Supplementary Material

Hans G. Wallraff :

Ratios among atmospheric trace gases together with winds imply exploitable information for bird navigation: a model elucidating experimental results

Suppl. 1 (Data base) Included chemical compounds

Table S1 enumerates the 46 airborne compounds on which the study is based. They are ordered in the sequence of gas-chromatographic retention times. The last column refers to results shown in Fig. 3 of the main text.

Table S1	46 volatile compoun	ds recorded in at	least 202 (=	90 %)	of 224 chromatogram
Table SI	40 volatile compoun	us recorded in at	least 202 (-	90 70)	of 224 chilomatograf

VOC	C index	Name	PeakArea	% finals
# 1 2 3	C03.54 * C04.00 **	i-butane n-butane	1008 2541 283	 65 ++
456	C04.74 ** C05.10 **	i-pentane isoprene	6226 885	15 40 +
7	C05.34 **	2,2-dimethylbutane	1416 51	100 + + + 20
9 10 11	C05.64 * C05.69 * C05.73 *	cyclopentane 2-methylpentane	729 1604 313	5 100 +++
12 13 14	C05.83 * C06.00 C06.14	3-methylpentane n-hexane	1157 1204 63	 55 ++
15 16 17	C06.27 * C06.45 ** C06.53	methylcyclopentane 1,1,1-trichloroethan 3 3-dimethylpentane	562 e 505 79	70 ++ 100 +++
18 19	C06.60 ** C06.67	benzene 2-methvlhexane	5329 664	30 +
20 21	C06.75 * C06.83	3-methylhexane	508 91	
22 23	C06.87 C06.92	i-octane	444 57	20 5
24 25 26	C07.00 C07.16 C07.33	n-heptane	489 25 109	10
27 28	C07.49 C07.53	to]	74 71	35 + 60 ++
29 30 31 32	C07.66 * C08.09 C08.62 C11.51	toluene	128 845 84	100 +++ 15
33 34	C14.38 C15.00	n-pentadecane	5 78	50 + 100 +++
35 36 37	C15.64 C15.71 C15.93		11 5 7	90 ++ 100 +++ 75 ++
38 39 40	C16.00 C16.22 C16.98	n-hexadecane	53 4 41	75 ++ 100 +++ 15
41 42 43 44 45 46	C17.70 C18.00 C18.11 C18.54 C19.06 C20 XX	n-octadecane	11 31 23 84 29 44	65 ++ 35 + 15 10 10
10	220.77			T 0

VOC # = running number of compound as used in the text.

C index = position on the gas-chromatographic time axis in relation to retention times of the two nearest n-alkanes (example: C06.60 = 60% time distance between n-hexane and n-heptane). 15 compounds evaluated by W&A are marked by asterisks, the six 'top compounds' selected earlier (W&A, Figs 6-8, and W2000) are labelled by two asterisks.

PeakArea = mean of 202–224 gas-chromatographic peak areas in relative units (100 = median of 46 VOCs). % finals = percentage of 20 final profiles of top *ItRuns* in which the respective compound persisted (green columns in Fig. 3f).

+ = >20 %, ++ = >50%, +++ = 100%.

It should be noted that identification and quantification of gas-chromatic peaks was hardly free of occasional errors. Such errors might slightly have contributed to noise, but could not have been frequent and systematic enough to falsify general outcomes. Decisions about inclusion or exclusion of a certain compound in some ambiguous cases were made, in view of the individual chromatograms, before the 46 VOCs were selected and used for any calculations. One of the 16 VOCs evaluated by W&A (C05.00 = n-pentane) has now been omitted owing to more rigorously applied criteria (too many unclear quantifications of chromatographic peaks due to overlaps).

Suppl. 2 (Procedural algorithms) Standardization of initial weights of VOCs

Concentrations of some VOCs are always quite high, of others always manifold lower (see W&A, Fig. 3, e.g. C6.5 vs C6.6; see also Table S1, column 'PeakArea'). Perceived signal intensity, however, may be differ from this relation. In order to give all VOCs the same starting condition, I calculated, for each VOC *i*, the arithmetic mean weight $\overline{w}_{ori,i}$ (i = 1, 2, ..., 46) of *m* original chromatographic peak areas $w_{ori,isk}$ (k = 1, 2, ..., m; m = 224 if the peak area of the respective VOC has been determined in all CGMs). Standardized weight per CGM is then $w_{st,i,k} = w_{ori,i,k}/\overline{w}_{ori,i}$ and the initial standard average weight $\overline{w}_{st,i}$ of all VOCs is 1. Proportional relations of a given individual VOC between CGMs remain unchanged.



Suppl. 3 (Iteration loops) Examples of VOC profiles

Fig. S1. Arbitrary examples of average VOC profiles from CGMs of 15 and 6 (winds from north and south) air samples collected at the central site (= home) under two different wind conditions (left) and single VOC profiles from air samples collected on two days at two distant sites (right). A: Initial status during the first iteration loop with 46 VOCs and a resulting overall homeward component $\bar{c}_{H} = -0.034$. B: Final status of the last iteration loop (*ItStep* 10 000) of the best *ItRun* with 16 remaining VOCs and $\bar{c}_{H} = +0.624$ (see Fig. 2). Note the different scales of the ordinate.

Suppl. 4 (Phase A) Determination of wind-related VOC ratios at the home site

The relative signal weight of each VOC in percent of the sum of all VOCs included in a CGM fitting in a given half-circle sector was modified by a factor $f_i = \cos (\beta_i - \omega_s) \cdot v_i$, where β_i is the direction of the wind, $\omega_s (s = 1, 2, ..., 12)$ the mid-sector bearing, and v_i velocity of wind in km/h (i = 1, 2, ..., n; n = number of CGMs obtained with winds arriving within the respective sector). Fig. S2 illustrates how the weighting factors f_i were determined. For each included VOC the modified values under n wind conditions were added up. Final VOC ratios per wind sector were then computed as percentages of the sum per VOC of the total sum of all VOC sums. They are independent of the number of participating CGMs which varied per sector between 6 and 15.



Fig. S2. Schematic example illustrating the determination of wind components within a half-circle sector including winds arriving from a directional range of $\omega_s = 300^\circ \pm 90^\circ$ (see text).

Suppl. 5 (Phases A and B) Vector calculations

In one case of phase A application and in all applications of phase B of each iteration loop mean vectors from a sample of either *n* angles α_i or vectors (angles α_i + vector lengths d_i) were calculated. As a result, in either case, a mean compass vector with direction $\overline{\alpha}_N$ and length \overline{a}_N (0° = 360° = north) or a mean home-related vector $\overline{\alpha}_H$, length \overline{a}_H (0° = homeward), was determined (positive angles always clockwise) according to

$$\bar{\alpha} = \arctan \frac{\sum_{i=1}^{n} d_i \sin \alpha_i}{\sum_{i=1}^{n} d_i \cos \alpha_i} \qquad \qquad \bar{a} = \frac{\sqrt{\left(\sum_{i=1}^{n} d_i \sin \alpha_i\right)^2 + \left(\sum_{i=1}^{n} d_i \cos \alpha_i\right)^2}}{n}$$

If only directions were involved, d_i is always 1 or omitted. In detail, applications were as follows:

- Phase A, last paragraph: equation inputs $\omega_s \to \alpha_i$, $w_{s,i} \to d_i$ (s = 1, 2, ..., 12 wind sectors, i = 1, 2, ..., n VOCs); outputs per VOC number $\overline{\omega}_i$ and $\overline{a} \to \overline{r}_i$.
- Phase B, Mode 1: inputs: α_H per CGM \rightarrow_i (*i* = 1,2, ...*n*; *n* = number of included CGMs), d_i omitted or = 1; outputs: $\bar{\alpha} \rightarrow \bar{\alpha}_H, \bar{\alpha} \rightarrow \bar{\alpha}_H$.

Phase B, Mode 2:

First application (see Suppl. 6, item 4): inputs: $\omega_s \rightarrow \alpha_i$, $\bar{d}_{P,s} \rightarrow d_i$ (s = i = 1, 2, ..., n = 12 home-wind sectors, P = CGM at current position); outputs: $\bar{\alpha} \rightarrow \bar{\alpha}_{P,N}$, $\bar{\alpha} \rightarrow \bar{\alpha}_P$.

Second application (see Suppl. 6, item 5): inputs: $\bar{\alpha}_{P,H} \rightarrow \alpha_i$ (*i* = 1,2,..., *n* = number of included CGMs), d_i omitted or = 1; outputs: $\bar{\alpha} \rightarrow \bar{\alpha}_H$, $\bar{a} \rightarrow \bar{a}_H$.

Phase B, Mode 3:

First application (see Suppl. 12, item 7): inputs: $\overline{\omega}_i \to \alpha_i$, $d_{P,i} \to d_i$ (i = 1,2,..., n = number of included VOCs, P = CGM at current position); outputs: $\overline{\alpha} \to \overline{\alpha}_{P,N}$, $\overline{a} \to \overline{a}_P$.

Second application (see Suppl. 12, item 8): inputs: $\bar{\alpha}_{P,H} \rightarrow \alpha_i$ (*i* = 1,2,..., *n* = number of included CGMs), d_i omitted or = 1; outputs: $\bar{\alpha} \rightarrow \bar{\alpha}_H$, $\bar{a} \rightarrow \bar{a}_H$.

Suppl. 6 (Phase B) Mode 2 algorithm, detailed description

The determination of the overall homeward component \bar{c}_{μ} within each iteration loop operates as follows:

- (1) Dealing with a certain CGM obtained at a certain position *P*, we determine for each included VOC *i* the difference $d_{P,i} = w_{P,i} \overline{w}_{H,i}$ (*i* = 1, 2, ..., *n*; *n* = number of VOCs), where *w* is the relative stimulus weight (in percent of the sum of *n* weights) at the current position *P* and, on average, at the home site *H*. Additionally, we determine $d_{s,i} = w_{s,i} \overline{w}_{H,i}$, where *s* refers to one of the 12 wind directions measured at home.
- (2) If the sign of $d_{P,i}$ coincides with that of $d_{s,i}$, final $d_{P,i}$ becomes positive irrespective of its original sign. If the signs are different, final $d_{P,i}$ becomes negative.
- (3) Per home-wind sector *s*, the *n* $d_{P,i}$ values are added up and averaged to $\bar{d}_{P,s}$. The higher this value is, the clearer deviates the profile of signals at position *P* from the average home profile in a way corresponding to equivalent deviations obtained with winds at home coming from direction ω_s . The lower (most negative) $\bar{d}_{P,s}$ is, the more are the deviations at variance.
- (4) Finally for this CGM gained at position *P*, a mean vector is calculated from the 12 sector directions ω_s of winds at home and related 12 scalar values $\bar{d}_{P,s}$. The direction of the resulting vector gives the VP's assumed compass direction from home, its reversal (±180°) the initial compass direction of flight $\bar{a}_{P,N}$ and its angular deviation $\bar{\alpha}_{P,H}$ from the direction towards home (see Suppl. 5). Also obtained vector length \bar{a}_P is not used for further calculations.
- (5) Finally for a larger set of CGMs gained at many positions around home, a second-order mean vector is calculated either from all directions $\bar{\alpha}_{P,H}$ alone or from the complete first-order vectors including length \bar{a}_P as well. If not otherwise stated, the first alternative (directions only) was applied resulting in an overall mean deviation from home, $\bar{\alpha}_H$, vector length \bar{a}_H referring to the unit circle with radius 1 (Suppl. 5), and the overall homeward component $\bar{c}_H = \bar{a}_H \cos \bar{\alpha}_H$.

Suppl. 7 (Phase C) Optimized weighting

At the beginning of each iteration run, standardized average weight of each VOC *i* is $\overline{w}_{st,i} = 1$ (see Suppl. 2) which is now renamed as $\overline{w}_{old,i} = \overline{w}_{st,i}$. In the first iteration step, this value is modified by adding a random quantity e_i within the range of a Gauss distribution with $\sigma = \pm 0.2$, thereby creating $\overline{w}_{new,i} = \overline{w}_{old,i} + e_i$. Thus, about two thirds of the 46 new values $\overline{w}_{new,i}$ are between 0.8 and 1.2, the rest with decreasing frequency somewhat lower or higher. Then we calculate a factor $f_i = \overline{w}_{new,i} / \overline{w}_{old,i}$. All single weights w_i in each of the 224 CGMs are then multiplied by this factor f_i , so that the proportions of a given VOC *i*, as varying *among* the CGMs, remain unchanged, whereas the proportions of the 46 different VOCs within each CGM are altered.

In the following iteration loops, it depends on the result of the previous step whether $\overline{w}_{old,i}$ of all *n* VOCs remain as before or whether they are changed according to $\overline{w}_{old,i} = \overline{w}_{new,i}$. This change occurs if $\overline{c}_{H,new} > \overline{c}_{H,old}$. Newly created random values e_i within the range of $\sigma = \pm 0.2$ are again added to or subtracted from $\overline{w}_{old,i}$, thus new values $\overline{w}_{new,i}$ and f_i produced, and so on in all consecutive iteration steps. As soon as a certain $\overline{w}_{old,i}$ falls below zero, the respective VOC *i* is cancelled from the list of included compounds and *n* accordingly reduced.

Suppl. 8 (Orientation at 90-200 km from home) Range of resulting homeward components

Fig. S3 shows the ranges of \bar{c}_H levels reached by the three modes of iterations. With Mode 2, one of 500 *ItRuns* ended up with $\bar{c}_H > 0.600$, 50 = 10 % with $\bar{c}_H > 0.564$. In addition to these 500, about 6000 more *ItRuns* were started of which many were stopped before reaching 10^4 *ItSteps*, because progress of \bar{c}_H increase (cf. Fig. 1) made a final achievement of $\bar{c}_H > 0.5$ unlikely (though not definitely impossible). At the end, a series of 20 top *ItRuns* was available with $\bar{c}_H > 0.600$.



Fig. S3. Homeward components \bar{c}_{H} finally achieved by three *ItSeries* consisting of 500 or 100 *ItRuns*, respectively, each of them including 10⁴ *ItSteps*. M1, M2, M3 = algorithm Mode 1, 2, 3. Dotted lines below zero indicate starting positions with the three modes. Left dotted vertical line separates uppermost 10%.

Suppl. 9 (Orientation 90-200 km) Comparison with real pigeons



Fig. S4. Frequencies of angular deviations from a direct homeward course as achieved with the top 20 *ltRuns* (with $\bar{c}_H > 0.6$). Thin curves show percentages of 118 vectors per *ltRun*, which are averaged in the bold curve. The blue curve shows corresponding results obtained with real pigeons (mean of nine samples; from Wallraff 2000, sources given there).

Fig. S4 shows frequency distributions of angular deviations of initial bearings from homeward. On average, 61 % of the bearings deviated less than 45° to the left or right, 88 % were in the home-oriented semicircle (random expectations 25 and 50 %). Note, however, that the indicated comparison with real pigeons in this graph is not completely adequate. The blue curve summarizes nine fairly heterogeneous data samples obtained with differently experienced pigeons from seven home lofts displaced over varying distances. It shows, at least, that the performance level of the model is in a realistic range.

Suppl. 10 (Orientation 90-200 km) Effects of VOC selection

Fig. S5 gives results of secondary *ItRuns* which did not start with the whole sample of 46 VOCs, but with reduced samples containing only those compounds that had survived in a certain percentage of the 20 primary top ItRuns (e.g. 10% = 29 compounds persisting in at least 2 of 20 primary *ItRuns*, 50% = 16 compounds present in 10 or more final profiles, 100% = 7 compounds present in all 20 final profiles; see green columns in Fig. 3f).



Fig. S5. Results of secondary top 10% ItRuns starting with a sample of VOCs selected according to the frequency of their persistence in the 20 primary top *ItRuns* reaching \bar{c}_H >0.600. Black open symbols: Starting \bar{c}_H level of iteration loops. Black filled symbols: Mean final \bar{c}_H level (±99% confidence range) of uppermost 10 of 100 *ItRuns* after 10⁴ *ItSteps*. Green symbols (right ordinate): number of included VOCs at start (open circles) and at the end of iterations (triangles, range covered by the 10 *ItRuns*). 'ALL' refers to the 20 primary top *ItRuns* starting with the total of 46 VOCs.

Home-wind directions $\overline{\omega}_i$ enhancing specific VOCs are distributed all around (Fig. S6, circles). In most cases, neither the particular direction nor the degree of the directional influence (expressed by vector length r_i ; see loop phase A) is very variable in the course of an *ItRun* during which the number of VOCs decreases and ratios among VOCs change. Nevertheless, both types of change do occur, change of direction and change of the degree of wind dependence (see dark blue columns and triangles showing changes during one *ItRun* as an example).



Fig. S6 Open circles (right ordinate, duplicated) show the upwind direction $\overline{\omega}_i$, open columns (left ordinate) the vector length \overline{r}_i indicating the degree to which wind from that direction causes an increasing ratio of the respective VOC at the beginning of each *ltRun*. If the respective VOC survived in at least 10 (= 50%) of the 20 *ltRuns*, it is drawn green. Dark-blue filled triangles and narrow columns give the 17 corresponding final values of *ltRun* rank 5 (see Fig. 3e) as an example.

Suppl. 11 (Statistical validation) Random probabilities

Fig. S7 presents results of tests performed with randomly allocated data in comparison with the realistically arranged original data. The black dots represent outcomes obtained with the latter. Each of the 100 symbols gives the mean of the three uppermost \bar{c}_H values of an *ItSeries* of 30 *ItRuns*, each *ItRun* consisting of 4000 *ItSteps*. The computational procedure and the data used were identical to those that had led to the results shown Fig. S3 (curve M2, left of the dotted 10 % line) with the difference that the ItRuns were restricted to 4000 (instead of 10⁴) ItSteps. As the initial arrangement of data was identical during all 100 *ItSeries*, it was solely a matter of chance where between the lowest and highest value (0.548 – 0.591) the outcome of a particular *ItSeries* was located. The starting conditions of the calculations leading to the coloured dots were, in contrast, not equal, not even within the same colour. Either the VOC profiles obtained at the distant sites were, in each *ItSeries*, in different ways randomly linked to the wind conditions measured there on different days (dark red symbols). Each of the coloured dots roughly corresponds to only one of the black control dots. As it represents only one *ItSeries* under the same condition, however, an individually corresponding black partner cannot be determined; it could be closer to the left or to the right end of the chain of black dots.



Fig. S7 Means of the uppermost three \bar{c}_{H} values of 50 – 200 *ItSeries*, as described in the text, in descending order. Black symbols (100 *ItSeries*): Results obtained with the originally arranged data. Blue filled (50 *ItSeries*): Air samples distributed to the peripheral sites at random (two per site remaining together). Blue open (50 *ItSeries*): Air samples remaining fixed on their axes, but axes re-arranged at random. Dark red (200 *ItSeries*): Linkage between air samples and winds at the central place re-arranged at random. Small reddish (three 50-*ItSeries*): the same, but winds shifted by +90°, -90° and 180°.

Fig. S7 makes clear that spatially randomized positioning of VOC profiles never led to \bar{c}_H values reaching the range achieved with the natural distribution. The filled blue circles indicate results obtained with VOC data attributed to the 96 sampling sites distributed completely at random, each circle originating from another random pattern. The open blue circles are from distributions in which the eight original crosses along which air samples were collected (Fig. 1 in W&A) remained intact, but changed their angular positions among each other and rotated, additionally, in 90° steps at random. These results are better than the complete-random results, but still clearly weaker than those achieved with the original data. Obviously, the achievable level of homeward directedness of the computed vectors depends on the degree of regularity in the proportional spatial distributions of trace gases as expressed in ratio gradients.

Random coupling of VOC profiles obtained at the centre with concurrently measured wind direction led most frequently to lower levels of homeward orientation as well, but decline against the control results was not as drastic as with spatial randomization (dark red in Fig. S7). 194 out of 200 triples of top \bar{c}_H values per *ItSeries* are also below the control range, but five are inside and one above. The less clear-cut results against spatial randomization do not necessarily mean that dependency of VOC patterns on wind direction is less clear-cut than dependency on position. The difference may merely be a consequence of available sample sizes. For correlation with winds at the central site, only 21 data sets (CGMs) had been obtained with wind velocity ≥ 4 km/h, whereas spatial randomization included 64 peripheral sites and 118 evaluated data sets. Thus, in the case of wind, noise is much less compensated and similarities on a random basis are easier achievable.

The wind-correlated data have been inspected in more detail. Each dark-red dot in Fig. S7 represents the top 10% \bar{c}_H values achieved with one *ItSeries* using a certain random combination of wind and VOC conditions. Additional *ItSeries* computed with the same combination provide other results (in analogy to the chain of black dots obtained with the original data). Fig. S8 shows such additional \bar{c}_H values resulting from those random combinations that had produced the uppermost 15 dark-red dots in Fig. S7. It is obvious that one out of 200 random combinations produced outstandingly better results than all others, the original true combinations (black dots) included. Three random combinations provided results largely within the original range (one median slightly above, two medians slightly below the original). The remaining 196 (= 98 %) random-based \bar{c}_H chains range clearly below the level achieved with the natural combinations.



Fig. S8 Grey zone and black symbols are identical to those in Fig. S7, scale of ordinate magnified. Small dark-red symbols replicate the left part of the large dark-red symbols in Fig. S7. The other line-connected symbols, distinguished by colour, show results (corresponding to the black original) from 10 - 50 *ItSeries* based on those random couplings between winds and VOC profiles at the central site which had produced the 15 uppermost results drawn dark-red, here marked by surrounding small squares. Each small square is connected by a thin dotted line to a larger square which shows the position of the Fig. S7 *ItSeries* within a sequence of *ItSeries* obtained under identical random conditions.

Suppl. 12 (Homing to various sites)

As a basis for correlations, the 500 *ItRuns* were used whose \bar{c}_H outcomes under Mode 3 are shown in Fig. S3 (M3). Those 28 \bar{c}_H values were selected that came closest to steps at intervals of 0.010 (0.540, 0.530, 0.520,, 0.290, 0.280, 0.260; $\bar{c}_H \approx 0.270$ does not exist). Each such \bar{c}_H had been created by application of a particularly weighted VOC profile (cf. Fig.3, a - e).

The samples of data points in the graphs Ab - Cb of Fig. 7 have been created according to the following sequence of decisions and calculations:

- (1) Choose one of the site-selection schemes Aa, Ba or Ca.
- (2) Choose one of the 28 selected primary \overline{c}_H values and the appendant VOC profile = 'profile (2)'.
- (3) Choose one of 16 peripheral goal sites ('H' in Fig. 7) symmetrically distributed at a certain distance around the centre.
- (4) Multiply each of the 46 standardized VOC values obtained from the two air samples collected there by the corresponding value of profile (2). The means of the two resulting values add up to the current home-site profile with a reduced number of VOCs (*n*) according to profile (2).
- (5) Choose one of the relevant starting sites (drawn green or red in Fig. 7).
- (6) Multiply each of the 46 standardized VOC values obtained from one of the two air samples collected there by the corresponding value of profile (2). The result is the current starting-site profile in analogy to (4).
- (7) According to the Mode 3 description, compute the difference $d_{P,i}$ between starting-site and home-site value of VOC *i* (*i*=1,2,..,*n*) and create a vector consisting of length $d_{P,i}$ and wind-gradient direction $\overline{\omega}_i$ which is adopted, together with the weighted VOC profile, from the original central-site calculations in loop phase A. From *n* such vectors, calculate the VMP's flight directions $\overline{\alpha}_{P,N}$ and $\overline{\alpha}_{P,H}$ at the particular starting site. Save the bearing $\overline{\alpha}_{P,H}$ and the achieved number of such bearings, *N*. If only the first of two chromatograms originating from this starting site has been used, go back to item (6) and treat the other one.
- (8) Go back to item (5) until all relevant starting sites with all N relevant VOC profiles are incorporated. If accomplished, calculate a 2nd-order mean vector with home-related direction $\bar{\alpha}_H$, length \bar{a}_H and homeward component \bar{c}_H . Save these values and reset N to 0.
- (9) Go back to item (3) until all 16 relevant goal/home sites are incorporated. If accomplished, calculate the mean of 16 \bar{c}_H values, thus creating a third-order homeward component. Now insert a data point in the relevant graph of Fig. 7 (Ab Cb) with this homeward component on the ordinate and \bar{c}_H of item (2) on the abscissa.
- (10) Go back to item (2) until all 28 selected *ItRuns* with descending values are implemented. Now determine the related correlation coefficient. If this is done, go back to item (1) until all geometrical constellations are considered.