



HAL
open science

Uncertainty component estimates in transient climate projections.

Benoit Hingray, Juliette Blanchet, Guillaume Evin, Jean-Philippe Vidal

► **To cite this version:**

Benoit Hingray, Juliette Blanchet, Guillaume Evin, Jean-Philippe Vidal. Uncertainty component estimates in transient climate projections.: Precision of estimators in a single time or time series approach. *Climate Dynamics*, 2019, 53 (5-6), pp.2501-2516. 10.1007/s00382-019-04635-1 . hal-02327452

HAL Id: hal-02327452

<https://hal.science/hal-02327452>

Submitted on 26 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Uncertainty component estimates in transient**
2 **climate projections**

3 **Precision of estimators in a single time or time series**
4 **approach**

5 **Benoit Hingray · Juliette Blanchet ·**
6 **Guillaume Evin · Jean-Philippe Vidal**

7
8 Received: date / Accepted: date

9 **Abstract** Quantifying model uncertainty and internal variability components
10 in climate projections has been paid a great attention in the recent years. For
11 multiple synthetic ensembles of climate projections, we compare the precision
12 of uncertainty component estimates obtained respectively with the two Analy-
13 sis of Variance (ANOVA) approaches mostly used in recent works: the popular
14 Single Time approach (STANOVA), based on the data available for the con-
15 sidered projection lead time and a time series based approach (QEANOVA),
16 which assumes quasi-ergodicity of climate outputs over the available simula-
17 tion period.

18 We show that the precision of all uncertainty estimates is higher when more
19 members are used, when internal variability is smaller and/or the response-to-
20 uncertainty ratio is higher. QEANOVA estimates are much more precise than
21 STANOVA ones: QEANOVA simulated confidence intervals are roughly 3 to
22 5 times smaller than STANOVA ones. Except for STANOVA when less than
23 3 members is available, the precision is rather high for total uncertainty and
24 moderate for internal variability estimates. For model uncertainty or response-
25 to-uncertainty ratio estimates, the precision is low for QEANOVA to very low
26 for STANOVA. In the most unfavorable configurations (small number of mem-
27 bers, large internal variability), large over- or underestimation of uncertainty
28 components is thus very likely. In a number of cases, the uncertainty analysis
29 should thus be preferentially carried out with a time series approach or with a

B. Hingray
Univ. Grenoble Alpes, CNRS, IGE UMR 5001, Grenoble, F-38000, France
E-mail: benoit.hingray@ujf-grenoble.fr

J. Blanchet
Univ. Grenoble Alpes, CNRS, IGE UMR 5001, Grenoble, F-38000, France

G. Evin
Univ. Grenoble Alpes, Irstea, UR ETNA, Grenoble, F-38000, France

J.-P. Vidal
Irstea, UR RiverLy, centre de Lyon-Villeurbanne, F-69625 Villeurbanne, France

30 local-time series approach, applied to all predictions available in the temporal
31 neighborhood of the target prediction lead time.

32 **Keywords** Uncertainty Sources · Climate Projections · ANOVA · Internal
33 Variability · Model Uncertainty · Scenario Uncertainty · Precision of Estimates

34 1 Introduction

35 A critical issue in climate change studies is the estimation of uncertainties in
36 projections along with the contribution of the different uncertainty sources,
37 including scenario uncertainty, the different components of model uncertainty,
38 and internal variability (e.g. Hawkins and Sutton, 2009).

39 Scenario uncertainty is related to the poorly known future of greenhouse
40 gas emissions. Model uncertainty corresponds to the dispersion between the
41 different climate responses obtained with different models for the same forcing
42 configuration. Model uncertainty concerns Global Climate Models (GCMs) but
43 also all subsequent models of the climate change impact modeling chain, such
44 as Regional Downscaling Models (regional climate models and/or statistical
45 downscaling methods) and impact models like Hydrological Models (HMs).
46 Internal variability first originates from the chaotic variability of the climate
47 at the global scale (e.g. Räisänen, 2001; Deser et al, 2012). At the local scale, it
48 also results from the fact that very similar large scale atmospheric circulation
49 configurations can lead to very different meteorological observations (Braun
50 et al, 2012; Lafaysse et al, 2014).

51 Estimating and partitioning uncertainty in future climate projections is
52 first intended to help evaluating the significance of estimated changes for adap-
53 tation purposes. Besides, this is intended to highlight the most important un-
54 certainty sources. This thus allows estimating the fraction of total uncertainty
55 that could be narrowed via scenario refinement and model improvement, and
56 its irreducible fraction pertaining to natural variability (e.g. Hawkins and Sut-
57 ton, 2011; Lafaysse et al, 2014).

58 Over the recent years, uncertainty has been mostly explored and parti-
59 tioned based on Multiscenarios Multimodel Multimember Ensembles (MMEs)
60 of transient climate projections. Various methods have been proposed for this,
61 most of them based on an Analysis of Variance (ANOVA) of projections avail-
62 able for the specific projection lead time considered (Hingray et al, 2007; Yip
63 et al, 2011; Bosshard et al, 2013; Giuntoli et al, 2015; van Pelt et al, 2015;
64 Paeth et al, 2017). In this single time approach – called STANOVA in the
65 following –, and provided multiple members are available for each modeling
66 chain, the model uncertainty components are estimated from the dispersion
67 between the climate responses of the different modeling chains, obtained for
68 each chain from the multimember mean of the projections. Similarly, the in-
69 ternal variability component for each modeling chain is estimated from the
70 inter-member variance of the projections.

71 The different uncertainty components can also be estimated from a time
72 series approach, benefiting from the long time series now available for most

73 climate experiments (e.g. Johns et al, 2011; Jacob et al, 2014). A quasi-ergodic
74 assumption for climate simulations in transient climate is often used in this
75 context (Hingray and Saïd, 2014). It considers that the temporal variations
76 of the climate response of a particular simulation chain are necessarily grad-
77 ual and smooth, the higher frequency variations of the time series being due
78 to internal variability alone. It considers next that the internal variability
79 can be assumed to remain roughly constant over the considered period or to
80 vary as a gradual and smooth function of the climate response of the chain.
81 These assumptions were used for instance by Räisänen (2001); Hawkins and
82 Sutton (2009, 2011); Charlton-Perez et al (2010); Lafaysse et al (2014); Brace-
83 girdle et al (2014); Reintges et al (2017). In this time series approach called
84 QEANOVA in the following, the climate change response of each simulation
85 chain and its possible evolution with time is obtained from the long term trend
86 estimated from the time series of the chain. The variance over time of the de-
87 viation from the climate response defines the internal variability of each chain.
88 The climate responses of the different chains can be used to estimate the com-
89 ponents of model uncertainty for any projection lead time, thanks again to a
90 usual ANOVA.

91 In both approaches, a potentially critical step is the possibility to identify
92 a precise estimate of the climate response of each modeling chain. A number of
93 previous works advocated the use of multiple runs to improve this identification
94 (e.g. Kendon et al, 2008; Deser et al, 2012; Kew et al, 2011). Unfortunately
95 a limited number of members have until now been usually available for most
96 modeling chains, logically deriving from the small number of GCM runs (e.g.
97 Johns et al, 2011; Jacob et al, 2014; Paeth et al, 2017; Reintges et al, 2017). In
98 a number of analyses, the discrimination between internal climate variability
99 and the climate response of each modeling chain is thus expected to be quite
100 inaccurate owing to the limited ensemble size especially when the internal
101 variability is non-negligible compared to the chain's climate responses (e.g.
102 Deser et al, 2012; van Pelt et al, 2015).

103 These inaccuracies translate to inaccuracies in estimates of uncertainty
104 components and of derived characteristics (e.g. response to uncertainty ratio,
105 significance of projected changes, time of emergence of climate change, etc.).
106 In specific analysis configurations, the precision of uncertainty component es-
107 timates is potentially low to very low, but to our knowledge this issue has
108 never been considered.

109 In the following, we characterize the precision of uncertainty components
110 estimates when obtained either with the single time or the time series ap-
111 proaches introduced previously. The precision is obviously expected to depend
112 on the multimodel multimember ensemble of projections (MME) available for
113 the analysis and on the climate variable considered for the analysis. In the
114 following, both ANOVA approaches are applied to multiple synthetic MMEs
115 obtained via Monte Carlo simulations. We consider a simplified configuration
116 where MME account for two uncertainty sources: model uncertainty and in-
117 ternal variability. The analysis framework we follow here is rather crude but it
118 allows comparing both ANOVA approaches for a large number of MME config-

119 urations with different characteristics. We especially discuss how the precision
 120 depends on the number of members used in the analysis, and we characterize
 121 the gain in precision obtained with the time series based approach, owing to
 122 the larger size of the data set accounted for. We also discuss the results for
 123 different levels of internal climate variability contribution to total uncertainty
 124 variance. We first focus on the precision of estimates obtained for model un-
 125 certainty and internal variability. We next consider the precision of estimates
 126 obtained on the one hand for the mean climate response from the model en-
 127 semble and on the other hand for the response to uncertainty ratio frequently
 128 used to estimate the significance of estimated changes.

129 All recent analyses of uncertainty sources in climate projections make use of
 130 the moment estimators of variance components. The present study also makes
 131 use of such estimators in order to put the precision of uncertainty estimates
 132 derived from these analyses into perspective.

133 The statistical bases for the estimation of uncertainty components in a
 134 given MME are summarized in section 2 for both the single lead time ANOVA
 135 and the quasi-ergodic ANOVA approaches. The Monte Carlo experiment for
 136 the simulation of multiple synthetic MMEs is presented in section 3. The
 137 precision of uncertainty estimates is presented in section 4, and discussed in
 138 section 5. Section 6 concludes.

139 2 ANOVA methods

140 2.1 Transient ensembles of climate experiments

141 Let us consider a hypothetical ensemble of climate experiments where M dif-
 142 ferent members are available for each of G different climate modeling chains.
 143 For the sake of simplicity, we assume that the same number of members is
 144 available for each chain. A chain refers for instance to a given GCM and the
 145 members to the different runs available for each GCM. A chain could also re-
 146 fer to a given GCM/RDM combination or to a given GCM/RDM/HM model
 147 combination. Members would respectively refer to the potentially multiple re-
 148 alizations obtained with the different runs for each GCM/RDM combination
 149 or for each GCM/RDM/HM combination (e.g. Lafaysse et al, 2014; Vidal et al,
 150 2016).

151 Note $Y(g, m, t)$ the climate projection obtained for the m^{th} member of
 152 chain g for time t . $Y(g, m, t)$ can be expressed as:

$$Y(g, m, t) = \lambda(g, t) + \nu(g, m, t), \quad (1)$$

153 where $\lambda(g, t)$ is the climate response of model g for year t and $\nu(g, m, t)$ is the
 154 deviation from the climate response obtained with the member m for this year
 155 as a result of internal variability (IV).

156 As in most climate impact studies, the uncertainty analysis can be carried
 157 out for the change variable X . In the following, we consider absolute changes,

158 but a similar analysis could be obtained for relative changes. Absolute changes
159 are expressed as:

$$X(g, m, t) = Y(g, m, t) - Y(g, m, t_C), \quad (2)$$

160 where $X(g, m, t)$ are the experiment outputs of the change variable for a given
161 member m of each chain g , each year t and where t_C refers to a given reference
162 year (or n -yr reference period centered on year t_C). These change outputs can
163 be written as:

$$X(g, m, t) = \varphi(g, t) + \eta(g, m, t), \quad (3)$$

164 where $\varphi(g, t)$ is the climate change response of chain g for t and where $\eta(g, m, t)$
165 are the residuals obtained for the m^{th} member of chain g as a result of IV.
166 In the following, we assume that the climate change response can be further
167 decomposed as:

$$\varphi(g, t) = \mu(t) + \alpha(g, t), \quad (4)$$

168 where $\mu(t)$ is the mean climate change response from all available chains at
169 projection lead time t , $\alpha(g, t)$ is the deviation from $\mu(t)$ for the climate change
170 response of chain g .

171 By assumption, no correlation is expected between the change responses
172 and the residuals η 's. If the residuals are additionally independent and iden-
173 tically distributed (i.i.d.), the total uncertainty of the change variable at t is
174 $\sigma_X^2(t) = s_\alpha^2(t) + \sigma_\eta^2(t)$ where $s_\alpha^2(t)$ is the sample variance of the α 's in equa-
175 tion (4) and $\sigma_\eta^2(t)$ the variance of the η 's in equation (3). They respectively
176 correspond to the model uncertainty and IV components of total uncertainty
177 variance for X . Note that, for any given chain g , no correlation is expected
178 between the values of the residuals $\nu(g, m, t)$ obtained from the raw variable
179 Y in the reference and in a future climate. The IV of the change variable X
180 is thus the sum of the IV estimates for Y obtained for the reference and the
181 future periods respectively.

182 In the following, we detail the two ANOVA approaches used to estimate (1)
183 the mean climate change response $\mu(t)$, (2) the deviations $\alpha(g, t)$ for the differ-
184 ent models, (3) the total uncertainty variance $\sigma_X^2(t)$, (4) its model uncertainty
185 $s_\alpha^2(t)$, and (5) its IV component $\sigma_\eta^2(t)$. We will also consider the Response-to-
186 Uncertainty ratio $R2U(t)$ defined as $R2U(t) = \mu(t)/\sigma_X(t)$, the fraction $F_\alpha(t)$
187 of total variance explained by model uncertainty and the fraction $F_\eta(t)$ of total
188 variance explained by internal variability. Moment estimators of variance
189 components are used here for both ANOVA approaches.

190 2.2 Single Time ANOVA method

191 The Single Time ANOVA method (STANOVA in the following) is by far the
192 most frequently used approach for the last years. It consists in estimating an
193 ANOVA model on the projected changes $X(g, m, t)$ for each time step t in
194 turn. For any given t , climate projections data available for the time steps
195 that precede and follow t are disregarded.

196 The STANOVA method uses an ANOVA model with fixed effects of the
 197 form of equations (3) and (4). The residuals are assumed to be i.i.d. In this
 198 configuration, an unbiased estimator of the residual's variance σ_η^2 at time t is
 199 the multimodel mean of the inter-member variance at t , expressed as (Mont-
 200 gomery, 2012):

$$\hat{\sigma}_{\eta,ST}^2(t) = \frac{1}{G(M-1)} \sum_{g=1}^G \sum_{m=1}^M \{X(g, m, t) - \hat{\varphi}_{ST}(g, t)\}^2 \quad (5)$$

201 where $\hat{\varphi}_{ST}(g, t)$ is an estimate of the mean climate change response for model
 202 g at t estimated over all M members as $\hat{\varphi}_{ST}(g, t) = \frac{1}{M} \sum_{m=1}^M X(g, m, t)$.

203 The model uncertainty variance is the inter-model sample variance of the
 204 deviations $\alpha(g, t)$ for all g , namely $s_\alpha^2(t) = \frac{1}{G-1} \sum_{g=1}^G \{\alpha(g, t)\}^2$. An unbiased
 205 estimator of $s_\alpha^2(t)$ under constraint $\sum_{g=1}^G \hat{\alpha}_{ST}(g, t) = 0$ is (Montgomery, 2012):
 206

$$\hat{s}_{\alpha,ST}^2(t) = \frac{1}{G-1} \sum_{g=1}^G \{\hat{\alpha}_{ST}(g, t)\}^2 - \frac{1}{M} \hat{\sigma}_{\eta,ST}^2(t) \quad (6)$$

207 where $\hat{\alpha}_{ST}(g, t)$ is an estimate of the deviation $\alpha(g, t)$ for model g , namely
 208 $\hat{\alpha}_{ST}(g, t) = \hat{\varphi}_{ST}(g, t) - \hat{\mu}_{ST}(t)$ with $\hat{\mu}_{ST}(t)$ the overall mean climate change
 209 response estimate at t , namely $\hat{\mu}_{ST}(t) = \frac{1}{G} \sum_{g=1}^G \hat{\varphi}_{ST}(g, t)$.

210 An unbiased estimator of the total variance for $X(g, m, t)$ finally reads:
 211 $\hat{\sigma}_{X,ST}^2(t) = \hat{s}_{\alpha,ST}^2(t) + \hat{\sigma}_{\eta,ST}^2(t)$.

212 2.3 Quasi-Ergodic ANOVA method

213 A time series ANOVA approach consists first in estimating for each chain g the
 214 climate response from the raw projections $Y(g, m, t)$. This is typically achieved
 215 by fitting a trend model to one of the members available for each chain (e.g.
 216 Hawkins and Sutton, 2009; Reintges et al, 2017). In the Quasi-Ergodic ANOVA
 217 method (QEANOVA in the following) considered here, the trend model is
 218 not obtained from one given member but from all the M members $Y(g, m, t)$
 219 available for the chain (Hingray and Saïd, 2014). The trend model is assumed
 220 to give an unbiased estimator of the climate response function $\lambda(g, t)$ of the
 221 chain (the limitations linked to the choice of the trend model are discussed in
 222 section 5.2). In this case, the trend model for Y gives an unbiased estimator
 223 of the climate change response function of the chain $\varphi(g, t)$, as:

$$\hat{\varphi}_{QE}(g, t) = \hat{\lambda}_{QE}(g, t) - \hat{\lambda}_{QE}(g, t_C) \quad (7)$$

224 where $\hat{\lambda}_{QE}(g, t)$ and $\hat{\lambda}_{QE}(g, t_C)$ are the trend estimates of the raw projections
 225 Y for the future and reference periods respectively.

226 Under the quasi-ergodic assumption, the IV of chain g for the raw data Y is
 227 estimated as the variance over time of the residuals from the climate response
 228 estimated for the chain, namely $\{Y(g, m, t) - \hat{\lambda}_{QE}(g, t)\}$. Note that all residuals

229 from all members are used here. If the experiment period covers T time steps,
 230 the IV of chain g is thus estimated from a sample of size $T \times M$. Assuming
 231 no correlation between residuals obtained for the reference and future periods
 232 and assuming that the IV variance for Y is constant over the entire experiment
 233 period, the IV of chain g for the change variable X is twice that of the raw
 234 variable Y (Hingray and Saïd, 2014).

235 An unbiased estimator of the IV component for X is derived from the
 236 multimodel mean of the IV estimates obtained for the G different chains re-
 237 spectively (Hingray and Blanchet, 2018, see also Appendix A, Eq.A.3 and
 238 Eq.A.6). When for each chain g , the trend model for Y can be expressed as a
 239 linear combination of L functions of time, it reads :

$$\hat{\sigma}_{\eta, QE}^2 = \frac{2}{G(TM - L)} \sum_{g=1}^G \sum_{m=1}^M \sum_{k=1}^T \left\{ Y(g, m, t_k) - \hat{\lambda}_{QE}(g, t_k) \right\}^2. \quad (8)$$

240 When the trend model is a simple linear function of time, $L = 2$.

241 An unbiased estimator of the sample model uncertainty variance $s_{\alpha}^2(t)$
 242 under constraint $\sum_{g=1}^G \hat{\alpha}_{QE}(g, t) = 0$ is (Hingray and Blanchet, 2018, see also
 243 Appendix A, Eq.A.5):

$$\hat{s}_{\alpha, QE}^2(t) = \frac{1}{G-1} \sum_{g=1}^G \left\{ \hat{\alpha}_{QE}(g, t) \right\}^2 - \frac{A(t, \mathcal{C})}{M} \hat{\sigma}_{\eta, QE}^2, \quad (9)$$

244 where $\hat{\alpha}_{QE}(g, t)$ is as previously an estimate of the deviation $\alpha(g, t)$ for model g
 245 expressed as $\hat{\alpha}_{QE}(g, t) = \hat{\varphi}_{QE}(g, t) - \hat{\mu}_{QE}(t)$ with $\hat{\mu}_{QE}(t) = \frac{1}{G} \sum_{g=1}^G \hat{\varphi}_{QE}(g, t)$
 246 the overall mean climate change response at t . $A(t, \mathcal{C})$ depends on the projec-
 247 tion lead time and on the different functions of time used for the trend model
 248 of Y (Hingray and Blanchet, 2018). It is positive and smaller than 1. The
 249 expression of $A(t, \mathcal{C})$ is given in Appendix A when the trend model is a simple
 250 linear function of time (Eq. A.7).

251 An unbiased estimator of total variance for $X(g, m, t)$ finally also reads:
 252 $\hat{\sigma}_{X, QE}^2(t) = \hat{s}_{\alpha, QE}^2(t) + \hat{\sigma}_{\eta, QE}^2$.

253 3 Simulations

254 In the following, we mainly focus on the precision of uncertainty estimators
 255 obtained with STANOVA and QEANOVA methods for the change variable
 256 $X(g, m, t_e)$ obtained between some future prediction lead time t_e and the refer-
 257 ence period t_C .

258 We construct via Monte Carlo simulations multiple synthetic MMEs where
 259 each MME is composed of $M \times G$ times series of raw projections $Y(g, m, t)$,
 260 with M the number of members for each of the G chains (see Appendix C
 261 for details). For reasons of simplification, each MME is simulated based on
 262 the assumptions that the climate response function of each modeling chain g

is a linear function of time and that the random deviations $\nu(g, m, t)$ due to internal variability come from independent and normally distributed random variables: $\nu(g, m, t) \sim \mathcal{N}(0, \sigma_\nu^2)$.

Each MME simulated for $Y(g, m, t)$ allows deriving the corresponding time series for $X(g, m, t)$. In the present analysis, for graphical simplification purposes, each MME for Y is constructed so that it leads, for X at time $t = t_e$, to a prescribed value of the Response-to-Uncertainty ratio $[\text{R2U}(t_e)]$ and to a prescribed value of the fractional variance $F_\eta(t_e)$ due to internal variability (Appendix C).

In the following, simulations were produced over the 1960-2100 period for a 10-yr average climate variable (see illustrative MMEs in figure 1). Decadal time series were next aggregated to 20-yr time series for all analyses presented in section 4, in order to match with the usually preferred 20-yr temporal resolution in climate analyses. $Y(g, m, t)$ time series are thus constituted from 7 values each ($T = 140/20$). The reference period is $t_C = 1990$.

For the sake of simplicity, a unique number of chains is considered in the following section ($G = 5$). The influence of the number of chains on the results is discussed in section 5.1. Most results are finally presented for prediction lead time $t_e = 2050$. t_e will be thus omitted in most following notations.

[FIGURE 1 HERE]

For each MME, both STANOVA and QEANOVA models are applied, allowing estimating the ratio $\mathbf{R}(\theta) = \hat{\theta}/\theta_{th}$ between the estimated value $\hat{\theta}$ and the theoretical (prescribed) value θ_{th} of different parameters θ : the grand ensemble mean response $\mu(t_E)$, the uncertainty variances $s_\alpha^2(t_E)$, $\sigma_\eta^2(t_E)$, $\sigma_X^2(t_E)$, the Response-to-Uncertainty ratio $[\text{R2U}(t_E)]$ and the contribution to total variance of internal variability and model uncertainty, $F_\eta(t_E)$ and $F_\alpha(t_E)$.

We run the simulations for different combinations of prescribed R2U and F_η values and for different numbers of members M . For each combination, we simulate 1,000 MMEs. We subsequently estimate, for each parameter θ , the mean $\mathbf{E}[\mathbf{R}(\theta)]$ and the standard deviation $\text{SD}[\mathbf{R}(\theta)]$ of the corresponding 1,000 ratios $\mathbf{R}(\theta)$.

Application of a QEANOVA method requires fitting a trend model to the raw climate projections Y . A parametric model is usually used for convenience. For simplicity, we here also assume that the trend is a linear function of time for the whole transient period $[t_1, t_T]$ (the influence of the trend model on the results is discussed in section 5.2). The estimation is made using ordinary least squares.

Equations (5) and (8) give unbiased estimators of σ_η^2 . Respectively, equations (6) and (9) give unbiased estimators of s_α^2 . $\mathbf{E}[\mathbf{R}(\theta)]$ obtained for σ_η^2 and s_α^2 with both STANOVA and QEANOVA approaches are therefore expected to equal one whatever the $(M, F_\eta, \text{R2U})$ configuration. This actually also applies to all other STANOVA and QEANOVA estimators (not shown).

We therefore focus on the precision of estimators from STANOVA and QEANOVA approaches, i.e. on results obtained for $\text{SD}[\mathbf{R}(\theta)]$, denoted SD below for shortness. The smaller the standard deviation, the better the precision of

the estimator. We will additionally look at the ratio SD_{ST}/SD_{QE} , denoting the gain in precision obtained with the QEANOVA approach with respect to the STANOVA approach. The larger this ratio, the larger the gain.

The uncertainty analysis with the STANOVA method requires at least two members for each chain. This is not the case for the QEANOVA method which can be carried out with no more than one single member for each chain. In the following, we also present, for information, results obtained with the QEANOVA method in case of single member MMEs.

4 Results

In our case and whatever the $(M, F_\eta, R2U)$ configuration, the distribution of the $R(\theta)$ ratios obtained respectively from the 1000 synthetic data sets was always found to be roughly normal (not shown). In a first approximation, the SD value of $R(\theta)$ thus determines the probability that the $R(\theta)$ ratio is outside of any $[1 \pm e]$ interval, that is the probability that the estimated value of the considered parameter is at least 100e% smaller or 100e% greater than the theoretical value. For instance, the probability that the ratio $R(\theta)$ is outside the $[0.5, 1.5]$ interval is 1%, 32% or 62% when SD is 0.2, 0.5 or 1 respectively. When the SD value is greater than 0.5, the probability for a large underestimation (let say $R(\theta) < 0.5$) or for a large overestimation (let say $R(\theta) > 1.5$) is thus non negligible (i.e. $p > 0.32$). When SD is greater than 1, a large underestimation or overestimation of the considered parameter is very likely (i.e. $p > 0.62$).

4.1 Variance components, total and fractional variance

SD results obtained for variance parameters $(\sigma_\eta^2, s_\alpha^2, \sigma_X^2, F_\eta, F_\alpha)$ are roughly independent on the R2U value prescribed for the simulation of the synthetic MMEs (not shown). They conversely depend on the value prescribed for F_η and on the number of members M . We here therefore only present results via the SD response functions in the (M, F_η) domain.

For the different variance parameters and both ANOVA methods, SD values are logically higher for small M numbers. Except for internal variability, they are also higher for large F_η values, highlighting the lower precision of the estimators in these configurations (figure 2). For all parameters, STANOVA SD values are also logically higher than QEANOVA SD values.

[FIGURE 2 HERE]

For internal variability estimates, σ_η^2 , SD values do only depend on M . For configurations with $M \leq 3$, STANOVA SD values are greater than 0.4, highlighting the possibility for significant errors in σ_η^2 estimates. Whatever the (M, F_η) configuration, QEANOVA SD values are at least 3 times smaller than STANOVA SD values. QEANOVA estimates are thus expected to be much closer to the theoretical value.

347 For model uncertainty estimates s_α^2 , the precision gain obtained with QEANOVA
 348 is at least 2.5. It is up to 4 for high F_η values (e.g. > 90%) or even more when
 349 M is smaller than 5. Whatever the (M, F_η) configuration, SD values are much
 350 larger than those obtained for internal variability. In the most unfavorable
 351 configurations (very low M and very high F_η values), STANOVA SD values
 352 are noticeably higher than 6! s_α^2 estimates are very likely to be largely over
 353 or underestimated. In those cases, a misestimation can also be obtained with
 354 QEANOVA but it would be somehow less critical.

355 Compared to s_α^2 and σ_η^2 , the potential error for total variance estimates,
 356 σ_X^2 , is likely to be small for both methods. SD values are globally smaller than
 357 0.2 for QEANOVA. For STANOVA, they are smaller than 0.3 except in the low
 358 M and high F_η configurations. This suggests that for any given dataset, errors
 359 obtained for s_α^2 and σ_η^2 partly compensate. The gain in precision obtained with
 360 QEANOVA is around 2.5, whatever the (M, F_η) configuration.

361 SD response functions obtained for fractional variances estimates, F_η and
 362 F_α , are very similar to those obtained for σ_η^2 and s_α^2 estimates respectively (not
 363 shown). Whatever the method, errors in F_η estimates are thus not expected
 364 to be very large. For F_α , large errors are conversely expected with STANOVA
 365 in the most unfavorable (M, F_η) configurations. The gain in precision obtained
 366 with QEANOVA is again at least 2.5.

367 4.2 Mean response and Response-to-Uncertainty ratio

368 We focus here on the estimates obtained for the mean change response μ and
 369 for the Response-to-Uncertainty ratio R2U. For both ANOVA methods and
 370 both parameters, SD values are again logically higher for smaller M and/or
 371 higher F_η values. SD values additionally depend on the theoretical value pre-
 372 scribed for the Response-to-Uncertainty ratio R2U. SD values decrease when
 373 the theoretical value for R2U increases.

374 Results are presented in Figure 3 for μ estimates and three theoretical
 375 R2U values (0.2, 1, and 5). They illustrate typical configurations encountered
 376 in climate analyses. A R2U value as high as 5 can be for instance typically
 377 obtained in the 2050s for large scale temperature variables (e.g. global mean
 378 temperature as in Hawkins and Sutton (2011)). A R2U value as low as 0.2
 379 can be conversely obtained for noisy systems, e.g. regional climate variables
 380 with a high natural variability (e.g. 20-yr average regional precipitation as in
 381 Hingray and Saïd (2014)).

382 For μ estimates, STANOVA SD values obtained when R2U = 0.2 are often
 383 greater than 0.5 and even exceed 2.5 for the most unfavorable (M, F_η) con-
 384 figurations. When R2U = 1, SD values are much smaller. They are lower than
 385 0.15 except for the low M and high F_η configurations. For larger prescribed
 386 R2U values, SD values tend to be negligible: they are always smaller than 0.03
 387 when R2U = 5. QEANOVA SD values present a similar dependence on M , F_η
 388 and R2U. They are however again much smaller than STANOVA SD values
 389 (roughly 2.5 times smaller whatever the M , F_η and R2U configuration).

[FIGURE 3 HERE]

For R2U estimates, the structure of SD response functions are very similar to those obtained for μ estimates (not shown). SD values are however systematically larger (from +20% for R2U=0.1 to more than +100% for R2U=5). For noisy systems and the most unfavorable (M, F_η) configurations, large errors are therefore expected with STANOVA. Errors are conversely expected to be low to moderate with QEANOVA estimates, which are at least 2.5 times more precise.

5 Discussion

5.1 Dependence to the setup and size of the projection ensemble

For the sake of simplicity, we considered that MMEs are composed of 5 modelling chains. This is much less than the number of chains currently available in recent CMIP MMEs. We also considered that the same number of members is available for all simulations chains. This is obviously never the case in real MMEs which are mostly unbalanced. In real MMEs, a large number of model combinations is also typically missing when the projections are produced with chains composed of different models (e.g. not all GCM/RCM combinations are available when projections are produced with different RCMs from different GCM outputs).

Whatever the ANOVA method, estimates of the different uncertainty components are expected to be more precise when more members are available, when the number of chains is larger and/or when the number of missing model combinations is smaller. In most of the cases however, a time series approach is still expected to produce more precise estimates of uncertainty components than a single time approach. This obviously results from the larger sample size used for the estimation which allows for a more precise estimate of the climate response and of the internal variability of each simulation chain.

As an illustration, the previous analysis was repeated with MMEs composed of 10 simulation chains instead of 5 (see Supplementary Material, Figure A1). As expected, the precision of all uncertainty components estimates is better with 10 chains (for both approaches) and QEANOVA estimates are more precise than STANOVA ones. In the present case, the gains in precision are actually roughly the same than those presented previously with 5 chains.

5.2 The trend estimation issue in the QEANOVA approach

In the present case and for convenience, time series of the synthetic MMEs are simulated based on the assumption that the climate response of each chain is a linear function of time. The QEANOVA trend function used to estimate the climate response of each chain is also assumed to be a linear function of time.

428 In real datasets of projections, the climate response likely evolves in a
429 more complex way, perhaps differently in different chains. In practice, the
430 precise form of the climate response is not known, and for small ensembles and
431 noisy systems it may be quite difficult to guess and/or fit. If the chosen trend
432 model were incorrect for some chain, the response function estimate could be
433 locally biased. Incorrect trend models could then lead to significant errors in
434 the estimation of the local effects of the chains, and additional uncertainty
435 would arise, particularly if internal variability is small. Because of this trend
436 estimation issue, the results of a time series approach can thus be sub-optimal.
437 The linear assumptions retained for practical reasons in Section 3, make the
438 QEANOVA results likely too optimistic since the assumed trend model is the
439 true one. Thus the precision gain with QEANOVA in Section 4 is likely to
440 overestimate the gain to be obtained with this method for real MMEs.

441 For real MMEs, non-linear models can be considered for the estimation of
442 the trend (e.g. 4rd order polynomial used in Hawkins and Sutton (2009) and
443 in Hingray and Said (2014) for temperature changes) and the risk of choosing
444 non relevant trend models can be reduced with dedicated statistical tests and
445 a visual check of the fit (if any expert knowledge can be used for instance to
446 assess the climatic relevance of the estimated trend).

447 The trend estimation issue in a time series ANOVA approach, which may
448 be critical for some MME configurations, requires thus specific care. Notwith-
449 standing, a time series approach is likely to present a significant advantage
450 in many cases, even when the trend is not linear, especially when internal
451 variability is large. For illustration, we considered the same set of MMEs as
452 previously but, instead of fitting a linear trend for estimating the climate re-
453 sponse of each simulation chain, we fitted a 3rd order polynomial function.
454 Results are presented in Figure A2 of Supplementary Material. As expected,
455 whatever the configuration in the (M, F_7) domain, the precision of QEANOVA
456 estimates is lower with the polynomial trend model, especially for noisy MMEs
457 and for small ensembles. Indeed, with a polynomial trend, the degrees of free-
458 dom are larger and in case of noisy systems, the local trend is often dominated
459 by internal variability which logically leads to a lower precision of the esti-
460 mated climate response. Nevertheless, QEANOVA is still much more precise
461 than STANOVA. Note that a more relevant experiment would have been to
462 simulate synthetic MMEs with non linear trends for the true trends but such
463 a simulation is unfortunately not convenient and would have required to make
464 different arbitrary and non representative choices (e.g. shape and amplitude
465 of non linearity to be retained for each chain). The experiment discussed here
466 thus likely again underplays the trend estimation issue in real MMEs where
467 the true trends are not linear. Fitting higher order polynomials when the true
468 trend is linear is indeed expected to introduce little error especially when
469 the sample is large (the fit essentially yields the linear trend), but, as men-
470 tioned previously, the opposite does not hold: fitting a linear trend when the
471 true trend is non linear (e.g. quadratic or cubic) will always yield local biases
472 which may be large.

473 Finally, another possible limitation of a time series ANOVA approach is
474 when internal variability is assumed to be constant over time. This assumption
475 may be not really valid as suggested by the recent work of Olonscheck and
476 Notz (2017). As presented by Hingray and Saïd (2014) or Bracegirdle et al
477 (2014), this assumption is not necessarily required and can also be relaxed.
478 The internal variability can for instance be assumed to vary gradually from
479 one lead time to the other (Hingray and Saïd, 2014). It can also be estimated
480 from the variance of the residuals within a moving temporal window centered
481 on the targeted prediction lead time (Bracegirdle et al, 2014).

482 5.3 Toward a local-QEANOVA approach

483 The potential limitations discussed above could make the apparent higher
484 attractiveness of a time series approach somehow reduced for some MMEs
485 configurations (e.g. MMEs with complex trend shapes and low noise).

486 However, if difficult to exploit over the whole available simulation period,
487 the QEANOVA principles can be at least used locally, in the temporal neigh-
488 borhood $\Omega(t)$ of each time point t . In such a local QEANOVA approach, the
489 climate response of any given chain g can be obtained for instance from lo-
490 cal linear regression, in which the trend is assumed to be smooth so that the
491 climate response $\lambda(g, t)$ of the chain can be approximated by a straight line
492 in the neighborhood of t (Cleveland, 1979). The residuals estimated from all
493 data in $\Omega(t)$ can next be used to assess the local internal variability of the
494 projections. When the uncertainty analysis is produced for the climate change
495 variable, the local linear regression and the estimation of the local internal
496 variability is done for both the reference time t_c and the future lead time t_e .
497 In this configuration, note that the internal variability no longer has to be
498 assumed constant between both periods.

499 The theoretical developments of this local QEANOVA approach, given in
500 Hingray and Blanchet (2018) are summarized in Appendix B. For illustration,
501 the local QEANOVA approach was also applied to the different simulation
502 ensembles considered in the previous sections. The neighborhood $\Omega(t)$ was here
503 chosen to be composed from the 3 consecutive time steps centered around t . In
504 this context, the risk that the linear trend is dominated by natural variability
505 is likely important for noisy systems and small ensembles. This configuration
506 would have led to a risk of underestimating the natural variability but this is
507 not the case here thanks to the unbiased estimator derived for this variable
508 (see Appendix B).

509 Results of the local QEANOVA are presented for internal variability, model
510 uncertainty and for the Response-to-Uncertainty ratio in figure 4. Results are
511 not a surprise. For all uncertainty parameters, the precision obtained with the
512 local QEANOVA is significantly higher than that obtained with STANOVA
513 (confidence intervals are 2 to 3 times smaller for the local QEANOVA). The
514 precision of the local QEANOVA is conversely smaller than that obtained with

515 QEANOVA. These results again logically derive from the different sizes of the
 516 samples used by the three methods.

517 A local QEANOVA is thus a promising alternative to STANOVA. The risk
 518 for a large misestimation of uncertainty components is indeed much lower.
 519 Despite its smaller precision, a local QEANOVA can be also a good alternative
 520 to a time series approach. When the time series of the simulation chain presents
 521 a low to moderate noise but has a complex and nonlinear climate response
 522 for instance, a local QEANOVA actually reduces the risk of having a biased
 523 estimate of the climate response, which may occur in a time series approach
 524 because of the choice of an inappropriate trend function (see section 5.2).

525 [FIGURE 4 HERE]

526 5.4 Temporal variations of uncertainty component estimates

527 Whatever the estimation method, the uncertainty estimates are typically de-
 528 rived for the different prediction lead times of a given climate simulation pe-
 529 riod. The temporal behavior of these estimates is worth considering as it may
 530 be a relevant indicator of their precision. In most cases, the climate response
 531 of each simulation chain is indeed expected to evolve gradually with time. In
 532 most cases, the different uncertainty components of a given MME are thus
 533 also expected to evolve gradually with time.

534 Similarly, uncertainty components for model uncertainty and scenario un-
 535 certainty should start from zero when the climate variable considered for the
 536 analysis is a climate change variable. The climate change response of any sim-
 537 ulation chain is indeed expected to start from zero when the prediction lead
 538 time corresponds to the control period.

539 Large temporal fluctuations of uncertainty estimates and/or large values
 540 of model uncertainty for the control period should warn for the possible lack
 541 of robustness of the analysis. The possibility to consider the temporal varia-
 542 tions of uncertainty estimates to assess their precision is illustrated for three
 543 synthetic data sets in Figure 5 for s_α^2 and σ_η^2 and in Figure 6 for F_η . The
 544 three MMEs have been simulated following the simulation process described
 545 in section 3. Confidence intervals (colored bounds) of uncertainty estimates
 546 (white lines), obtained from Monte Carlo simulations similar to those carried
 547 out in section 4, highlight the much lower precision of STANOVA, even when
 548 the fractional variance due to internal variability is low. This is also clear from
 549 the temporal variations of uncertainty estimates. When only 3 members are
 550 available, the variations of STANOVA estimates are erratic and non-negligible
 551 whatever the value of F_η . They are much smoother with a local QEANOVA
 552 and they completely vanish with QEANOVA. As expected also, the model
 553 uncertainty component estimated with QEANOVA starts from zero at the be-
 554 ginning of the century and then gradually increases. This is thus fully coherent
 555 with the expected behavior of this uncertainty component (limit to zero for
 556 the control period). This is not the case for the two other methods, especially

557 for STANOVA in the case of small ensembles and noisy systems (e.g. signifi-
558 cant non zero values of model uncertainty can be achieved at time $t = 2000$
559 in Figure 5, STANOVA, 2nd line).

560 [FIGURE 5 HERE]

561 [FIGURE 6 HERE]

562 6 Conclusion

563 6.1 Precision of different methods for partitioning uncertainty

564 Two ANOVA approaches are currently used for partitioning model uncertainty
565 and internal variability components in ensembles of climate projections: the
566 popular Single Time approach (STANOVA), based on the data available for
567 the considered projection lead time and time series approaches such as the
568 QEANOVA approach considered here. For multiple synthetic ensembles, we
569 compared the precision of uncertainty component estimates obtained with
570 both approaches. The main findings are as follows:

- 571 • In both cases, the precision is logically lower for noisy MMEs and/or small
572 ensembles of members. The precision of the mean climate response and of
573 the Response-to-Uncertainty ratio estimates are additionally lower when
574 the theoretical (prescribed) Response-to-Uncertainty ratio is lower.
- 575 • In many configurations, QEANOVA estimates are expected to be more
576 precise than STANOVA ones. This does not depend on the setup and size
577 of the MME considered (unbalanced MMEs, MMEs with missing chains,
578 MMEs with more GCMs,...). The larger sample size used for the estimation
579 is indeed expected to produce a better estimate of the climate response and
580 of the internal variability of each simulation chain.
- 581 • Whatever the estimation method, the temporal variations of the uncer-
582 tainty estimates should be considered as a relevant indicator of the preci-
583 sion of estimates. Large temporal fluctuations should call for caution and
584 should warn for the likely low relevance of the uncertainty analysis.

585 Under the specific setup and assumptions of our analysis framework (e.g.
586 length of time series, number of simulations chains, linearity of trend, con-
587 stancy of internal variability over time...), we find that:

- 588 • Simulated confidence intervals obtained with QEANOVA are at least 2.5
589 times smaller than those obtained with STANOVA.
- 590 • The precision of estimates is rather high for total uncertainty variance
591 and for the multichain mean response (except for STANOVA for noisy
592 MMEs and/or for MMEs with small ensembles). The two main features
593 of any given ensemble of projections are thus likely to be rather correctly
594 estimated in most cases.

- 595 • This is obviously not the case for the individual uncertainty components.
596 For model uncertainty variance for instance, the precision of estimates is
597 low to very low for small ensembles and/or noisy MMEs. In the latter
598 configurations, a large over- or underestimation of uncertainty components
599 is thus very likely preventing any precise partition of uncertainty sources.
600 This problem is particularly critical for STANOVA.
- 601 • STANOVA actually appears to be inappropriate in a large number of con-
602 figurations as a result of the limited number of members, e.g. GCM runs,
603 usually available for most modeling chains up to now. In the case of small
604 ensembles and noisy MMEs, improving the relevance of uncertainty esti-
605 mates requires using more data. A time series analysis approach as the
606 QEANOVA method offers such an opportunity.

607 For a time series approach, the trend estimation can be an issue. For MMEs
608 where internal variability is small for instance, there is some risk to misrep-
609 resent the forced trend. In such a case, a local QEANOVA approach offers
610 an interesting alternative as it is not subject to the same limitation as a full
611 time series approach. The precision of a local QEANOVA is additionally al-
612 ways better than that of STANOVA. Note also that, similarly to a time series
613 approach, a local QEANOVA can be also applied when only a single member
614 is available for the chains. It can therefore make use of all climate experiments
615 produced in the recent years. This is another advantage over STANOVA which
616 requires multiple members of each chain.

617 6.2 Precision of uncertainty estimates for real ensembles of projections

618 Most ensembles of projections have potentially more uncertainty sources than
619 those considered in the present work. They may actually include scenario un-
620 certainty, different components of model uncertainty (e.g. associated respec-
621 tively to GCMs, Hydrological Models), of internal variability (with potentially
622 its large or small scale components) and, eventually, uncertainty associated to
623 interactions between models and/or scenarios (e.g. Hawkins and Sutton, 2011;
624 Vidal et al, 2016; Paeth et al, 2017). In all cases, and whatever the ANOVA
625 method used for the analysis, an idea of the precision of uncertainty compo-
626 nents should be provided along with estimates of uncertainty components.

627 Results presented in the present work give already some benchmark for
628 simple ensembles of projections. They were obtained with moment estimators
629 of variance components. Likelihood-based estimation methods such as max-
630 imum likelihood, restricted maximum likelihood or Bayesian methods could
631 (should) be used instead. They indeed allow for the estimation of confidence
632 or credibility intervals associated to each uncertainty component. They also
633 present next a number of interesting properties (e.g. Hartley and Rao, 1967;
634 Gelman, 2005). Conversely to method-of-moment based ANOVA approaches,
635 they can especially not produce negative estimates of variance components.
636 For the two ANOVA methods considered here, negative estimates are indeed
637 obtained for model uncertainty variance when the second terms in the right

638 hand sides of equations (6) or (9) are greater than the first terms. The prob-
639 ability of having a negative estimate for s_α^2 is actually directly related to the
640 precision that can be achieved with the method for this component. For our
641 work, this probability is roughly 2.5%, 16% or 31% when the SD value ob-
642 tained for s_α^2 is 0.5, 1 or 2 respectively. It can be up to 50% in the most
643 critical ($M, F_\eta, R2U$) configurations.

644 Likelihood-based estimation methods are also expected to use the data
645 more efficiently than moment-based estimation methods. They are for instance
646 easily applicable for unbalanced MMEs and/or MMEs with missing chains.
647 The work of Northrop and Chandler (2014) and of Evin et al (2018) give some
648 nice illustrations of their uses in a Bayesian context. The use of likelihood-
649 based estimation methods is thus to be preferred to the use of moment-based
650 estimation methods. Even with such methods however, our results strongly
651 suggest that a full- or partial time series approach should be used instead of
652 a standard single time approach.

653 **Author contributions.** BH designed the analysis, developed the local-
654 QEANOVA and the synthetic simulations. JB derived the theoretical expres-
655 sions for unbiased estimators of uncertainty components and wrote the ap-
656 pendixes. All authors contributed to write the manuscript and discuss results.

657 **Acknowledgements** We thank the three anonymous reviewers for their constructive sug-
658 gestions which helped to significantly improve the content of our manuscript.

659 Appendix A QEANOVA Estimates of uncertainty components

660 We summarize the theoretical developments of the QEANOVA approach to
 661 achieve unbiased estimators of uncertainty components for a simplified con-
 662 figuration where the trend model is a simple linear function of time. The full
 663 developments are given in Hingray and Blanchet (2018) for the general config-
 664 uration where the trend model is a linear combination of L functions of time.
 665 Here, we first consider the case where the number of members differs from one
 666 chain to the other; the simplified equations obtained when all chains have the
 667 same number are then given. For the sake of conciseness, we omit the subscript
 668 "QE" related to the QEANOVA approach.

669 *Model.* We first consider the raw projections $Y(g, m, t)$ with M_g members for
 670 each of the G chains, assuming that, for all $t_s \leq t \leq t_f$:

$$Y(g, m, t) = \lambda(g, t) + \nu(g, m, t), \quad (\text{A.1})$$

671 where $\lambda(g, t)$ is the trend model expressed as a linear function of time:
 672 $\lambda(g, t) = \Lambda_{g1} + \Lambda_{g2}(t - t_s)$ and where the $\nu(g, m, t)$ are independent and
 673 homoscedastic random variables (with variance $\sigma_{\nu_g}^2$). Second let consider the
 674 change variable at future prediction lead time $t \in [t_c, t_f]$:

$$X(g, m, t) = Y(g, m, t) - Y(g, m, t_c) = \alpha(g, t) + \eta(g, m, t), \quad (\text{A.2})$$

675 where t_c is the reference period. We have thus $\alpha(g, t) = \Lambda_{g2}(t - t_c)$ and
 676 $\eta(g, m, t) = \nu(g, m, t) - \nu(g, m, t_c)$.

Unbiased estimation of the model parameters for the raw variable Y . We dis-
 cretize $[t_s, t_f]$ into T time steps (from $t_s = t_1$ to $t_f = t_T$) and write t_c as the
 K th time step (i.e. $t_c = t_K$). We are interested in the future prediction lead
 time $t_k \in [t_K, t_T]$. Let consider the regression model (A.1) for a particular g .
 Unbiased estimators of the regression parameters $(\Lambda_{g1}, \Lambda_{g2})$ are given by the
 least square estimates

$$(\hat{\Lambda}_{g1}, \hat{\Lambda}_{g2})' = \mathbb{V} \mathbb{R}' \left(\frac{1}{M_g} \sum_{m=1}^{M_g} Y(g, m, t_1), \dots, \frac{1}{M_g} \sum_{m=1}^{M_g} Y(g, m, t_T) \right)'$$

677 where "''" denotes the transpose, \mathbb{R} is the $T \times 2$ matrix of covariates whose
 678 k th row is $(1, t_k - t_1)$, for $1 \leq k \leq T$, and $\mathbb{V} = (\mathbb{R}'\mathbb{R})^{-1}$. Covariance matrix
 679 of the estimators $(\hat{\Lambda}_{g1}, \hat{\Lambda}_{g2})$ is given by $\widehat{\sigma_{\nu_g}^2} M_g^{-1} \mathbb{V}$ where $\widehat{\sigma_{\nu_g}^2}$ is an unbiased
 680 estimator of $\sigma_{\nu_g}^2$ given by

$$\widehat{\sigma_{\nu_g}^2} = \frac{1}{TM_g - L} \sum_{m=1}^{M_g} \sum_{k=1}^T \left\{ Y(g, m, t_k) - \hat{\lambda}_{QE}(g, t_k) \right\}^2 \quad (\text{A.3})$$

681 where $L = 2$ and $\hat{\lambda}_{QE}(g, t_k) = \hat{\Lambda}_{g1} - \hat{\Lambda}_{g2}(t_k - t_1)$.

682 In particular, an unbiased estimator of A_{g2}^2 is

$$\widehat{A}_{g2}^2 = \widehat{\Lambda}_{g2}^2 - \widehat{\sigma}_{\nu_g}^2 M_g^{-1} V_{22}, \quad (\text{A.4})$$

683 where V_{22} is the element (2, 2) of \mathbb{V} . Considering t_1, \dots, t_T regularly spaced
684 on $[t_s; t_f]$, we have $V_{22} = 12(T-1)/\{T(T+1)(t_T - t_1)^2\}$.

685 *Unbiased estimation of the sample variance of the α 's in the change vari-*
686 *able.* Given (A.2) and (A.4), an unbiased estimator of the sample variance of
687 $\alpha(g, t_k)$, i.e. of $s_\alpha^2(t_k) = \frac{1}{G-1} \sum_{g=1}^G \{\alpha(g, t_k)\}^2$, is

$$\widehat{s}_\alpha^2(t_k) = s_{\widehat{\alpha}}^2(t_k) - \frac{12T-1}{T(T+1)} \left(\frac{t_k - t_K}{t_T - t_1} \right)^2 \left(\frac{1}{G} \sum_{g=1}^G \frac{\widehat{\sigma}_{\nu_g}^2}{M_g} \right),$$

where

$$s_{\widehat{\alpha}}^2(t_k) = \frac{(t_k - t_1)^2}{G-1} \sum_{g=1}^G \widehat{\Lambda}_{g2}^2.$$

688 When all GCMs have the same number of runs (M), this expression reduces
689 to:

$$\widehat{s}_\alpha^2(t_k) = s_{\widehat{\alpha}}^2(t_k) - \frac{A(t_k, \mathcal{C})}{M} \widehat{\sigma}_\eta^2, \quad (\text{A.5})$$

690 where $\widehat{\sigma}_\eta^2$ is an unbiased estimator of internal variability variance for X

$$\widehat{\sigma}_\eta^2 = \frac{2}{G} \sum_{g=1}^G \widehat{\sigma}_{\nu_g}^2 \quad (\text{A.6})$$

691 and where

$$A(t_k, \mathcal{C}) = \frac{6(T-1)}{T(T+1)} \left(\frac{t_k - t_K}{t_T - t_1} \right)^2 \quad (\text{A.7})$$

692

693 Appendix B Estimates with a local QEANOVA approach

694 We here summarize the expressions of the different uncertainty estimators
695 obtained with a local-QEANOVA approach. The full developments, similar to
696 those presented in appendix A for the QEANOVA approach, are detailed in
697 Hingray and Blanchet (2018).

698 *Model.* A regression model is still considered to estimate the response function
699 $\lambda(g, t)$ for Y in Eq.A.1 but $\lambda(g, t)$ is assumed to be only locally linear in time,
700 in the neighborhoods of t_c and t_e respectively, i.e. on $[t_c - \omega, t_c + \omega]$ and
701 $[t_e - \omega, t_e + \omega]$, where $t_e \in [t_c, t_f]$ is the future prediction lead time under
702 consideration. $\lambda(g, t)$ can thus be expressed as

$$\lambda(g, t) = \begin{cases} \lambda_c(g, t) = A_{g1,c} + (t - t_c)A_{g2,c} & \text{for } t_c - \omega \leq t \leq t_c + \omega, \\ \lambda_e(g, t) = A_{g1,e} + (t - t_e)A_{g2,e} & \text{for } t_e - \omega \leq t \leq t_e + \omega. \end{cases} \quad (\text{B.1})$$

703 The change variable $X(g, m, t)$, for $t_e - \omega \leq t \leq t_e + \omega$, in Eq.A.2 is such
704 that $\alpha(g, t) = (A_{g1,e} - A_{g1,c}) + (t - t_e)A_{g2,e}$.

705 Each interval $[t_c - \omega, t_c + \omega]$ and $[t_e - \omega, t_e + \omega]$ is discretized into T^* regular
706 periods of length $dt = 2\omega/(T^* - 1)$, with T^* odd, giving respectively the
707 sequences t_1, \dots, t_{T^*} and $t_{T^*+1}, \dots, t_{2T^*}$. The values of Y for these different
708 times are further considered to estimate the regression coefficients of the linear
709 trend models in Eq.B.1. For the illustration given in section 5.3, $dt = \omega =$
710 20yrs and $T^* = 3$.

711 *Unbiased estimators of model uncertainty and internal variability variance.*
712 Following Hingray and Blanchet (2018), an unbiased estimator of model un-
713 certainty variance, i.e. of the sample variance of $\alpha(g, t_e)$, is

$$\widehat{s}_\alpha^2(t_e) = s_\alpha^2(t_e) - \frac{4}{T} \left(\frac{1}{G} \sum_{g=1}^G \widehat{\sigma}_{\nu_g}^2 \right) \quad (\text{B.2})$$

714 where $T = 2T^*$ is the total number of time steps considered in the analysis
715 and where $\widehat{\sigma}_{\nu_g}^2$ is an unbiased estimator of $\sigma_{\nu_g}^2$ given by

$$\widehat{\sigma}_{\nu_g}^2 = \frac{1}{TM_g - 4} \sum_{m=1}^{M_g} \left[\sum_{j=1}^{T^*} \left\{ Y(g, m, t_j) - \hat{\lambda}_c(g, t_j) \right\}^2 \right. \quad (\text{B.3})$$

$$\left. + \sum_{j=T^*+1}^{2T^*} \left\{ Y(g, m, t_j) - \hat{\lambda}_e(g, t_j) \right\}^2 \right] \quad (\text{B.4})$$

716 with $\hat{\lambda}_c(g, t_j) = \hat{A}_{g1,c} + \hat{A}_{g2,c}(t_j - t_c)$ and $\hat{\lambda}_e(g, t_j) = \hat{A}_{g1,e} + \hat{A}_{g2,e}(t_j - t_e)$
717 where $\hat{A}_{g1,c}$, $\hat{A}_{g2,c}$, $\hat{A}_{g1,e}$ and $\hat{A}_{g2,e}$ are the regression coefficients of the two
718 linear models in Eq.B.1.

719 When all GCMs have the same number of runs (M), the expression of
720 model uncertainty in Eq.B.2 reduces to

$$\widehat{s}_\alpha^2(t_e) = s_\alpha^2(t_e) - \frac{1}{MT^*} \widehat{\sigma}_\eta^2 \quad (\text{B.5})$$

721 where $\widehat{\sigma}_\eta^2$ is an unbiased estimator of internal variability variance for X

$$\widehat{\sigma}_\eta^2 = \frac{2}{G} \sum_{g=1}^G \widehat{\sigma}_{\nu_g}^2. \quad (\text{B.6})$$

722 Appendix C Simulation of MMEs

723 Each MME is simulated for $Y(g, m, t)$ assuming that, for all $t_s \leq t \leq t_f$:

$$Y(g, m, t) = \lambda(g, t) + \nu(g, m, t), \quad (\text{C.1})$$

724 where the climate response function $\lambda(g, t)$ is a linear function of time
 725 and where the $\nu(g, m, t)$ are independent and homoscedastic random variables
 726 (with variance $\sigma_{\nu_g}^2$). For convenience we further assume that $\lambda(g, t)$ can be
 727 decomposed as $\lambda(g, t) = w(t) + d(g, t)$, for $t = t_1, \dots, T$ where the mean
 728 climate response $w(t)$ of the G chains and the deviations $d(g, t)$ of chain g are
 729 linear functions of time, expressed as: $w(t) = B + P.(t - t_1)/(t_e - t_C)$ and
 730 $d(g, t) = D(g).(t - t_1)/(t_e - t_C)$ with the constraint $\sum_{g=1}^G D(g) = 0$.

731 For graphical simplification purposes, each MME for Y is constructed so
 732 that the parameters $(\sigma_{\nu}^2, P, D(g), g = 1, \dots, G)$ lead, for the change variable
 733 X at time $t = t_e$, to a prescribed value of the Response-to-Uncertainty ratio
 734 $[R2U(t_e)]$ and to a prescribed value of the fractional variance $F_{\eta}(t_e)$ due to
 735 internal variability.

736 For X , we have by definition $\varphi(g, t) = \lambda(g, t) - \lambda(g, t_C)$ and $\eta(g, m, t) =$
 737 $\nu(g, m, t) - \nu(g, m, t_C)$. We have thus for the change variable $\mu(t) = w(t) -$
 738 $w(t_C)$, $\alpha(g, t) = d(g, t) - d(g, t_C)$ and in turn $\mu(t) = P.(t - t_C)/(t_e - t_C)$ and
 739 $\alpha(g, t) = D(g).(t - t_C)/(t_e - t_C)$.

740 The theoretical values for $\mu(t_e)$ and $s_{\alpha}^2(t_e)$ are thus as follows : $\mu(t_e) = P$
 741 and $s_{\alpha}^2(t_e) = \text{Var}(D(g))$. Fixing P to 1, we thus simply require in turn

$$\sigma_X^2(t_e) = \frac{1}{[R2U(t_e)]^2}; \quad (\text{C.2})$$

$$\sigma_{\nu}^2(t_e) = \frac{1}{2}\sigma_{\eta}^2(t_e) = \frac{1}{2}F_{\eta}(t_e).\sigma_X^2(t_e); \quad (\text{C.3})$$

$$\text{Var}(D(g)) = s_{\alpha}^2(t_e) = (1 - F_{\eta}(t_e)).\sigma_X^2(t_e). \quad (\text{C.4})$$

742 For each MME simulation, the deviations of the different chains, $D(g), g =$
 743 $1, \dots, G$, are obtained from a sample of G realizations in a normal distribution.
 744 These realizations are scaled so that their mean is zero and their variance
 745 corresponds to the prescribed value $s_{\alpha}^2(t_e)$.

746 References

- 747 Bosshard T, Carambia M, Goergen K, Kotlarski S, Krahe P, Zappa M, Schaer
 748 C (2013) Quantifying uncertainty sources in an ensemble of hydrological
 749 climate-impact projections. *Water Resources Research* 49(3):1523–1536
- 750 Bracegirdle TJ, Turner J, Hosking JS, Phillips T (2014) Sources of uncer-
 751 tainty in projections of twenty-first century westerly wind changes over the
 752 Amundsen Sea, West Antarctica, in CMIP5 climate models. *Climate Dy-*
 753 *namics* 43(7-8):2093–2104
- 754 Braun M, Caya D, Frigon A, Slivitzky M (2012) Internal variability of the
 755 Canadian RCMs hydrological variables at the basin scale in Quebec and
 756 Labrador. *Journal of Hydrometeorology* 13(2):443–462
- 757 Charlton-Perez AJ, Hawkins E, Eyring V, Cionni I, Bodeker GE, Kinnison
 758 DE, Akiyoshi H, Frith SM, Garcia R, Gettelman A, Lamarque JF, Naka-
 759 mura T, Pawson S, Yamashita Y, Bekki S, Braesicke P, Chipperfield MP,

- 760 Dhomse S, Marchand M, Mancini E, Morgenstern O, Pitari G, Plummer D,
761 Pyle JA, Rozanov E, Scinocca J, Shibata K, Shepherd TG, Tian W, Waugh
762 DW (2010) The potential to narrow uncertainty in projections of strato-
763 spheric ozone over the 21st century. *Atmospheric Chemistry and Physics*
764 10(19):9473–9486
- 765 Cleveland S W (1979) Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* 74(368):829–836
- 766 Deser C, Phillips A, Bourdette V, Teng H (2012) Uncertainty in climate change
767 projections: the role of internal variability. *Climate Dynamics* 38(3-4):527–
768 546
- 769 Evin G, Eckert N, Hingray B, Verfaillie D, Morin S, Lafaysse M, JBlanchet
770 (2018) Traiter l’incertitude des projections climatiques: l’exemple des condi-
771 tions meteorologiques et d’enneigement dans les Alpes francaises. *Schweizerische Zeitschrift fur Forstwesen* 169(4):203–209
- 772 Gelman A (2005) Analysis of variance why it is more important than ever. *Annals of Statistics* 33(1):1–53
- 773 Giuntoli J, Vidal JP, Prudhomme C, Hannah DM (2015) Future hydrological
774 extremes: the uncertainty from multiple global climate and global hydrologi-
775 cal models. *Earth Syst Dynam* 16:267–285
- 776 Hartley HO, Rao JNK (1967) Maximum-likelihood estimation for the mixed
777 analysis of variance model. *Biometrika* 54(1/2):93–108
- 778 Hawkins E, Sutton R (2009) The potential to narrow uncertainty in regional
779 climate predictions. *Bulletin of the American Meteorological Society*
780 90(8):1095–1107
- 781 Hawkins E, Sutton R (2011) The potential to narrow uncertainty in projections
782 of regional precipitation change. *Climate Dynamics* 37(1–2):407–418
- 783 Hingray B, Blanchet J (2018) Uncertainty components estimates in transient
784 climate projections. bias of moment-based estimators in the single time and
785 time series approaches. Tech. rep., IGE, Univ-Grenoble Alpes, Grenoble,
786 FR, <https://hal.archives-ouvertes.fr/hal-01738218>
- 787 Hingray B, Saïd M (2014) Partitioning internal variability and model uncer-
788 tainty components in a multimodel multimember ensemble of climate pro-
789 jections. *Journal of Climate* 27(17):6779–6798
- 790 Hingray B, Mezghani A, Buishand TA (2007) Development of probability dis-
791 tributions for regional climate change from uncertain global mean warming
792 and an uncertain scaling relationship. *Hydrology and Earth System Sciences*
793 11(3):1097–1114
- 794 Jacob D, Petersen J, Eggert B, Alias A, Christensen OB, Bouwer LM, Braun
795 A, Colette A, Deque M, Georgievski G, Georgopoulou E, Gobiet A, Menut
796 L, Nikulin G, Haensler A, Hempelmann N, Jones C, Keuler K, Kovats S,
797 Kroener N, Kotlarski S, Kriegsmann A, Martin E, van Meijgaard E, Moseley
798 C, Pfeifer S, Preuschmann S, Radermacher C, Radtke K, Rechid D, Roun-
799 sevell M, Samuelsson P, Somot S, Soussana JF, Teichmann C, Valentini R,
800 Vautard R, Weber B, Yiou P (2014) EURO-CORDEX: new high-resolution
801 climate change projections for European impact research. *Regional Envi-
802 ronmental Change* 14(2):563–578
803
804
805

- 806 Johns T, Royer JF, Höschel I, Huebener H, Roeckner E, Manzini E, May W,
807 Dufresne JL, Otteå O, Vuuren D, Salas y Melia D, Giorgetta M, Denvil S,
808 Yang S, Fogli P, Körper J, Tjiputra J, Stehfest E, Hewitt C (2011) Climate
809 change under aggressive mitigation: the ensembles multi-model experiment.
810 *Climate Dynamics* 37(9–10):1975–2003
- 811 Kendon EJ, Rowell DP, Jones RG, Buonomo E (2008) Robustness of future
812 changes in local precipitation extremes. *Journal of Climate* 21(17):4280–
813 4297
- 814 Kew SF, Selten FM, Lenderink G, Hazeleger W (2011) Robust assessment of
815 future changes in extreme precipitation over the Rhine basin using a GCM.
816 *Hydrology and Earth System Sciences* 15(4):1157–1166
- 817 Lafaysse M, Hingray B, Mezghani A, Gailhard J, Terray L (2014) Internal vari-
818 ability and model uncertainty components in a multireplicate multimodel
819 ensemble of hydrometeorological projections : the Alpine Durance basin.
820 *Water Resources Research* 50:3317–3341
- 821 Montgomery DC (2012) *Design and Analysis of Experiments.*, 8th edn. Wiley
- 822 Northrop PJ, Chandler RE (2014) Quantifying sources of uncertainty in pro-
823 jections of future climate. *Journal of Climate* 27(23):8793–8808
- 824 Olonscheck D, Notz D (2017) Consistently Estimating Internal Climate Vari-
825 ability from Climate Model Simulations. *Journal of Climate* 30(23):9555–
826 9573, DOI {10.1175/JCLI-D-16-0428.1}
- 827 Paeth H, Vogt G, Paxian A, Hertig E, Seubert S, Jacobeit J (2017) Quantifying
828 the evidence of climate change in the light of uncertainty exemplified by the
829 Mediterranean hot spot region. *Global and Planetary Change* 151(SI):144–
830 151, DOI {10.1016/j.gloplacha.2016.03.003}
- 831 van Pelt SC, Beersma JJ, Buishand TA, van den Hurk BJJM, Schellekens J
832 (2015) Uncertainty in the future change of extreme precipitation over the
833 Rhine basin: the role of internal climate variability. *Climate Dynamics* 44(7-
834 8):1789–1800
- 835 Räisänen J (2001) CO₂-induced climate change in CMIP2 experiments: Quan-
836 tification of agreement and role of internal variability. *Journal of Climate*
837 14(9):2088–2104
- 838 Reintges A, Martin T, Latif M, Keenlyside NS (2017) Uncertainty in twenty-
839 first century projections of the Atlantic Meridional Overturning Circulation
840 in CMIP3 and CMIP5 models. *Climate Dynamics* 49(5-6):1495–1511, DOI
841 {10.1007/s00382-016-3180-x}
- 842 Vidal JP, Hingray B, Magand C, Sauquet E, Ducharne A (2016) Hierar-
843 chy of climate and hydrological uncertainties in transient low-flow pro-
844 jections. *Hydrology and Earth System Sciences* 20(9):3651–3672, DOI
845 {10.5194/hess-20-3651-2016}
- 846 Yip S, Ferro CAT, Stephenson DB, Hawkins E (2011) A simple, coherent
847 framework for partitioning uncertainty in climate predictions. *Journal of*
848 *Climate* 24(17):4634–4643

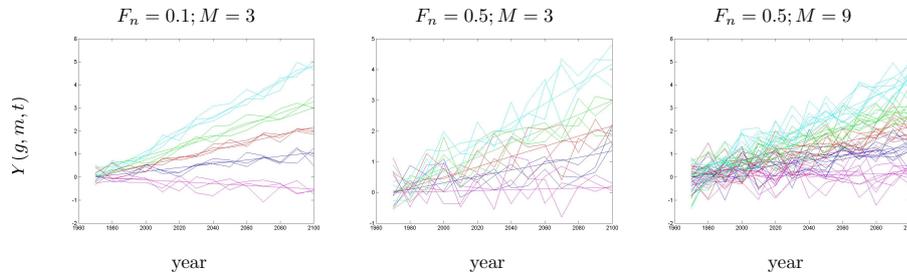


Fig. 1 Three synthetic multimodel multimember ensembles of time series projections (MMEs) obtained via simulation with different F_n ratios and M numbers (left: $F_n = 0.1$, $M = 3$; middle $F_n = 0.5$, $M = 3$; right: $F_n = 0.5$, $M = 9$). Time series over the 1960-2100 period for the raw projections $Y(g, m, t)$ of a 10-yr average climate variable. Cases on the left and in the middle are illustrative of typical configurations encountered in climate analyses: "small" number of members available for each GCM ($M=3$) with relatively low (resp. high) contribution of IV to total variance (case 1: $F_n = 0.1$; case 2: $F_n = 0.5$). The case on the right is similar as the middle case but in a configuration with a (unrealistic) "large" number of members ($M = 9$).

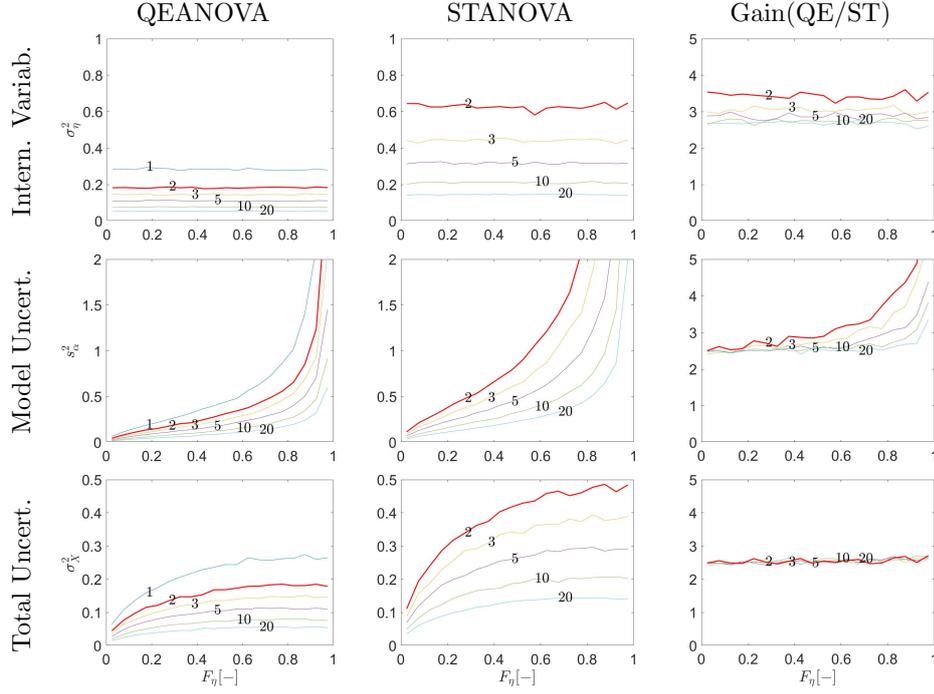


Fig. 2 Precision of QEANOVA (left) and STANOVA (middle) estimates for uncertainty variance components, and gain in precision between STANOVA SD values and QEANOVA SD values (right)(gain = ratio between STANOVA SD values and QEANOVA SD values). Top: Internal Variability, middle: Model Uncertainty, bottom: Total Uncertainty. SD values are given as a function of the fraction of total variance explained by internal variability (F_η) for a few representative values of number of members M : (1), 2, 3, 5, 10, 20. Results are presented for a theoretical ratio $R2U = 1$. For the sake of clarity, the upper limit of the figures for s_α^2 is truncated to 2. The highest values are greater than 10. Figures of the ratios between SD values obtained with both approaches are truncated to 5.

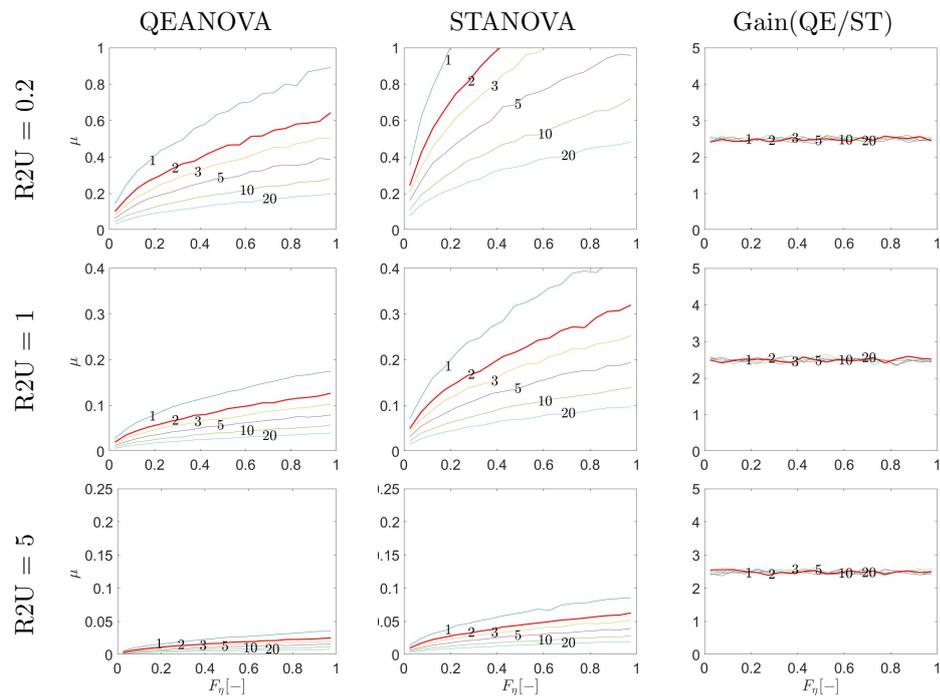


Fig. 3 Precision of QEANOVA (left) and STANOVA (middle) estimates for mean change response estimates (μ). Gain in precision between STANOVA and QEANOVA in the right panels. Results are presented for different values of the theoretical response-to-uncertainty ratio $R2U$ (top: $R2U = 0.2$, middle: $R2U = 1$, bottom: $R2U = 5$). For details see caption of Figure 2.

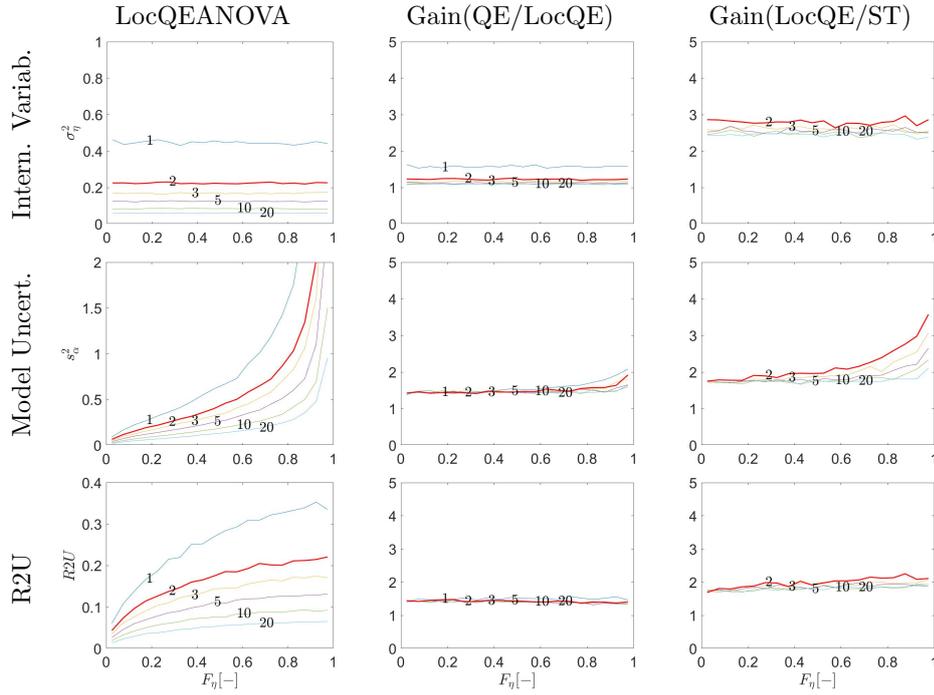


Fig. 4 Precision of uncertainty estimates obtained with the local QEANOVA approach. Results for internal variability variance (σ_{η}^2 , top), model uncertainty variance (s_{α}^2 , middle), response-to-uncertainty ratio (R2U, bottom). Left: SD values, middle: gain in precision between local QEANOVA and QEANOVA approaches (ratio between local QEANOVA SD values and QEANOVA SD values), right: gain in precision between STANOVA and local QEANOVA approaches (ratio between STANOVA SD values and local QEANOVA SD values). Results for a theoretical ratio $R2U = 1$. For details see caption of Figure 2.

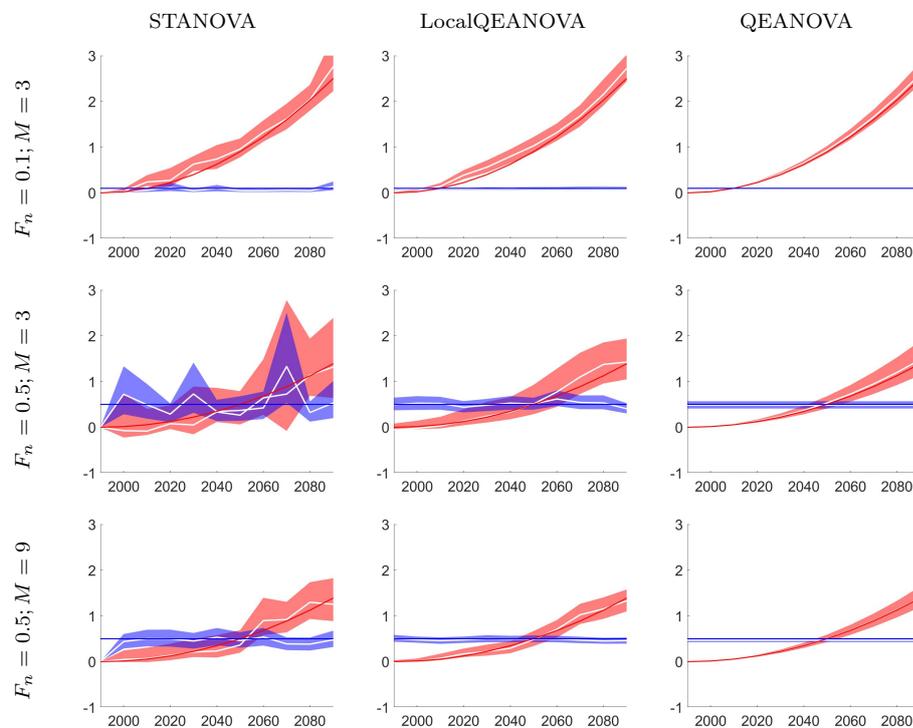


Fig. 5 Temporal variations and 90% confidence intervals of model uncertainty and internal variability estimates for the three reference datasets presented in **Fig. 1** (top: $F_\eta = 0.1$, $M = 3$; middle $F_\eta = 0.5$, $M = 3$; bottom: $F_\eta = 0.5$, $M = 9$). Left: with STANOVA; middle: with local QEANOVA where $\Omega(t)$ is composed of the 3 consecutive time steps centered around t ; right: with QEANOVA. s_α^2 in red, σ_η^2 in blue. The blue and red lines correspond to the theoretical variance components used for the generation of the reference dataset. The variance component estimates (white curves) obtained with each ANOVA approach are presented with their 90% confidence intervals.

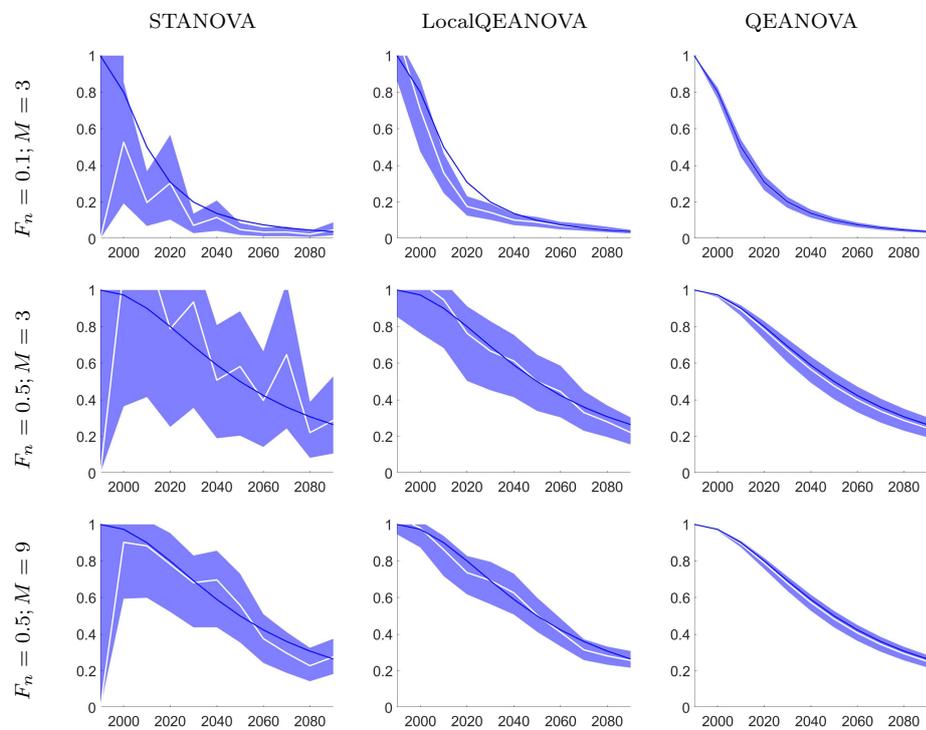


Fig. 6 Temporal variations and 90% confidence intervals of uncertainty estimates for the fractional variance due to internal variability (F_n ratio) for the three reference datasets presented in Fig. 1. Left: with STANOVA; middle: with local QEANOVA; right: with QEANOVA. The blue lines correspond to the theoretical fractional variance used for the generation of the reference dataset. See Fig. 5 for caption details.