Author's response to reviews

Title: Using a Generalized Additive Model with Autoregressive Terms to Study the Effects of Daily Temperature on Mortality

Authors:

Lei Yang (yanglei_cq@126.com)
Guoyou Qin (gyqin@fudan.edu.cn)
Naiqing Zhao (nqzhao@fudan.edu.cn)
Chunfang Wang (cfwang@scdc.sh.cn)
Guixiang Song (gxsong@scdc.sh.cn)

Version: 8 Date: 20 September 2012

Author's response to reviews: see over
Reply to reviewers and editors.

Based on the manuscript we recently update, there are some new changes after more discussion and consideration. We are very sorry that we couldn’t incorporate all these modifications at last time. To reduce the time you need to revise the new version, we list all new changes, and you can only review these changes based on the previous version uploaded in Sep. 13rd. Hope not to bother you too much.

1. All “AutoRegressive” and “AutoCorrelation” were changed to “Autoregressive” and “Autocorrelation”.
2. The spaces between the equations and their corresponding numbers have increased.
3. The Appendix is simplified and clarified. The method is actually a modified Newton’s method, so the title is also changed. (P18, Appendix)
4. The section “natural cubic spline” is simplified and clarified. (P7, Line 11)
5. The description of GAMM in Chapter discussion is modified to be more clear and detailed. (P16, Line 3)

The remaining parts of article are just the same as the former one. There are some changes in the cover letter below, based on the former cover letter. These changes are meant to express our gratitude to the reviewers, for their comments and suggestions which we believe will greatly improve our manuscript.

******

Thank you very much for your suggestions.

In response to your suggestion, now the manuscript is proofread by a native speaker. He has modified the article much as well as gave us some important suggestions. As a result, there are some important changes listed as follows:

1. The manuscript’s title and the model’s name were altered. The model’s name is now GAMAR under his suggestion.
2. The deductions in the section “Newton’s method” was moved into Appendix
3. Several tables of less significance were moved to Additional files
4. The language of the manuscript was improved. And many parts of the manuscript were simplified

There are also other minor modifications.

Changes made according to the suggestions of reviewers are given separately below. These suggestions help us a lot to improve manuscript.

Reply to Xiangming Fang

Thank you very much for your review suggestions.

Following content in bold are my corresponding modifications and replies.

Major Compulsory Revisions:

1) The authors mentioned in the third paragraph on page 5 that GLMM/GAMM are unsuitable for time series without clustering. This is not true. Although GLMM and GAMM are often used for longitudinal, clustered, and functional data, they can definitely be applied to time series as well. Just treat the time series as a single cluster/group.

I still think it would be more comparable if the authors could compare
GAMwithAR with GAMM in their simulation study 1. They could simulate the data from the GAMwithAR model and then fit both GAMM and GAMwithAR with the same AR order. Since in practice we usually do not know how the AR correlation was formed (through regression errors as in GAMwithAR or through random errors as in mixed models), so for practical purpose, it is of interest to see how the two model fits differ from each other.

**Your opinion is right. And this idea is very interesting and enlightening.** We have learnt a lot from it and thus look at this problem from a total new perspective. We used to think such autocorrelation structure is modeled by between-group errors, in fact, it is modeled by within-group errors. And it can be modeled! It is a degenerative case of GAMM and can be modeled by GAMM. However, it is difficult to simulate data from GAMM with correlation structure (except for normal distribution). And it is more practical to simulate data from GAMAR(3). So the comparison between GAMM and GAMAR is based on data from GAMAR(3) in simulation study 1.

We've found GAMM gives quite good estimates as shown in Table 1. However, it is rather computing extensive, the time for GAMM and GAMAR(3) for a same data set from simulation 1 with sample size 1641 are 1 hour 7 minutes and 0.02 seconds on a same PC. So we only fit 50 samples by GAMM.

The results are reported in Table 1 and other related parts of the manuscript are also modified according to this change.

2) Below Equation (9) on page 17, it says the degrees of freedom of all natural spline functions are determined simultaneously by minimizing AIC. I am not sure if this can be done easily. Are the reported values the minimizer of AIC for all the possible combinations of df values? Or, are they just the best values among the combinations that the authors had tried?

**Just as you pointed out, “simultaneously” is not an accurate description of parameter determination procedure.** In fact, we first find the parameters which minimize AIC locally in a sequence. Then a procedure using this set of parameters as starting value to find a set of parameters which minimize AIC locally. This part is modified and the corresponding program is provided in Additional files. Your query has helped us to improve the explanation of this optimization process.

Minor Essential Revisions:
Line 7 of Page 7: Please change “doesn’t” to “does not”.
It is changed to be so.
Line 8 of Page 9: “L” in the definition of \( \theta \) should be “...”.

**The definition of \( \theta \) is clarified**
Last line of Page 9: Equation reference error.

**Equation reference is clarified.**
Line 8 of Page 10: Equation reference error.

**Equation reference is clarified**.
5) Line 8 of Page 11: “[i=1,2,L,mp” should be “[i=1,2,...,mp”.

**Equation reference is clarified**.
6) Line -6 (6th from the bottom) of Page 13: Please change “What’s more” to “Moreover”.
This modification has been made
7) Line -5 of Page 13: “coefficient” should be “coefficients”.
This modification has been made
8) Line -4 of Page 13: “calculated, a higher...” should be “calculated. A higher...”.
I don’t quite understand this change, please see in our manuscript whether I have also modified this.
9) Line -4 of Page 13: “the fitted value is” should be “the fitted values are”.
This modification has been made
10) Line -3 of Page 13: “more dependent on the true value” should be “closer to the true values”.
This confusing sentence is deleted.
11) Line 4 of Page 15: “Table 1 show...” should be “Table 1 shows...”.
This modification has been made
12) Line 5 of Page 15: “GAM are...” should be “GAM is...”.
This sentence means the “biases of GAM are”, not the “GAM is”. Since the original sentence is ambiguous, we change it into: “Table 1 shows that the biases of the mean parameter estimates from GAM are almost the same as GAMAR.”
13) Line 11 of Page 16: “which demonstrate...” should be “which demonstrates...”.
This sentence has been changed into a clearer shape: “In contrast, the ACF and PACF of GAMAR(3) on the same data are both very close to 0.”
14) Line 13 of Page 16: “the Pearson correlation coefficient” should be “the Pearson correlation coefficients”.
This modification has been made

Reply to Michael Hayn
Thank you very much for your review suggestions. You are so kind and careful that always point out our mistakes and non-standard parts in the details. We believe your valuable comments have helped a lot to improve our manuscript.
Following content in bold are my corresponding modifications and replies.

Minor Essential Revisions:
p. 2, l. 5: 
"[...], GAM fundamentally assumes errors are mutually independent":
correct to 
"[...], the GAM fundamentally assumes that errors are mutually independent” or 
"[...], the GAM fundamentally assumes mutually independent errors"
This sentence is changed to be” GAM assumes that errors are mutually independent”
p. 3, l. 3:
"[...] studies have done to [...]":
correct to 
"[...] studies have been done to [...]"
This sentence is changed to be “Recently, many studies have been done to analyse the way and the extent temperature influences health outcomes.”
p. 7, l. 8
"[...] or negative values":
correct to
"[...] or negative values."
(Please finish the sentence. Many other not finished sentences appear, especially when sentences finish with mathematical expressions
Examples: p. 8, l. 21 (finishing with "PL=[...]" ); p. 9, l. 1 (finishing with
"PL=[...]" ), 15 (finishing with "x_{m+1}=[...]" ), 18 (finishing with
"\theta_{m+1}=[...]" ), and many further examples in the section "Newton's
method", p. 12, l. 10 (finishing with "parameters are: [formula]" ), and
further examples.)
All such mistakes have been corrected.
p. 8, l. 4:
"[...] is a cubic.":
correct to
"[...] is a cubic function."
This modification has been made
p. 8, l. 19:
"[...] a sequence of conditional likelihoods f(yt \mid Xt, [...], t=1,...,n, where
[...])":
correct to
"[...] a sequence of conditional likelihoods f(yt \mid Xt, [...], t=1,...,n), where
[...]
(Close the parentheses.)
The parentheses is close after \theta, not n.
p. 9, l. 19; p. 10, l. 8:
Please correct for the newly appeared reference errors.
These are corrected.
p. 13, l. 13:
"Same as simulation study 1, sample statistics are calculated over time points
[...]":
correct to
"As in simulation study 1, the sample statistics are calculated over the time points
[...]"
This is modified.
p. 13, l. 15:
"The first sample are analysed just the same way as simulation study 1.":
correct to
"The first sample is analysed just the same way as in simulation study 1."
This sentence is corrected to be so.
p. 13, l. 25:
"[...] similar to relative error in Simulation 1.":
correct to
"[...] similar to the relative error in simulation 1."
This sentence is deleted.

p. 14, l. 4: Please replace "humility" by "humidity".

This change has been made.

p. 15, l. 4:
"Insofar as general performance of [...] Table 1 show that [...]":
correct to
"Concerning the general performance of [...] Table 1 shows that [...]"

This change has been made.

p. 16, l. 7:
"Just as the former study, we can see ACF tails off and PACF [...]":
correct to
"Just as in the former study, we can see that the ACF tails off and the PACF [...]"

This change has been made.

p. 16, l. 11:
"[...], which demonstrate its Pearson residual is nearly white noise. From Figure 5, we can see the estimated curve [...] is closer [...]":
correct to
"[...], which demonstrates that its Pearson residual is nearly white noise. From Figure 5, we can see that the estimated curve [...] is closer [...]"

This sentence has been changed into a clearer shape: “In contrast, the ACF and PACF of GAMAR(3) on the same data are both very close to 0.”

p. 19, l. 6:
"[...], which further justify the choice or AR order.":
correct to
"[...], which further justifies the choice or the AR order."

This sentence has been deleted.

p. 19, l. 9:
"[...] to finally ensure our requirement is [...]":
correct to
"[...] to finally ensure that our requirement is [...]"

This sentence has been deleted.

p. 19, l. 18:
"[...] while the real model is GAMwithAR(3). GAMwithAR with AR(1), [...]":
Correct to
"[...] while the real model is GAMwithAR(3), GAMwithAR with AR(1), [...]"

I don’t quite understand your meaning.

p. 21, l. 11:
"[...] many days. 3) cumulative [...]":
correct to
"[...] many days; 3) cumulative [...]"

This change has been made.

Furthermore, I recommend to avoid starting sentences with symbols.

Recommended changes:

p. 6, l. 23 - p. 7, l. 1
"[formula]\newline [formula] are autoregressive terms [...]": correct to
"[formula], [formula] are autoregressive terms [...]"
Similarly at p. 7, l. 13.
p. 7, l. 6,19; p. 8, l. 10:
"[formula]\newline [formula] is a positive threshold parameter.": correct to
"[formula], [formula] is a positive threshold parameter."
p. 9, l. 7:
"[formula]\newline [\theta=(\beta_1, \ldots, \beta_m, c_1, \ldots, c_p)^T is the model parameter vector.": correct to
"[formula], [\theta=(\beta_1, \ldots, \beta_m, c_1, \ldots, c_p)^T is the model parameter vector."
Please also replace "L" by dots. Also at p. 11, l. 8 ("i=1,2,L,mp")
p. 12, l. 6:
"[formula]\newline x_t is daily average temperature [...]": correct to
"[formula], [x_t is the daily average temperature [...]"

All the above changes have been made to ensure that sentence won’t begin with symbols.
p. 12, l. 7:
"s_{i5} […] form the": correct to
"The terms s_{i5} […] form the"

This change has been made.

p. 8, l. 2:
"an interval[a,b]. [a,b] is [...]": correct to
"an interval [a,b]. The interval [a,b] is [...]"
This has been changed into “an interval [a,b], where [a,b] is…”
p. 8, l. 7
"ns is often […]": correct to
"Natural splines are often […]” or "The ns is often […]"
Similarly at p. 8, l. 11
"ns is usually expressed by linear […]” correct to
"Natural splines are usually expressed by a linear […]”

When “natural cubic splines” first occur in the manuscript, it abbreviation ns is also given. So I think it can avoid misunderstandings.
p. 10, l. 11:
"[formula], is information matrix:" correct to
"The information matrix $[\text{formula}]$ is given by:"

This sentence is changed into “The information matrix is”

p. 17, l. 10:

"$[\text{formula}]$ stands for week effect. $[\text{formula}]$ represents corresponding day [...]":
correct to

"The term $[\text{formula}]$ stands for the week effect, and $[\text{formula}]$ represents the corresponding day [...]"

This change has been made.