 Proportions of significant ($p \leq 0.05$) grids of all data in panel analysis of SOS trend and climate sensitivity

Land cover types	Trend	Mean temperature during			Total precipitation during		
		Spring	Winter	Summer	Spring	Winter	Summer
ALL	95.6%	99.1%	95.4%	94.2%	94.0%	93.0%	94.0%
ENF	89.6%	96.8%	88.6%	88.8%	84.4%	83.6%	85.5%
DBF	93.6%	98.9%	91.2%	91.7%	91.1%	91.1%	91.2%
CRO	91.8%	98.3%	90.2%	89.9%	85.8%	87.2%	87.8%
WET	85.2%	94.3%	84.8%	82.4%	80.1%	75.8%	80.1%

4. Matlab scripts

Panel analysis on climate sensitivity.

```
function [slope, p_mod, var_mod,par_rho] = panel(Y,X1,X2)
%Input: Y(n,m): the Dela_DOY of SOS of m years and n pixels in a panel of climate grid (1st-order difference in SOS)
%        X1(m),X2(m):two climate variables, e.g. Dela_TMP and Dela_PPT of m years of the climate grid (1st-order difference in TMP and PPT)
%Output: slope: regression slope; p_mod: p-level, correct for heteroscedasticity and autocorrelation; var_mod: variance of slope, correct for heteroscedasticity and autocorrelation; par_rho: partial correlation, by controlling other factors, including time.

X1=repmat(X1,n,1);
X1_mean = nanmean(X1,2);
X1 = bsxfun(@minus,X1, X1_mean); % deduction of group mean
X2=repmat(X2,n,1);
X2_mean = nanmean(X2,2);
X2 = bsxfun(@minus,X2, X2_mean); % deduction of group mean,
Y_mean = nanmean(Y,2);
Y = bsxfun(@minus,Y, Y_mean); % deduction of group mean

Ind=find(~all(isnan(Y),2) & ~all(isnan(X1),2)& ~all(isnan(X2),2)); %unbalanced panel data, some years have data, some years have not
if  length(Ind) < 20 %no process if less 20 pixels (each has 17 years data)
   slope = nan(2,1);
   p_mod = nan(2,1);
   var_mod = nan(2,1);
   par_rho = nan(2,1);
% Baltagi, B.H., Kao, C., & Liu, L. (2014). Test of Hypotheses in a Time Trend Panel Data Model with Serially Correlated Error Component Disturbances. Advances in Econometrics, 33, 347-394
else
   Ytmp=Y(Ind,:);
   X1tmp=X1(Ind,:); % for unbalanced panel
   X2tmp=X2(Ind,:);
   ind=find(~any(isnan([Ytmp(:),X1tmp(:),X2tmp(:)]),2));
   par_rho = partialcorri(Ytmp(ind),[X1tmp(ind),X2tmp(ind)]);
   beta =
   ([X1tmp(ind),X2tmp(ind)]'*[X1tmp(ind),X2tmp(ind)])\([X1tmp(ind),X2tmp(ind)]'*Ytmp
   (ind));
   r = Ytmp(ind) - [X1tmp(ind),X2tmp(ind)]* beta; % within group residual
% first OLS to estimate AR(1) process of residuals
   slope = beta; %OLS univariate fixed effect coefficient, only get slope
   epslonit = nan(size(Ytmp));
   epslonit(ind) = r;
   rho = nansum(nansum(epslonit(:,2:m).*epslonit(:,1:m-
   1)))/nansum(nansum(epslonit(:,1:m-1).^2.* ~isnan(epslonit(:,2:m))));
   %Baltagi, B.H.; Kao, C.; Na, S. Test of hypotheses in panel data models when the regressor and disturbances are possibly nonstationary. Advances in Statistical Analysis 2011, 95, 329-350.
   if abs(rho) >= 1
      fprintf('Non-stationary process!\n');
      beep; end
   df = length(find(~isnan(Y(:))));
   tmp = [nansum(X1tmp.*epslonit,2),nansum(X2tmp.*epslonit,2)];
   meat=tmp'*tmp;
   bread = ([X1tmp(ind),X2tmp(ind)]'*[X1tmp(ind),X2tmp(ind)])\eye(2);
   covb = df/(df-1) * bread * meat * bread /(1-rho^2); % sandwich
   var_mod = diag(covb);
   tval= slope ./ sqrt(var_mod); %t test p value
   p_mod = 2 * tcdf(-abs(tval), df);

end
```